Data Management Plan

December 2008

Abstract

Every clinical study should have a data management plan to ensure and document adherence to good clinical data management practices for all phases of a study. This chapter identifies data management plan components and provides information on acceptable criteria for various sections of the plan. Although the clinical data manager will not personally perform all the tasks or prepare all the sections of the data management plan described in this chapter, the data manager should ensure all of these tasks and sections are completed according to good clinical data management practices.

Introduction

Although a study protocol contains the overall clinical plan for a study, separate plans, such as a data management plan (DMP) or statistical analysis plan, should be created for other key areas of emphasis within a study. Before data collection begins, all clinical studies should have a DMP in place to document the relevant conventions for that particular study. A well-designed DMP provides a road map of how to handle data under any foreseeable circumstances and establishes processes for how to deal with unforeseen issues.

The optimal end result for a clinical data manager is to provide a study database that is accurate, secure, reliable, and ready for analysis. Many people will be involved in handling data throughout the course of a clinical study, so it is imperative that all parties refer to the DMP for a consistent approach to the processes and guidelines for conducting data management activities.

The DMP is an auditable document often asked for by regulatory inspectors and should be written in a manner that is professional and of high quality.

During an audit, the inspectors may also seek to ascertain the degree to which the project team adheres to the processes described in the DMP.

Scope

Although style and format may differ from one organization to the next, this chapter gives a broad overview of the components and processes that make up a DMP. Whether the DMP document itself contains all of the elements or refers the reader to other study documents for further detail, this chapter provides the data manager with the minimal components that should be addressed within the overall study documentation.

Minimum Standards

- Complete a draft of the DMP prior to enrollment of the first subject.
- Ensure the DMP supports compliance with applicable regulations and oversight agencies.
- Identify and define the personnel and roles involved with decision making, data collection, data handling and data quality control.
- Ensure data management processes are described and defined from study initiation until database closeout.

Best Practices

- Develop the DMP in collaboration with all stakeholders to ensure that all responsible parties understand and will follow the processes and guidelines put forth in the DMP from study initiation to database closeout.
- Develop and maintain a DMP template for the organization that ensures consistency and standardization across all projects.
- Ensure the DMP for each study is kept current, including proper versioning, and that all responsible parties are aware of and agree to the current content.
- Ensure that an approved, signed version of the DMP is completed prior to starting on the work it describes. The job functions or titles that must

approve and sign the DMP may vary between organizations and depending on the type of study.

Purpose of the DMP

The DMP documents the processes and procedures employed by organizations to promote consistent, efficient and effective data management practices for each individual study. A primary goal of the DMP is to communicate to all stakeholders the necessary knowledge to create and maintain a high-quality database ready for analysis. The DMP serves as the authoritative resource, documenting data management practices and decisions that are agreed to at study initiation. The DMP should comply with all applicable regulatory guidelines (e.g., FDA, ICH, GCP) or local laws of the country; as well as the standard operating procedures (SOPs) of the organization. The DMP should also address any procedural or protocol updates that are made during conduct of the study.

Creation and Maintenance

For each new study, clinical data management (CDM) personnel should compose a detailed DMP based on the protocol, work scope, contract, analysis plans, dataflows, case report forms (CRFs), other supporting documents, and data management standards and practices. The entire DMP should be drafted and approved by all responsible parties prior to commencement of the work it describes. The clinical data manager should ensure the DMP is kept current, including proper version control, and that all parties involved agree with the content. Upon conclusion of the study, the DMP should be archived with all other pertinent study documentation.

The DMP should be created during the setup phase of each study and should contain information relating to all aspects of data management activities to be performed. The DMP should be considered a living document throughout the life cycle of a study, capturing any changes impacting data management made to the protocol or processes being used. The DMP must be uniquely identifiable, carry such identification on each page (e.g., study code/title) and be subject to version control. Each version should be documented and include date, author, reason for version change and an individual version identifier.

Organization of a DMP

The organization, structure and order of topics presented in a DMP may differ between organizations. The following sections of this chapter cover the components that typically make up a DMP. Some of these components may be contained in documents referenced by the DMP rather than being detailed within the DMP itself. In either case, these components should be addressed within the overall study documentation.

Approval Page

The approval page should detail the study identifiers and primary reviewers or signatories. The signature line(s) should include dates of approval. For companies allowing e-signatures, company requirements for e-signatures must be followed. The work detailed in the DMP should not begin until signatures are present from all relevant stakeholders.

Protocol Summary

Many organizations may include a short synopsis of the study protocol, visit schedule, or critical data analysis variables within the DMP. This summary or synopsis gives a broad overview of the protocol and should refer the reader to the full protocol for more detailed information. Just as a DMP typically omits a full version of the study protocol, the DMP also typically omits a record of each protocol change or amendment. However, in some organizations the DMP may maintain a list of major protocol revisions and associated version numbers.

Dictionary and Coding Management

The DMP should indicate which medical coding dictionaries (e.g., MedDRA, WHO Drug, SNOMED) and versions of the dictionaries will be used for the study. The DMP should reference documents providing instructions for how to handle dictionary updates or changes and define all quality control measures, validation methods, and user acceptance testing (UAT) for the dictionary. The DMP should also describe any auto-encoding or study-specific conventions used, as well as listing appropriate SOPs. Some examples of different types of coding include medication coding (prior/concurrent), adverse event (AE) coding, medical history coding, non-AE medical event coding (primarily for observational studies), and physical exam coding.

Please refer to the "Dictionary Management" and "Safety Data Management and Reporting" chapters of *Good Clinical Data Management Practices* for more information, including recommendations, minimum standards and best practices.

Definitions and Acronyms

The DMP should include a list of acronyms that are specific to the protocol and DMP. Acronyms can be very helpful, but if their meaning is obscure they can become a hindrance. The DMP should also provide definitions of terms that may be misinterpreted or misunderstood.

Personnel/Role Identification/Training

The DMP should specify key personnel with roles and responsibilities for the associated protocol and study activities, or the DMP may refer to external documents or related SOPs containing this information. The DMP should also refer to documents related to project-specific training requirements for various roles and functions.

Timelines

The timeline included in the DMP or document referenced by the DMP lists expected completion targets for all deliverables. For example, database validation could be targeted for completion a specified number of weeks from the time the protocol is finalized.

Some organizations may have more detailed timelines, including more interim, internal activities; other organizations may have less detail, only tracking critical path activities. Timelines may also vary based on parameters of the study, such as between paper-based studies and those utilizing electronic data capture (EDC). Following are examples of milestones that may appear on a study timeline and be detailed in a DMP or associated documentation.

- Protocol finalization
- CRF development
- Database design and UAT

- Data validation, programming and UAT
- First patient first visit
- Last patient last visit
- Last CRF/data element received/entered
- Last query/discrepancy form received/completed
- Final SAE reconciliation completed
- Medical coding completed and approved
- Interim analysis, when applicable
- Database audit
- Database lock
- Study data and documentation archiving

Case Report Forms

According to ICH E6, a CRF is defined as "A printed, optical, or electronic document designed to record all of the protocol-required information to be reported to the sponsor on each trial subject."[1] The following are specific areas that should be elucidated within the DMP or other documents referenced by the DMP.

- CRF design—Provide a detailed description of the CRF design process or refer to the organization's SOPs relating to CRF design and development.
- CRF instructions—Include general guidelines for CRF completion as well as protocol-specific guidelines.
- CRF changes—Describe the process for managing changes to the CRF design or reference the organization's appropriate SOP. Changes to CRFs may also involve metadata changes, which should be governed by the same SOPs or one SOP designed specifically for the description of that process.

Database Design, Creation and Maintenance

The DMP should refer to an in-depth study-specific database validation plan and include a brief description of how the database is created and maintained, a description of the system that is holding the data and table naming conventions. Title 21, Code of Federal Regulations Part 11 (21 CFR Part 11) mandates that procedures and controls be in place to ensure appropriate control of and access to documentation as well as revision and change control procedures to maintain an audit trail of modifications to documentation.[2]

Database Archive

The DMP should outline specific information regarding the organization's procedures for archiving the electronic records.

Database Roles and Privileges

The DMP should include profiles for available database roles within the system being used to support the study. Assign privileges to roles based upon the duties performed in the study. At a minimum, the roles should be listed or a reference should be made to a document where the roles are described. A detailed description of each role and the associated privileges is optimal.

Database Security

The DMP should describe or refer to documents that describe the security of networked equipment and servers as well as security features of the electronic records within the clinical data management system (CDMS). The database security section of the DMP should also address:

- Maintenance of user roles and access—Describe the procedure(s) or refer to the organization's SOPs for defining, creating and maintaining system user roles and access. This description should include the process for revoking access.

- Database backup—Outline database backup procedures, frequency and routines. The disaster recovery plan and database backup SOPs should also be referenced in this section.

Data Entry and Processing

The DMP or referenced documents should define data entry and processing plans. Data handling guidelines provide details of general study rules, which may cover acceptable abbreviations, symbol conversions, incomplete dates, illegible text, allowed changes and self-evident corrections. Ensure the DMP or DMP-referenced documents provide clear guidance for all of the following areas where applicable:

- Data entry guidelines—Describe proper entry of various data elements, proper handling of data anomalies, proper handling of missing data, and proper notation of self-evident changes. A comprehensive list of accepted abbreviations as well as symbols and their translations should be included in the guidelines. This list may be presented using a table within the DMP or by referring to a separate document.

- Data discrepancy conventions—Develop guidelines to provide consistency in classifying and processing data discrepancies.

- Data receipt—Specify the type of receipt (paper CRF or EDC), the expected frequency of data receipt, and how data receipt will be tracked. This also refers to data transfers from any third-party vendors.

- Data processing—Describe how data will be processed upon receipt at the organization (either electronic or paper-based data).

- Data entry—Indicate who will perform data entry and whether single or double entry will be used.

- Self-evident corrections—Specify the criteria for self-evident corrections and identify authorized data management personnel who will make these corrections to the data as necessary. A self-evident correction is a change to data or resolution of a query that can easily and obviously be made on the basis of other existing information on the CRF without sending a query to the investigative site. The most common self-evident corrections are obvious spelling errors. Self-evident corrections, like all other data changes, must be clearly documented and audited via the audit trail within the organization's database system. A list of approved self-evident corrections must be included in the DMP or exist in a separate document to be attached or referenced. Ensure the investigators associated with the study are in agreement with the self-evident correction process and that

the method of additional documentation (e.g., generation of reports for sign off) is thoroughly described. Self-evident corrections might not be applicable to all data management systems and types of data (e.g., source records).

- Data reconciliation—Provide details about the data fields and external databases requiring reconciliation per the study protocol.

- Database lock—Provide details defining the criteria for database lock, who will be responsible for database lock, and processes that will be employed in locking the database. Refer to the organization's SOPs on study closeout as well. The DMP may also contain or refer to other SOPs for the unlocking and relocking processes if required.

Please refer to the "Data Entry and Data Processing" chapter of *Good Clinical Data Management Practices* for more information, including recommendations, minimum standards and best practices.

Data Validation and UAT

The DMP should define validation test procedures to ensure integrity of study-specific components such as programming/algorithms, data entry/EDC screens, online logic/data-checking routines, security, backups, and archiving. If the DMP does not contain this information, it should reference a separate validation plan and/or validation and UAT SOPs. Please refer to the "Database Validation, Programming, and Standards" chapter of *Good Clinical Data Management Practices* for more information, including recommendations, minimum standards and best practices.

In addition to ensuring data entered into the database are complete, correct, allowable, valid, and consistent, other types of data quality checks may be applied. Once these checks have been identified, appropriate and verified programs are created to help identify discrepancies. All derivation and validation procedures may be fully tested and documented in the DMP or a referenced validation plan.

Data quality checks include:

- Manual review specifications—Describe all types of manual review specifications. Some aspects of these checks may be identified electronically depending on the features of the CDMS utilized. Other manual reviews (e.g., medical history, adverse events, concomitant medications reports, header information) may be generated via the CDMS; however, reviews of these data are usually accomplished through visual inspection.

- Discrepancy management—Describe the query process in detail, including how data clarification forms for paper studies or electronic queries for EDC studies are to be raised, tracked and handled when resolved, the annotation of any working copy CRFs and the documentation to be filed or retained. If different statuses are used for discrepancies, they should be defined.

- Electronic data discrepancy management—Define and describe processes to resolve electronic data discrepancies for the dataset or module being checked. These processes should include presentation of information which may include the CRF module, variable description, name of the edit check, processes for the use of test cases, a description of the edit check, an output message that would translate to a data query, other associated variables in the case of cross-checking data, and processes for documentation of these testing and validation activities.

SAE Data Reconciliation

The DMP should describe or refer to documents that describe the protocol specific SAE reconciliation plan.

Quality Assurance/Control Processes

The DMP should define quality assurance (QA) plans and quality control (QC) process steps. As defined by ICH E6, quality control is “the operational techniques and activities undertaken within the quality assurance system to verify that the requirements for quality of the study-related activities have been fulfilled.”[1]

Because studies of differing levels of regulatory importance are undertaken, occasionally a study will not be carried out within the established quality system. If this is the case, the study may not follow any SOPs in place or may only follow some of them. Complete an SOP compliance checklist indicating which SOPs are applicable to the study. Document in the comments section of the SOP compliance checklist any justification for opting out of all or part of the SOPs.

The DMP should address:

- Level of checks—Decide on and specify the required level of checking to be performed before data collection begins. Depending on the type and regulatory importance of a study, different levels of checking may be implemented. For example, an observational study may need only a minimal level of checking, whereas a highly regulated drug or device study requires a much more stringent level of QC checking.

- Frequency of quality control checks—Specify the frequency of QC checks in the DMP. According to ICH E6, "Quality control should be applied to each stage of data handling to ensure that all data are reliable and have been processed correctly."[1]

- QC check documentation processes—Define the means by which QC checks are documented and how this documentation is maintained throughout the course of the study.

For more information about quality assurance and quality control, please refer to the chapters entitled "Assuring Data Quality" and "Measuring Data Quality."

External Data Transfers

For external data transfers, the DMP should describe the data type (e.g., safety lab data), the entity providing or receiving the data and any applicable agreements, the format, the frequency of transfers, and contact information for all those involved with the data transfer. Good practice is to have an established data transfer plan and to conduct a test data transfer prior to the need for a live transfer.

Specific data transfer details may include, but are not limited to, the following:

- Variable/element specifications
- Format of transfer (SAS® datasets, ASCII files, XML files, etc.)
- Method of transfer (encrypted e-mail, FTP, CD, DVD, etc.)
- Recipient of data (site, sponsor, data safety monitoring board (DSMB), statisticians, etc.)
- Frequency of transfer
- Quality control/validation steps performed to maintain integrity

The DMP should describe procedures used for collecting and handling laboratory data. If data comes from any combination of central labs, core labs, local labs, or specialty labs, there should be a short section differentiating between procedures for collecting and handling different types of lab data. Include or reference guidelines on how to transport, track, clean and report upon the various types of laboratory data.

Please refer to the "External Data Transfers" chapter of *Good Clinical Data Management Practices* for more information, including recommendations, minimum standards and best practices.

Audit Plans

The DMP should either define the on-site audit and corrective action plans, or refer to those documents that do cover these processes. All interim and final study database audits should also be defined. As defined by ICH E6, quality assurance is "all those planned and systematic actions that are established to ensure that the study is performed and the data are generated, documented (recorded), and reported in compliance with GCP and the applicable regulatory requirements(s)."[1]

The DMP should also define how often during the course of a study QA will take place. Please refer to the "Assuring Data Quality" chapter of *Good Clinical Data Management Practices* for more information, including recommendations, minimum standards and best practices.

Metrics

The DMP should include the metrics that will be used for the study. Please refer to the "Metrics for Clinical Trials" chapter of *Good Clinical Data Management Practices* for a list of commonly used metrics.

Reports

The DMP should include a list of available reports for dissemination throughout the life of the study. For each report, specify the target audience, content of the report, level of detail provided, date of data extraction, frequency of generation and the mechanism used for distribution (e.g., e-mail, posting electronically). Additions and deletions to the report listing may occur throughout the life of the study and should be updated in subsequent versions of the DMP.

Communications

The DMP should describe the types of communications or correspondence used in the study. Detail where records of these communications (whether paper or electronic) will reside, as well as any associated archiving requirements. Document how communications will be conducted and outline regularly scheduled communications. Indicate where to find communications after the fact. For example, if there is a particular form that must be signed and faxed, an auditor could see this in the DMP and not waste time searching through e-mails.

The DMP should include information on:

- Frequency of communication—Describe how frequency of communication may vary throughout the course of a study. For example, the communication may be more frequent in the setup and early stages of the study, then become less frequent as the study progresses. During the study conduct, many communications may be limited to study maintenance issues. During the closeout and lock portion of a study, communication frequency may increase again. Although most studies will have communication variability of this nature, specify any regularly scheduled communications in the DMP.

- Medium (e.g., face-to-face vs. conference call vs. Web conference)—Describe the estimated amount and timing of meetings, as well as which medium will be used. Try to schedule one or two face-to-face meetings (or more depending on length) during the course of a study. Web conferences are a good medium to share information in real time, such as when collaborating with the study team to edit a document or modify a process.

- Escalation process—Determine if issues need to be moved up the chain of command, when is it appropriate, and which parties should be involved.

Other Processes

Every study is unique to some degree, and there may be processes within a particular study that have not been covered within this chapter. If a study involves other processes, they should always be described in detail somewhere within the protocol or DMP. Some additional processes that may need to be examined include the following:

- DSMB requirements—Describe any requirements pertaining to DSMB meetings that may occur during the course of the study. What preparation is expected to be performed prior to these meetings? Will this preparation be treated as a lock in regards to having all data clean and reported upon prior to the meeting? Will the DSMB be focusing on a sample of the data or the complete data set?

- Business rules—Specify business rules that may have an impact on data handling or data integrity in the DMP. For example, regularly scheduled IT maintenance that limits server access, organization-wide observed holidays or an anticipated change of address during the course of the study may affect data handling.

- Flowcharts and forms (e.g., CRFs, source documents, adjudication and query forms)—Include applicable flowcharts or sample forms that may be required by your organization.

- Problems and resolutions—Document the process of identifying, discussing, resolving and filing problems arising and resolved during the study.

- Change control processes—Evaluate if other change control processes may be encountered during the course of the study and describe them in the DMP.
- Blind data review specifications—Describe the expectation from data management if a blind data review will be conducted.
- Archival and record retention process—Describe when and how the archival process occurs. The processes described revolve around current organizational and governmental regulations. There are certain requirements that must be met according to applicable regulatory and/or sponsor requirements.[1] Document the record retention timeframe and communicate this timeframe to site personnel.

Recommended Standard Operating Procedures

- CRF Design and Development
- Database Design and Testing
- Data Management and Systems Roles and Responsibilities
- Coding Dictionary Management
- System Security
- Change Control
- Data Entry
- Internal Data Handling
- External Data Handling
- Data Cleaning
- SAE Data Reconciliation
- Quality Control
- Database Lock and Unlock

- Study Data and Documentation Archival

References

1. International Conference on Harmonisation. *Guideline for Good Clinical Practice, E6*: Good Clinical Practice; Consolidated Guideline. Geneva, Switzerland: Author; 1996.

2. *Code of Federal Regulations, Title 21, Volume 1, Part 11*. Washington, DC: US Government Printing Office; 1998.

Further Reading

Guidance for Industry: Computerized Systems Used in Clinical Investigations. U.S. Department of Health and Human Services, Food and Drug Administration (FDA), Office of the Commissioner (OC); May 2007.

Chapter Revision History

Publication Date	Comments
December 2008	Initial publication.