FDA/MHRA Good Clinical Practice Workshop

Day 2- October 24, 2018

Session 2 - Case Studies on BE/GCP

Date Management Case Study (participant copy)

<u>Section 1 (15 Minutes)</u>

A multicenter clinical trial was commenced by the sponsor using non-carbon copy two-part paper (NCR paper) case report forms (CRFs). Each CRF visit page had final sign-off by the site's research nurse. The sponsor's clinical research associates (CRAs) obtained the top page of the original CRF from the investigator, leaving the bottom copy at the site, and brought the originals back to the sponsor. The sponsor's Clinical Data Analysts (CDAs) entered the data into the database management system located on the sponsor's server.

The CDA amended the database where they felt it appropriate based on previous data received by the site and, afterwards, emailed the investigator to inform them of the change they had made. Data queries from data verification and validation checks were also emailed to the investigator. The investigator replied via email and the CDA filed the email as evidence of investigator approval of the change, informed the investigator to amend their copy of the CRF and the CDA would do the same to the original.

Later, the sponsor decided to reduce workload and improve the process. The investigators were given the option of remote access to the database and the sites entered data directly into the database rather than the CDAs doing this activity, though the CDAs continued to do so for some sites. There were no updates to the database computer system validation documentation. Queries by the CDA were dealt with via email in the same way as before because there was no functionality within the database to raise queries.

After the trial ended, a regulatory agency inspection of the sponsor occurred. It was noticed that the paper CRF data did not always reflect what was seen in the database. Often, only the query response value was entered into the database and not corrected on the original paper CRF. It was also noticed from the audit trail that the data changes were made by the CDA prior to either sending the email to the site with data queries or before a response from the investigator was received. The investigator had not signed any "authorized changes" document that allowed the CDA to make changes to specific data.

What GCP and data integrity issues are you able to identify?

When, if ever, would it be appropriate to make changes to the data without the prior approval of the site investigator?

Section 2 (10 Minutes)

At an investigator site of the trial inspected by the regulatory agency, it was seen that not all the required data was being captured in the patient medical record/chart. Some data was entered directly onto the paper CRF. The protocol did not define any data to be directly entered into the CRF. Additionally, the paper CRFs were not always amended by the investigator after a data query based on the emails received from the CDA. When the investigator was given access to the database to enter the data, they were unsure what to do with the paper CRFs and sometimes filled these in as well.

The primary endpoint of the trial relied upon data entered into a paper study diary by the patient concerning cold/flu infection symptoms. None of the trial diaries were at the investigator site as the clinical investigator was told by the sponsor that they were not required to be copied or retained. Instead, the diary was to be used as a "memory aide" with the investigator writing the information into the patient's study chart. This was not always being done. Review of one patient's study record showed that a visit by the patient took place for an infection on Tuesday, December 5, 2017, but there were no symptoms documented in the study record/chart. The database audit trail showed the December 5th visit with three symptoms (rhinitis, cough, conjunctivitis) entered into the database by the investigator on Monday, January 8, 2018 at 2:20 pm; the data had then been amended on Friday January 12, 2018 at 11:15 am. Found in the medical record/chart was an email of Monday, January 8, 2018 at 2:25 pm that was sent to the patient from the investigator asking for details of the symptoms. The patient's response to the email, giving three symptoms, was received Friday January 12, 2018 at 10:20 am. The data in the database for one of the symptoms (conjunctivitis) was not consistent with the email (itching eye/ocular pruritus).

There was no source data location agreement.

What further GCP and data integrity issues are you able to identify with respect to source data?

What is the overall impact on the clinical trial results and are there any corrective and preventative actions that could be taken?

<u>Section 3 (10 Minutes)</u>

The sponsor decided for later trials to utilize independent vendors for electronic case report forms (CRFs) and electronic diaries (eDiaries). The eCRFs are transmitted via the Internet through a firewall into the data management software system. Pre-programmed edit checks provide automatic feedback when data are entered incorrectly. Review of one of these trials at the vendor site identified the following:

- The contract stated that the eCRFs, which had been hosted by the vendor, would be provided on a compact disc (CD) as portable document format (pdf) files to the sponsor for onward distribution to the investigator sites. Metadata was not covered by the contract or procedures. The subjects were given password protected access to the database to enter their eDiary information directly. The eDiary data was provided to the sponsor. Changes to the eDiary data during the trial were made by the vendor upon request of the investigator, provided the sponsor authorized the change.

- In preparation for study database lock, the sponsor asked the vendor to conduct a blinded review of the primary endpoint (i.e., the response rate) based on the eDiary. One site that enrolled 60 subjects had an almost 100% response rate, which was an extreme outlier.

- In addition, a review of the audit trail of the eCRF database showed that investigator sites were entering data after the database lock. The data extracted for the statistical analyses used in the clinical study report was done before this additional data had been entered.

Can there be data clarification or queries for patient-reported outcome (PRO) such as subject diary data?

Are there any GCP or data integrity issues apparent in the new processes?

What additional actions would you do to investigate the high efficacy results at the one site?

GCP REFERENCES

All clinical information shall be recorded, handled and stored in such a way that it can be accurately reported, interpreted and verified, while the confidentiality of records of the trial subjects remains protected. **2005/28/EC Article 5**

Any change or correction to a CRF should be dated, initialed, and explained (if necessary) and should not obscure the original entry (i.e. an audit trail should be maintained); this applies to both written and electronic changes or corrections. Sponsors should provide guidance to investigators and/or the investigators' designated representatives on making such corrections. Sponsors should have written procedures to assure that changes or corrections in CRFs made by sponsor's designated representatives are documented, are necessary, and are endorsed by the investigator. The investigator should retain records of the changes and corrections. **CPMP/ICH/135/95: "Note for Guidance on Good Clinical Practice" (ICH GCP) 4.9.3**

The necessary procedures to secure the quality of every aspect of the trial shall be complied with **2005/28/EC**

The investigator/institution should maintain adequate and accurate source documents and trial records that include all pertinent observations on each of the site's trial subjects. Source data should be attributable, legible, contemporaneous, original, accurate, and complete. Changes to source data should be traceable, should not obscure the original entry, and should be explained if necessary (e.g., via an audit trail). **ICH 4.9.0**

The investigator should maintain the original source document or a certified copy. **ICH GCP 2.11, 5.15.1**

Source data should only be modified with the knowledge or approval of the investigator. **ICH GCP 4.9.3, 4.9.4 and chapter 8**

The sponsor should not have exclusive control of a source document. **ICH GCP 8.3.13**

Source Data: 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 trial. Source data are contained in source documents (original records or certified copies). **ICH GCP 1.51**

Source Documents: Original documents, data, and records (e.g., 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). **ICH GCP 1.52**

Certified Copy: A copy (irrespective of the type of media used) of the original record that has been verified (i.e., by a dated signature or by generation through a validated process) to have the same information, including data that describe the context, content, and structure, as the original. **ICH GCP 1.63**

When using electronic trial data handling and/or remote electronic trial data systems, the sponsor should: Ensure and document that the electronic data processing system(s) conforms to the sponsor's established requirements for completeness, accuracy, reliability, and consistent intended performance (i.e. validation). **ICH GCP 5.5.3**

The sponsor should base their approach to validation of such systems on a risk assessment that takes into consideration the intended use of the system and the potential of the system to affect human subject protection and reliability of trial results. **ICH GCP 5.5.3**

The SOPs should cover system setup, installation, and use. The SOPs should describe system validation and functionality testing, data collection and handling, system maintenance, system security measures, change control, data backup, recovery, contingency planning, and decommissioning. The responsibilities of the sponsor, investigator, and other parties with respect to the use of these computerized systems should be clear, and the users should be provided with training in their use. **ICH GCP 5.5.3**

Data reported on the CRF, that are derived from source documents, should be consistent with the source documents or the discrepancies should be explained. **ICH GCP 4.9.2**

Protocol Content: The identification of any data to be recorded directly on the CRFs (i.e. no prior written or electronic record of data), and to be considered to be source data. **ICH GCP 6.4.9**

The purposes of trial monitoring are to verify that: The reported trial data are accurate, complete, and verifiable from source documents. Checking the accuracy and completeness of the CRF entries, source documents and other trial-related records against each other. The monitor specifically should verify that: The data required by the protocol are reported accurately on the CRFs and are consistent with the source documents. The reported trial data are accurate, complete, and verifiable from source documents. **ICH GCP 5.18.1 1(b) 5.18.4 (a) and (m)**

The sponsor should ensure that the trials are adequately monitored. **ICH GCP 5.18.3**

Monitor's Responsibilities: Communicating deviations from the protocol, SOPs, GCP, and the applicable regulatory requirements to the investigator and taking appropriate action designed to prevent recurrence of the detected deviations. **ICH GCP 5.18.4 (q)**

The investigator should maintain the original source document or a certified copy. Sponsor should not have exclusive control of a source document. **ICH GCP 2.11, 5.15.1, 8.3.13**

If noncompliance that significantly affects or has the potential to significantly affect human subject protection or reliability of trial results is discovered, the sponsor should perform a root cause analysis and implement appropriate corrective and preventive actions. **ICH 5.20.1**

The sponsor and investigator/institution should maintain a record of the location(s) of their respective essential documents including source documents. The storage system used during the trial and for archiving (irrespective of the type of media used) should provide for document identification, version history, search, and retrieval. The sponsor should ensure that the investigator has control of and continuous access to the CRF data reported to the sponsor. The sponsor should not have exclusive control of those data. When a copy is used to replace an original document (e.g., source documents, CRF), the copy should fulfill the requirements for certified copies. The investigator/institution should have control of all essential documents and records generated by the investigator/institution before, during, and after the trial. **ICH 8.1**

The instrument should be created in a controlled manner to ensure that it conforms to the protocol and is validated. In addition, appropriate change control as part of ongoing validation is needed, in cases where protocol amendments require changes to the instrument. The fundamentals of clinical research include that patient rights, safety and well-being are the most important considerations and the integrity of the reported data must be confirmable. To this end all data generated in a clinical trial relevant to patient care must be

made available to the investigator at all times during and after the trial and all data held by the sponsor that has been generated in a clinical trial should be verifiable to a copy not held (or that has been held) by the sponsor. The requirements above are not met if data are captured in an electronic system and the data are stored on a central server under the sole control of the sponsor. This is because the investigator does not hold an independent copy of the data and therefore the sponsor has exclusive control of the data. A number of attributes are considered of universal importance to source data and the records that hold those data. These include that the data and records are: Accurate, Legible, Contemporaneous, Original, Attributable, Complete, Consistent, Enduring and Available when needed. Any transfer from paper to electronic CRF should be subject to quality control. The protocol should identify any data to be recorded directly into the CRFs that is considered to be source data. I and the level of control should be justified. **EMA Reflection paper on expectations for electronic source data and data transcribed to electronic data collection tools in clinical trials, 01 Aug 2010**

FDA Regulations and Guidances

21 CFR 11.10 – "Procedures and controls [for electronic records systems] shall include the following: (a) Validation of systems to ensure accuracy, reliability, consistent intended performance, and the ability to discern invalid or altered records."

21 CFR 11.10 – "Procedures and controls [for electronic records systems] shall include the following: (d) Limiting system access to authorized individuals.

21 CFR 11.10 – "Procedures and controls [for electronic records systems] shall include the following: (e) Use of secure, computer-generated, time-stamped audit trails to independently record the date and time of operator entries and actions that create, modify, or delete electronic records. Record changes shall not obscure previously recorded information. Such audit trail documentation shall be retained for a period at least as long as that required for the subject electronic records and shall be available for agency review and copying.

Guidance for Industry Part 11, Electronic Records; Electronic Signatures – Scope and Application. "…your decision to validate computerized systems, and the extent of the validation, [should] take into account the impact the systems have on your ability to meet predicate rule requirements. You should also consider the impact those systems might have on the accuracy, reliability, integrity, availability, and

authenticity of required records and signatures. Even if there is no predicate rule requirement to validate a system, in some instances it may still be important to validate the system.

Guidance for Industry Use of Electronic Records and Electronic Signatures in Clinical Investigations Under 21 CFR Part 11 – Questions and Answers "…processes should be in place to control changes to the electronic system and evaluate the extent of revalidation that the changes may necessitate. When changes are made to the electronic system (e.g., system and software upgrades, including security and performance patches, equipment or component replacement, or new instrumentation), sponsors and other regulated entities should evaluate the effect of the changes and validate the changes using a risk-based approach…Major changes may require additional re-validation and critical changes could trigger a re-validation of the entire system."
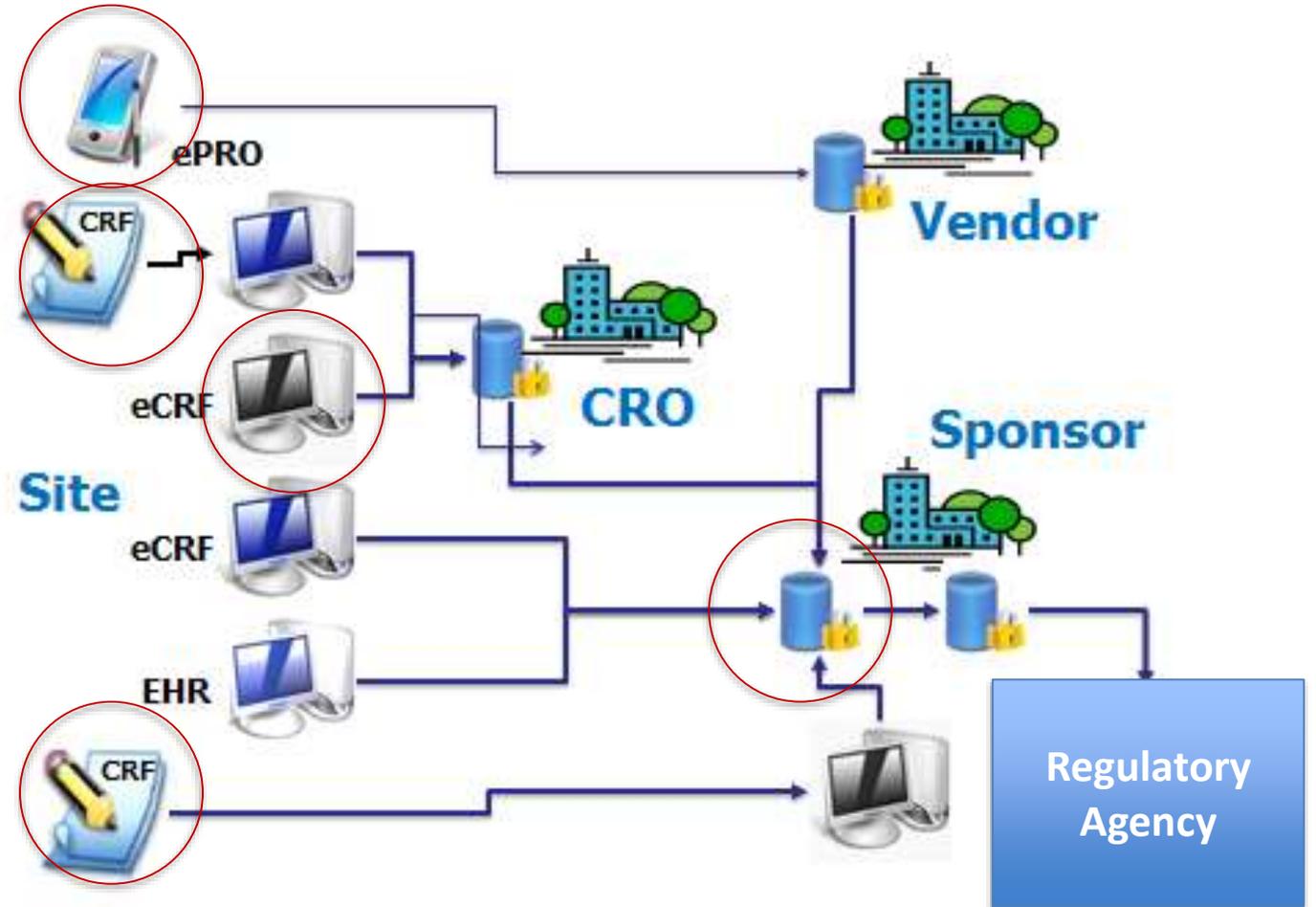
INTRODUCTION

Case Study – 3 Sections

- You should have already received the texts (and hopefully have read them!)
- The case is based on situations found on Good Clinical Practice (GCP) inspections.
- Appoint a table group leader who will be prepared to provide feedback.
- Review the description of the situation together and discuss the questions.
- Ask the facilitators for assistance, if needed.
- We will take feedback/discuss as a whole group after <u>each</u> section.

Overview: Concepts of Data Management

- Data Ownership
- Data Collection
- Data Processing/Coding
- Data Storage
- Data QA
- Data Protection
- Data Retention
- Data Analysis
- Data Sharing
- Data Reporting

SECTION 1

This scenario involves paper and electronic case report forms (CRFs), data entry, and making corrections to the data in a multi-center clinical trial. It also involves a change to functionality of a computer system during the trial.

- **What GCP and data integrity issues are you able to identify?**
- **When, if ever, would it be appropriate to make changes to the data without a prior approval of the clinical investigator?**

SECTION 1

What GCP and data integrity issues are you able to identify?

- Data were not controlled appropriately.
  - There was no formal data clarification process/form used for data correction.
  - Clinical data analyst (CDA) made changes prior to investigator's approval.
  - The investigator did not sign "authorized changes" document.
  - There was no signing off the CRFs by the investigator with the changes made.
    - Only the investigator should have final sign-off of the CRFs.
  - There was no documentation of changes to the CRFs, leading to discrepancies in values between the database and the CRFs.

SECTION 1 (cont.)

- There was no audit trail for changes.
  - The database reflected only the value entered by the CDA. It did not record a change shown in the paper CRF.
- There was no agreement between the investigator and sponsor as to which data could be amended without the investigator's approval.

SECTION 1 (cont.)

- The data entry procedure was not clear.

- Best practice is double data entry.

  - *Data entries, such as …recording of source data for inclusion of a patient in a clinical trial, should be verified by a second person, as appropriate for the intended use of this data.* (WHO GUIDANCE ON GOOD DATA AND RECORD MANAGEMENT PRACTICES)

SECTION 1 (cont.)

- System moved from being a database to being an eCRF – significant change in the functionality.

- The systems requirements for the new approach to investigator's access and data entry were <u>not</u> assessed as part of the change control process.

- New functionalities needed for the new process were not built.
  - There is no functionality for the CDA to raise queries to the investigator in a controlled and transparent manner.
  - The investigator should change the data, not the CDA.

SECTION 1 (cont.)

- There was loss of investigator control.
  - Sponsor held the investigator data on its server.
  - There was no segregation of investigator's data.
  - The CDA (sponsor) retained edit access to the investigator's entered data (same rights as the investigator).
  - Expectation is that the investigator maintains control of their data!

ADDITIONAL COMMENTS

1.

2.

3.

SECTION 1 (cont.)

When, if ever, would it be appropriate to make changes to the data without the prior approval of the site investigator?

- If the Data Validation Plan for the query management process includes:
  - Edit check specifications
  - Definitions of self-evident corrections (SECs)
  - Project-specific processing guidelines

➢ An Edit Check Specifications document should be developed and authorized by the investigator. It is recommended that the investigator approves what changes were made.

➢ It is recommended to have a library that addresses common data entry errors via standard checks.

➢ Note: Use of Data Entry Constraints and Restricted Value Sets prevents errors in the first place.

SECTION 1 (cont.)

- If an investigator is no longer available (e.g., died), procedures for change of site investigator should be completed, per Institutional Review Board/Ethics Committee policies/procedures.

- Some discrepancies will be considered as "irresolvable" and should be placed in the discrepancy database.

ADDITIONAL COMMENTS

1.

2.

3.

SECTION 2

This scenario involves direct data entry, source data, use of a subject diary, the use of an audit trail, and assessing data integrity issues.

- **What further GCP and data integrity issues are you able to identify with respect to source data?**

- **What is the overall impact on the clinical trial results and are there any corrective and preventative actions that could be taken?**

SECTION 2

What further GCP and data integrity issues are you able to identify with respect to source data?

- There is confusion over what is considered source data. There is no document to define source data location.

- Source data that is entered directly into the eCRF is not covered in the protocol.

- The CRF became a completed source document once the investigator entered the data. The change in process is not addressed as there is no change in data management activities.

SECTION 2 (cont.)

- Source data should be retained at the investigator site; however, there was loss of source, i.e. dose confirmations, patient diaries.

- Entries to the database were made a long time after the visit with no source (from memory!) and were amended by checking with the patient later.

- Changes made on the original CRF by the CDA would not reflect the investigator's copy where the investigator did not amend their copy, leading to discrepancies between the source data and the database.

ADDITIONAL COMMENTS

1.

2.

3.

SECTION 2 (cont.)

What is the overall impact on the clinical trial results and are there any corrective and preventative actions that could be taken?

- The Sponsor had the opportunity to amend the data that was out of the investigator's control.

- Reconstruction of data (at the investigator site) sent to the sponsor is impossible (investigator copy of data is a mess).

- This directly impacted the primary endpoints; therefore, the data is not reliable.

Whether to salvage the trial data will depend on the purpose of the trial.

- Is it part of a marketing authorization application? Is it a non-commercial trial that has important public health questions that need to be answered? Is it of low-importance with respect to commercial gain and public health? If important, then attempts should be made to rectify the situation.

- The extent of the issue needs to be determined. Are the investigator site issues widespread (source document confusion, failure to amend CRF, lack of source data)?

- Need to provide assurance that the investigator's source data matches the data in the CRF – this would require extensive auditing. The investigator must provide evidence that he/she agrees that their data matches the CRF.

ADDITIONAL COMMENTS

1.

2.

3.

SECTION 2 (cont.) What to Implement

- Target site visits to assess extent of issues discovered – i.e., monitor and audit to provide evidence of extent. It may be that there are a few sites where there are similar problems.

  – These sites could be excluded from the final per protocol analyses if the trial is powered sufficiently to allow this (i.e., the question can still be answered without the data from these sites).

- Review changes made in the database and by whom – ensure documented authorization by the investigator.

- Develop a source data definition document.

SECTION 2 Implement (cont.)

- There should be collection of all paper diaries.

- There should be re-validation of the system for eCRF functionality or revert to paper.

- A robust query management process should be developed (source document agreement, authorized changes agreement, training, etc.).

- There should be data review by the investigator to confirm data reported to the sponsor is consistent with source.

SECTION 2 Implement (cont.)

- A robust validation and lifecycle management process should be developed.

- There should be risk assessment, with appropriate monitoring.

- There should be additional, enhanced training.

ADDITIONAL COMMENTS

1.

2.

3.

This scenario involves the use of electronic diaries, central monitoring of data, use of an audit trail, database lock and archiving of data.

- **Can there be data clarification or queries for patient-reported outcome (PRO) data such as subject diary data?**

- **Are there any GCP or data integrity issues apparent in the new processes?**

- **What additional actions would you do to investigate the high efficacy results at the one site?**

ANSWER

SECTION 3

Can there be data clarification or queries for patient-reported outcome (PRO) data, such as subject diary data?

YES. Although patient-reported outcome (PRO) data is unusual in that it is not entered by the investigator but directly recorded or entered by the subject or patient, it can still be queried. There is a perception by some in clinical research that there is no data clarification or query process for PRO data or diary data.

SECTION 3 (cont.)

FDA's guidance *Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims* states:

- Sponsors should avoid direct PRO data transmission from the PRO data collection device to the sponsor, clinical investigator, or other third party without an electronic audit trail that documents all changes to the data after it leaves the PRO data collection device.

- The data maintained by the clinical investigator should include an audit trail to capture any changes made to the electronic PRO data at any point in time after it leaves the patient's electronic device.

ADDITIONAL COMMENTS

1.

2.

3.

Are there any GCP or data integrity issues apparent in the new processes?

- The investigator has not had continuous access to his/her data.
  - The use of electronic PRO instruments may pose a problem if direct control over source data is maintained by the sponsor or the contract research organization and not by the clinical investigator. Regulators consider the investigator to have met his or her responsibility when the investigator retains the ability to control and provide access to the records that serve as the electronic source documentation for the purpose of an inspection. The clinical trial protocol, or a separate document, should specify how the electronic PRO source data will be maintained and how the investigator will meet the regulatory requirements.

- All data must be provided to the investigator– the eDiary data was not.

SECTION 3 (cont.)

- The diary data could only be changed if approved by the sponsor.

- The Investigator does not have final control over changes to diary data to confirm source data.
  - The FDA guidance states that the sponsor should avoid the ability of any entity other than the investigator (and/or site staff designated by the investigator) to modify the source data.

SECTION 3 (cont.)

- Complete data was not archived at the site as there was no metadata.
  - Contracts with vendors must address storage and security, and must comply with GCP – i.e., ensuring that investigators' source data is not under the sole control of the sponsor. Must address archiving and what will be kept such as audit trails. Flat files that lose important data such as audit trails are not acceptable.

- There were insufficient data lock procedures.
  - Database lock should prevent any unauthorized changes.

ADDITIONAL COMMENTS

1.

2.

3.

SECTION 3 (cont.)

What additional actions would you do to investigate the high efficacy results at the one site?

- Conduct a blinded review of treatment assignment at the site to rule out any randomization error.

- Review data in a blinded manner to see if the site is an outlier for any other data.

- Retrieve the audit trails of the eDiaries to look at time and pattern of data entry and if changes were made.

SECTION 3 (cont.)

- Audit the site to look for any signs of misconduct.
  - Include documentation of the presence of real subjects.
  - Review training of staff and subjects.
  - Review how access passwords were communicated to the subjects and how they were kept confidential.
  - Confirm that the site had no knowledge of the subjects' access codes.

- Interview investigator and staff on all procedures at the site.
  - Ask them to offer an explanation for the efficacy results.

ADDITIONAL COMMENTS

1.

2.

3.

Wrap-up

- Appropriate clinical data management is a key to generation of high-quality, reliable, and statistically sound data from clinical trials.

- Data management should be considered for all stages of clinical trials; from inception to completion, including CRF designing, database designing, data-entry, data validation, discrepancy management, database locking, data transformation, and quality control procedures.

Wrap-up

- Effective methods for receiving, storing, disseminating, reporting, analyzing, and retaining regulatory data are required.

- The requirements for good data and record management apply equally to paper and electronic data.

- All changes prompted by the sponsor through queries should be captured in the audit trail.

Wrap-up

- A sponsor of clinical studies must establish an operating model to support its data needs.
  - Define decision making authorities, roles and responsibilities, and assign accountability.

- Hopefully, at the conclusion of this conference, you will go back to your office and do a gap analysis for compliance to best practice goals, taking into consideration the regulatory requirements and the guidance from regulatory agencies.

Thank you to all that contributed to the case study.

© Crown copyright 2018

About copyright
All material created by the MHRA, including materials featured within these MHRA presentation notes and delegate pack, is subject to Crown copyright protection. We control the copyright to our work (which includes all information, database rights, logos and visual images), under a delegation of authority from the Controller of Her Majesty's Stationery Office (HMSO).

The MHRA authorises you to make one free copy, by downloading to printer or to electronic, magnetic or optical storage media, of these presentations for the purposes of private research, study and reference. Any other copy or use of Crown copyright materials featured on this site, in any form or medium is subject to the prior approval of the MHRA.

Further information, including an application form for requests to reproduce our material can be found at **https://www.gov.uk/government/publications/reproduce-or-re-use-mhra-information/reproduce-or-re-use-mhra-information**

Material from other organisations
The permission to reproduce Crown copyright protected material does not extend to any material in this pack which is subject to a separate licence or is the copyright of a third party. Authorisation to reproduce such material must be obtained from the copyright holders concerned.

A Model Data Management Plan Standard Operating Procedure: Results From the DIA Clinical Data Management Community, Committee on Clinical Data Management Plan

14 authors, including:

Scott Brand
FEFA LLC

17 PUBLICATIONS   **403** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Evolving Approaches to Clinical Trial Data Quality View project

DIA
DEVELOP
INNOVATE
ADVANCE

A Model Data Management Plan Standard Operating Procedure: Results From the DIA Clinical Data Management Community, Committee on Clinical Data Management Plan

Scott Brand, PhD[1,*]**, Diana Bartlett, BA, CCDM**[2]**, Mary Farley, BA**[3]**, Miriam Fogelson, BS**[3]**, Jan Bart Hak, PhD**[4]**, Grace Hu, MS**[5]**, Olivia D. Montana, BA, CCDM**[6]**, Jan Holladay Pierre, MPH, CMQ/OE**[7]**, Johann Proeve, PhD**[8]**, Samina Qureshi, MD**[9]**, Anita Shen, MS, CCDM**[10]**, Peter Stockman, MSc**[11]**, Richard Chamberlain, PhD**[12]**, and Kristin Neff, MS**[13]

Abstract

Background: The DIA Clinical Data Management Community created a committee to develop a model standard operating procedure (SOP) for writing a data management plan. *Methods:* The goal of the committee was to develop a plan that could be used by industry and academic institutions. The model was based on contributed examples from committee members and their experiences with current practices and technologies. It is understood that as new clinical trial technology is implemented, the SOP will require modification. *Results:* The model SOP and associated templates are presented as a starting point, and each company or institution will need to modify them to meet its individual needs. *Conclusion:* The model DMP SOP produced addresses most data management issues that are present in any phase clinical trial while providing a flexible framework.

Keywords

data management plan, standard operating procedure, clinical trial, electronic data capture, case report form

Introduction

The data management plan (DMP) specifies all clinical trial data management tasks. This includes how the data are collected, quality control, data transformations, and final deliverables. The procedure for creating a DMP should be documented in a company or institution's standard operating procedure (SOP). This SOP is written to ensure consistency and compliance with good clinical practices. The difficulty in writing a DMP SOP is that the tasks defined can vary depending on the type of clinical trials that a company carries out and the resources that are available. A clinical trial that makes use of multiple electronic data collection systems (web based, smartphone, interactive voice recognition system) may require more DMP tasks than an early-phase study with few subjects that collects data on paper case report forms (CRFs). Similarly, a large company can assemble several groups to carry out multiple data-cleaning processes, while a small pharmaceutical

[1] QPS LLC, Newark, DE, USA
[2] Theravance Biopharma US Inc, South San Francisco, CA, USA
[3] KAI Research Inc, Rockville, MD, USA
[4] QPS LLC, Groningen, The Netherlands
[5] QPS Qualitix, Taipei, Taiwan, ROC
[6] SyneractHCR Inc, San Diego, CA, USA
[7] INC Research, Orlando, FL, USA
[8] Bayer Healthcare, Leverkusen, Germany
[9] PSI International Inc, Rockville, MD, USA
[10] Janssen of Johnson & Johnson, Beijing, China
[11] Merck, Sharp and Dohme, Oss, The Netherlands
[12] Extended Clinical Services LLC, Centreville, VA, USA
[13] InVivo Therapeutics Corp, Cambridge, MA, USA
*Committee chair.

Submitted 14-Jan-2015; accepted 02-Mar-2015

Corresponding Author:
Scott Brand, PhD, QPS LLC, Delaware Technology Park, 3 Innovation Way, Newark, DE 19711, USA.
Email: scott.brand@QPS.com

company or contract research organization will not have the same resources. While there are multiple templates and guides available for creating a DMP,[1-7] there is no guidance as to how to write a DMP SOP. Private pharmaceutical companies consider their SOPs proprietary information, and academic institution SOPs, while made available by their sponsors, are not geared to industry. The DIA Clinical Data Management Community (CDMC) decided that there was a need to create a model DMP SOP, which could be used as a starting point for small to midsize pharmaceutical companies and contract research organizations to create their own DMP SOPs. It is understood that no one DMP SOP will meet the needs of all studies and companies; however, the CDMC DMP committee attempted to include topics that would form the core of any DMP SOP.

Method

The DIA CDMC formed a committee of data managers, clinical researchers, clinical research associates, and quality assurance professionals. DMPs from several pharmaceutical and medical device companies and academic research institutions were reviewed. All client and personal identifiers in these example DMPs were redacted. The primary components of these DMPs were defined. Additionally, new technologies were assessed to see how they would influence clinical trial data management.

Topics in the DMP SOP were assigned to individuals who had expertise in those areas. These subject matter experts wrote the first draft of their assigned sections. These draft sections were reviewed by the entire committee and modified. When a section was considered final, it was turned over to the committee chairman. Final integration of the all the sections into one document was carried out by a smaller group. After integration, the document was sent to reviewers outside the committee. The outside reviewers had experience in data management, study design, monitoring, electronic data capture technology, and quality assurance. Modifications were based on this external review and the final document created.

When an annotation was thought to be needed, it was included with the label *Comment*.

Results

The resulting model DMP SOP is shown in the appendix and includes associated templates.

Discussion

The model SOP that is described should be considered a baseline document. It reflects current data management practices and technology. However, data management is being reshaped by new technologies,[8,9] changes in practice,[7,10,11] and the

pressure to contain clinical trial cost. Technologies are being advanced that place data collection in the hands of the subjects,[9] including the possible uses of in-home diagnostic testing.[12] The smartphone places a sophisticated device capable of clinical data collection in the hands of every subject. These systems are currently being used for patient report outcomes, but it is easy to see their potential use as a continuous monitoring device for activity, blood chemistries, and adverse event reporting. Likewise, the use of gathering data directly from the electronic medical record, specifically extracting data from an HL7 record, offers new challenges to creating a DMP.

Quality control practices are also being evaluated; 100% source data verification is being replaced by remote risk-based monitoring[13,14]; and the benefit of source data verification is itself being questioned.[13] These changes in data management will influence what is required in a DMP. This evolution in clinical trial technology and practice forces the DMP SOP to evolve. For these reasons, this model SOP will need to be reviewed regularly and modified to reflect the changes in the pharmaceutical research.

As stated, not all companies and institutions have the same resources. When writing a DMP SOP, the parties are faced with the conflict of meeting the highest level of regulatory agency[3,15] and industry expectations while being realistic about what can actually be accomplished. For example, companies with limited staff, proposing a multilevel quality control process in which various groups review certain aspects of data collection or data analysis can put a company in a position of not following its own SOP. Rather than writing a perfect SOP, it is better to write a logical DMP SOP that addresses processes in a realistic manner. In regard to the present model DMP SOP, the responsibility matrix may not need all the job titles included in the template, and if risk-based monitoring is not used, that section of the SOP can be eliminated. Similarly, low-risk studies—such as a bioequivalence studies, where there is one site entering safety data on a limited number of subjects—could use an abbreviated SOP. As stated previously, this model SOP and resulting DMPs should be modified to meet the needs of the company or institution.

When writing a DMP SOP, all stakeholders should be involved. The writing team should consist of representatives from data management, quality assurance, statistics, monitoring, and project management. This includes the people designing case report forms (CRFs) and electronic CRFs (eCRFs), creating the electronic data capture application, testing the electronic data capture application, using the CRF/eCRF, and carrying out statistical analysis (SAS programmers). Project management could represent the sites that will use the CRF/eCRF during data collection. When one group dominates the creation of the SOP, critical components are often ignored or dealt with in a limited fashion.

Appendix

Model Standard Operating Procedure for Creation of Data Management Plan

Note: This is not an SOP; it is a guide for writing the DMP SOP. The document is a mixture of instructions and sample text. It is meant to be modified to reflect a company's procedures.

1.0 Purpose

The purpose of this SOP is to provide guidance for creating a company's or institute's DMP. The ultimate goal is that the resulting DMPs meet regulatory, industry, and institutional standards.

2.0 Scope

The scope describes to what and whom the SOP applies. The SOP can list all divisions of a company or institution that the SOP affects. For example, a DMP SOP should apply, at a minimum, to

- Data managers
- Database developers
- Programmers
- Project managers
- Quality assurance staff
- Quality control staff
- Staff involved in the creation of CDISC SDTM files
- Sponsor/client who needs verification that its standards for the data life cycle are met

A general statement can be made that the DMP SOP applies to all {company/institution} persons involved in the creation of data collection systems, databases, data cleaning, data quality control, and quality assurance.

A larger company with multiple facilities may want to enter in the names of specific facilities or subsidiaries to which the SOP is applicable.

3.0 Responsibilities

All functional and any associated support staff (including contracted employees) will comply with this SOP as it pertains to their job functions. It is the responsibility of company {name of company or institution} management to ensure that the appropriate staff is trained on the DMP SOP before performing the activities described herein. It is the responsibility of the staff and company {name of company or institution} management to ensure that procedures described in this SOP are followed.

It is the responsibility of the {global} *functional head* to ensure the SOP is revised as needed, at a minimum the SOP should be reviewed and/or revised biannually.

Quality Assurance is responsible for monitoring compliance with the SOP and/or providing recommendations to management for any noncompliance observed.

The {global} *functional head* and the QA {department, representative, etc} are accountable for ensuring that process changes in other company {name of company or institution} SOPs affecting the DMP SOP are being synchronized.

4.0 Definitions

Comment: These are abbreviations and terms used in this model DMP SOP; this section should be modified per the terms used in your company's or institution's SOP.

AE: Adverse events

Back-end edits: Edits run on the database in batch mode, usually written in SAS (also known as *programmed edits*)

CDASH: Clinical Data Acquisition Standards Harmonization

CDISC: Clinical Data Interchange Standards Consortium

CDMS: Clinical data management system; this term is sometimes used interchangeably with EDC. It creates a clinical trial database, eCRFs, real-time data edits, and it can generate reports.

CRF: Case report form

COTS: Commercial off-the-shelf software

CRO: Contract research organization

DM: Data manager

DMP: Data management plan

eCRF: Electronic case report form

EDC: Electronic data capture system; see CDMS

Edit: A test of a data value to verify whether it is accurate or conflicts with another data value

EHR: Electronic health record

EMR: Electronic medical record

ePRO: Electronic patient-reported outcomes

Front-end edits: Automated edits built into the EDC that are triggered as the data are entered (also known as *real-time edits*)

IVRS: Interactive voice recognition system

ODM: Operational data model

PDF: Portable document format

PI: Principal investigator

Programmed edits: Edits run on the database in batch mode, usually written in SAS (also known as *back-end edits*)

QA: Quality assurance

QC: Quality control

Query: A request for data clarification

Real-time edits: Automated edits built into the EDC that are triggered as the data are entered (also known as *front-end edits*)

SAE: Serious adverse event
SAP: Statistical analysis plan
SDTM: Study data tabulation model

5.0 Procedures

5.1 Clinical Trial Database

5.1.1 Required Information Needed Prior to Creation of a Clinical Database

Before a DMP can be written, certain information must be available. This should be specified in the SOP and can include, but is not limited to, the following:

- The database standards on a global, therapeutic area, or project level must be available, including database structure, formats, code lists. Comment: It is recommended to make use of CDASH (Clinical Data Acquisition Standards Harmonization) standards when possible.
- The final or close-to-final study protocol must be available.
- The planned visit structure with dosing, exams, and tests to be recorded at specific time points.
- Sample CRFs from client or CRO
- External data sources specified (ECG, central lab, ePRO data, pharmacokinetic data, bioanalytic data) and if any data conversion is needed (data from local labs to common unit lab analytes)
- Data to be generated in house (eg, safety and drug coding, visit or subject validity, reasons for invalidity)
- Statistical analysis plan (final or near-to-final plan)

5.1.2 Creation of a Clinical Trial Database for Remote Data Capture or Centralized Data Entry Using an EDC System

Comment: These are suggested items to cover under this heading. The language is general so that it applies to any EDC. A company may assign tasks to different job categories. Responsibilities can be documented as shown in Figure A1.

- A study database will be built using an EDC system. The DM or designee is responsible for creating the database, setting up an operation environment, and managing subjects, CRFs, events, database, and users.
- The data capture rules (real-time data edits) and data parameters will be implemented by the DM using the functionality of the EDC system. All real-time data edits will be documented.
- Database development will include script-driven user acceptance testing.
  - Test scripts should undergo independent review before they are executed.
  - All completed test scripts and testing results must be maintained as part of study documentation.

- Blank CRFs are generated by the EDC and annotated.
  - Visual QC of CRFs is carried out to verify all data that needs to be collected to analyze primary and secondary endpoints as documented in the SAP are present.
  - Review of the CRFs by client or sponsor
- Access to all fields in eCRFs is verified, and navigation to fields is in logical order.
- Studies using hardcopy CRFs should undergo the same processes, with the exception of verifying access to fields in an eCRF.

5.1.3 Creation of a Clinical Trial Database Not Based on an EDC System

Comment: This would apply to a database that serves as a repository for clinical trial data uploaded from external sources (EHR, EMR, laboratory data, etc).

- A study database will be built using database design tools or a COTS system (a database associated with an EDC system could also be utilized without the data being collected through eCRFs).
- Tables and associated variables would be based on the protocol and organization of the source data.
- Data quality edits could be implemented as the following:
  - Built into a data loader (either a database importation utility or a separate system)
  - Stored procedures within the database
  - Programmed SAS edits
- Testing based on verification that the data stored in the database matches the source data
  - Full comparison
  - Statistical sampling

5.2 Access to EDC/Database System

Access to the EDC system or clinical database must be tracked with either an electronic or manual log. How the tracking is done should be described in the SOP. The information collected in the log should include the person getting access, the type of access (read, write, form approval only), the date when access was granted, the type of access, and the date that access was terminated. An example of differing access rights would be as follows:

- Data entry: Read/write privileges
- Monitor access: Read access and limited write access (ability to flag data problems and leave comments)
- PI access: Read access and write access limited to approving forms
- Project manager: Read-only access
- Investigator: Write access limited to form approval and read access

a. Task descriptions can vary per study; the list used is an example.

Task Description[a]	Data Manager	Project Manager	Statistician	Statistical Programmer	Database Administer	Other...	Comment
Create Data Management Plan (DMP)	☐	☐	☐	☐	☐	☐	
-- Statistical Analysis Plan --							
Prepare	☐	☐	☐	☐	☐	☐	
Approve	☐	☐	☐	☐	☐	☐	
-- Database Development --							
Create database and CRFs in EDC system	☐	☐	☐	☐	☐	☐	
Create User Acceptance Testing (UAT) Script	☐	☐	☐	☐	☐	☐	
Verify that data needed to determine end points are being collected in eCRFs	☐	☐	☐	☐	☐	☐	
Review annotated CRF	☐	☐	☐	☐	☐	☐	
-- Security --							
Create, control and manage user account access	☐	☐	☐	☐	☐	☐	
-- CRF Workflow --							
Data entry	☐	☐	☐	☐	☐	☐	
Ascertain error rate for individual entering data)	☐	☐	☐	☐	☐	☐	
-- Data Cleaning --							
Approve Edit Specification section of DMP	☐	☐	☐	☐	☐	☐	
Implement real time edit checks	☐	☐	☐	☐	☐	☐	
Create batch edit checks external to EDC	☐	☐	☐	☐	☐	☐	
Create manual queries	☐	☐	☐	☐	☐	☐	
Review of batch edit check reports	☐	☐	☐	☐	☐	☐	
Track resolution of queries	☐	☐	☐	☐	☐	☐	
-- Data Coding --							
Perform coding procedure	☐	☐	☐	☐	☐	☐	
QC coding	☐	☐	☐	☐	☐	☐	
--Uploading Data --							
Prepare documentation and electronic data for uploading	☐	☐	☐	☐	☐	☐	
Execute upload	☐	☐	☐	☐	☐	☐	
Verify uploaded data in database	☐	☐	☐	☐	☐	☐	
-- Database Lock --							
Provide SAE tracking log and protocol deviation log to DM for reconciliation	☐	☐	☐	☐	☐	☐	
Request approval, lock the study in EDC system	☐	☐	☐	☐	☐	☐	
Approvals obtained	☐	☐	☐	☐	☐	☐	
Extract all data from the system into SAS datasets or XML files	☐	☐	☐	☐	☐	☐	
-- CDISC SDTM --							
Verify validation of SDTM domain structure	☐	☐	☐	☐	☐	☐	
Verify data in SDTM is consistent with EDC database	☐	☐	☐	☐	☐	☐	
Validate the format of the SDTM domains	☐	☐	☐	☐	☐	☐	

Figure A1. Data management responsibilities.

5.3 Entering and Uploading Data

Specify how the data will be entered. The following methods could be used:

5.3.1 Single or Double Data Entry

- Training materials for data entry
- Specify if only certain fields require double entry and identify these fields

5.3.2 Loading Electronic Files

- How are the electronic files generated
- How transfer will take place
- Will the files be cumulative or incremental

Figure A2. Real-time (front-end) edits in electronic data capture.

5.3.3 EHR Data Extracted

- Source of EHR data (hospitals, managed care organization, clinic)
- Format of EHR data (HL7, SNOMED)
- How is it reformatted and loaded into study database

5.3.4 New Data Sources

Advances in technology will present new methods for collecting and entering data. These will need to be addressed in future versions of the model DMP SOP. For example, direct collection of physical examination and diagnostic data using attachments to smartphones are being tested for clinical trials. The data from such a device could be part of an external file or possibly uploaded into the clinical database at the time the data are collected.

5.4 External Files

External files sent from a laboratory, clinic, or vendor must be in an agreed-upon format. A document specifying the format should be created and signed by a representative of the file creator and the receiver of the file. This document will take the form of a record layout designating the variables or description of the data items and the data type (character, numeric, length, etc). The process and required documentation regarding external file transfer can be described in a separate SOP and referenced in the DMP.

External files can be kept separate from the clinical trial database or merged with it. External files can originate from multiple sources, such as various ePRO and mobile systems (eg, texted questions and responses sent and received using a smartphone, pain scale administered via the web, IVRS system to verify that medication was taken, and diagnostic results from mobile devices).

5.5 Data Cleaning

Data cleaning for a study using an EDC system usually occurs at two levels:

1. Real-time edits (front-end edits) built into the EDC system and executed during data entry

2. Programmed edits (back-end edits) that are run on the data currently stored in the database (these edits are most commonly run using SAS)

A study that loads data from external files runs edits either during the loading process or as back-end edits.

All edits (front-end or back-end) must be documented as part of the data management plan. Figure A2 shows a format for documenting front-end edits; Figure A3 shows a format for documenting back-end edits.

All front-end edits should be tested as part of the user acceptance testing of the EDC system and back-end edits tested by the DM or designee. All completed testing documentation must be maintained. The testing process should be addressed in a separate testing SOP.

When data cleaning is the responsibility of a CRO, the documentation of the edits (completed templates Figures A2 and A3) should be approved by the client or sponsor prior to implementation (Figure A4).

While data quality is addressed by the real-time and programmed edits, these edits trap data errors that were documented in only the edit specifications of the DMP and are created before the study is started. Complex studies can have unforeseen data errors not included in the edit specifications. These unforeseen data problems may be found only by reviewing actual collected data using methods that survey the entire study population, such as listings, SAS frequencies, and cross-tabulations. These general approaches to data quality can be listed separately as part of the programmed back-end edits.

5.6 Coding and Reconciliation

Coding

Comment: This section briefly explains that a company's or institute's common coding philosophy or approach is used for all studies conducted and for all SAEs reported to drug safety. The actual coding philosophy or approach can be in a separate document (medical coding SOP) and referenced in the DMP SOP.

			a. If an error message identifies a protocol deviation, this should be stated in the message			
Edit Number	CRF/File	Visit/Event	Fields	Specification	Message[a]	Action
1	External lab file, AE	Lab data reconciliation with AE	WBC	Lab value for XXX that is greater than ### subject must have an AE record	Lab value indicates AE but subject has no adverse event record	
2	InformC Medical History	Visit 1	NO_CANC, CANCER	Inclusion criteria is subjects cannot have cancer, medical history indicates subject had cancer	Inclusion criteria does not allow for subjects to have cancer in medical history. Possible protocol deviation	
3	Medication	Visit 1	DoseDT	Date received first dose cannot be before Med History Date (MedHisDT)	Date of dose 1 prior to Med History Date	

Figure A3. Programmed (back-end) edits.

a. If an error message identifies a protocol deviation, this should be stated in the message				
Data Management Plan Approvals	Name	Title	Signature	Date
Sponsor representative				
Project/data manager				
Management				

Figure A4. Edit spec approvals.

An example of documenting the coding approach is the following:

Coding is done on a by-text basis, not on a by-patient basis. Only the final text submitted for coding will be considered when a code is assigned. Therefore, only the information in the text will be assessed for assigning codes; background or historical information on the patient must not influence coding.

Each type of data (eg, AE, MedHist, Conmed, Surgeries, ECG, microorganisms) should document which glossaries and version were used.

When data from the clinical trial database is entered into a drug safety database for coding, the data between the two systems should be reconciled to verify the data in both systems are identical. The processes and frequency of reconciliation should be specified.

A matrix (Figure A5) can also be used to specify the sequence of events in the coding process.

5.7 Query Generation and Processing

Comment: Query generation and processing can depend on the technology used and the type of study being carried out. Some EDC systems generate queries automatically based on unresolved data edits (data rules). These edits are manually reviewed and directed to the person who is best able to resolve them. The responsible party then is notified that a query needs to be addressed, and he or she can then either resolve the query or respond back to the DM with a question or comment. During most clinical trials, queries are also generated manually during the QC process or monitoring visits. Other systems rely on manual query processing or a combination of automated and manual.

The following information is needed in the DMP:

- Methods of query generation (automated within EDC, manual, other)
- Responsible person (job title) for generating queries during the data-cleaning/QC process
- Query-tracking system used
- How the queries are distributed
  - Through queues in an EDC
  - Manually/automatically emailed to sites or sponsor
  - Notification to responsible person that a query has been generated
  - Other
- To whom the queries are distributed during the query resolution process
  - DM at site
  - Central DM at CRO or pharmaceutical company
  - Reviewer who determines query is resolved correctly
  - Other
- Document who is responsible for resolving queries and correcting the data in the database (eg, only DMs at the site, senior DM at CRO)
- Query reports generated

Task	Timing / Frequency	Comments
DM will set up the coding plan	Within ## weeks of First subject first visit.	
Manual queries resulting from medical coding will be posted in EDC system by designated DM	Within ## week after the receipt of query from Medical Coder	*[This is an example, alternatively queries can be loaded to EDC]*
Medical Coding queries will be loaded to EDC automatically in regular intervals shortly after coding	On a daily basis	*[This is an example, alternatively queries can be posted manually]*
DM will ensure that schedule of last coding run is in line with planned database lock *(in case coding run is not triggered automatically by pre-defined events)*	Before database lock	
Listings are generated to compare SAEs reported to the medical monitor and AEs listed as *serious* on the eCRF and entered in the clinical trial database. Fields that are compared may include verbartim term, coded term, onset date, end date, causality, etc. The comparisons are reviewed and discrepancies are reported directly to the site and sponsor by data queries. The {sponsor or CRO} works with the sites and the medical monitor to reconcile any discrepancies between the Safety Database and the clinical trial database.	This depends on the length of the study. Reconciliation for a long duration study could occur monthly, and for a short study it could take place prior to database lock	

Figure A5. Coding process matrix.

- Status of queries by site
- Business intelligence that signals trends in type of queries being generated and/or the source of queries
  - Monitors during source data verification
  - Generated by batch edits
  - QC review of clinical listings
  - EDC system
  - AE that could not be coded
- Previsit query reports for monitors
- Other

Who is responsible for verifying all queries have been resolved or closed prior to database lock?

5.8 Risk-Based Monitoring

Specific parameters (data elements) in the clinical trial database can be used to estimate the risk that a participating site is or is not performing tasks at a high level of quality. This approach can be used to shift monitoring from a purely schedule-driven effort to one that is more focused, efficient, and cost effective. The process for risk-based monitoring should be described in a separate monitoring SOP, which can be referenced in the DMP. The DMP should document which study-specific data elements are used to estimate risk and how they are used. The following should be addressed in the DMP (see Figure A6).

- List the data elements used in the estimation of risk
- What types of calculations or trend analysis based on the data elements is used to estimate risk? These could include, but are not limited to,

- High rate of AEs or low rate of AEs
- High number of procedures for multiple subjects performed within a limited number of days
- High rate of subject enrollment
- High percentage of missing data
- Data entered into EDC system for multiple subjects within a limited number of days
- Unusual distribution of values, excessively narrow or wide, for primary and secondary endpoint-related data elements
- Reports used for deriving risk and how are these provided
- What is the responsibility of the DM?
- Is there a system used to derive risk? If so, what is the system, and who has access?

5.9 Database Lock

Database locking and unlocking procedures are usually described in a separate SOP. This locking/unlocking SOP should list the data quality checks that take place prior to locking. This SOP can be referenced in the DMP; however, the database-locking information can also be placed in the DMP.

The locking procedures should include, but are not limited to, verifying the following:

- All data queries have been resolved
- The safety database has been reconciled with the clinical trial database
- All subjects have a final status
- All protocol deviations have been addressed

Data Elements Required			
Table	Data Element	Primary or Secondary Endpoint	Comment
		☐ P ☐ S ☐ No	
		☐ P ☐ S ☐ No	
		☐ P ☐ S ☐ No	
Calculations and Trends used to estimate risk (This may be placed in the monitoring plan)			
Calculation or Trend Analysis	System or Program Used	Presented in Report	Responsible Person

Figure A6. Risk-based monitoring data management information.

- All site access rights have been terminated
- All site signatures have been obtained
- Database locking approval signature from the sponsor and PM has been obtained

6.0 Archiving

Identify where the following will be stored (these can be paper and/or electronic documents and can refer to a network directory and/or physical storage area):

- DMP
- Final data files
  - Clinical trial database
  - SDTM data sets
- Documentation related to the data files (metadata)
  - Data dictionary
  - File description of any external files used during the study
  - PDF files of the CRFs (blank CRFs, no content)
  - Completed user acceptance testing scripts
  - Approvals for CRFs and user acceptance testing
  - PDF files of completed CRFs
- SAP
- SAS programs
- Other relevant documents or files

6.1 Conversion to CDISC SDTM

Comment: The methods to convert from the table structure used by a clinical trial database to the SDTM domains vary. Some systems store data directly into an SDTM structure, while others convert tables to an SDTM structure after database lock (back-end conversion). The back-end conversion can make use of commercial extract transform and load systems or create a CDISC ODM file as an intermediary step, while others use SAS code to directly convert tables to SDTM data sets. There are also applications used to validate the structure of the SDTM data sets. Because of this variability, it is difficult to write general data management procedures related to the creation of SDTM data sets. It is encouraged that the DMP include a description of all steps used to create the SDTM data sets and the associated QC procedures. The following information should be the minimum used to describe this process.

- Method used to convert data to SDTM structure
- Original data stored in SDTM structure
- EDC system converts data to ODM for latter conversion to SDTM
- Back-end conversion
  - Dedicated extract transform and load system
  - SAS programs
    - Conversion from original data to SDTM
    - Conversion from ODM to SDTM
  - QC steps used to verify that conversion does not corrupt data
  - QC process used to validate that the SDTM data set structures are correct
- Method used to validate structure of SDTM datasets

References

1. National Cancer Institute. Cancer Prevention Agent Development Program: for phase I and phase II clinical trials of cancer chemo preventative agents. http://prevention.cancer.gov/Files/clinical-trials/Master-DMP-Template.doc. Published February 2013.

2. Cardiff and Vale University Health Board. Data management in clinical trials: guidelines for researchers. http://www.cardiffand-valeuhb.wales.nhs.uk/opendoc/198842. Published 2012.

3. European Clinical Research Infrastructures Network, Translational Working Group. Deliverable D10: GCP-compliant data management in multinational clinical trials. http://www.ecrin.org/fileadmin/user_upload/public_documents/About_ecrin/downloads/ECRIN_Report_D10_Vers1_final_150908.pdf. Published 2008.

4. National Institute of Dental and Craniofacial Research. Tool summary sheet: clinical data management plan template, approved v10.doc. http://www.NIDCR.gov. Published 2014.

5. Ohmann C, Kuchinke W, Canham S, et al. ECRIN Working Group on Data Centres: standard requirements for GCP-compliant data management in multinational clinical trials. *Trials*. 2011;14:85.

6. Prokscha P. *Practical Guide to Clinical Data Management*. 3rd ed. Boca Raton, FL: CRC Press; 2012.

7. Society of Clinical Data Management. Good clinical data management practices, version 4.0. http://www.scdm.org/sitecore/content/be-bruga/scdm/Publications.aspx13. Published 2014.

8. Dayer L, Heldenbrand S, Anderson P, Gubbins PO, Matrin BC. Smartphone medication adherence apps: potential benefits to patients and providers. *J Am Pharm Assoc*. 2003;53(2):172-181.

9. Nixon A, Colman S, Sabounjian L, et al. A validated patient reported measure of urinary urgency severity in overactive bladder for use in clinical trials. *J Urol*. 2005;174:604-607.

10. FDA. Guidance for industry computerized systems used in clinical investigations. http://www.fda.gov/ohrms/dockets/98fr/04d-0440-gdl0002.pdf5. Published 2007.

11. FDA. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf. Published 2009.

12. Oncescu V, Mancuso M, Erickson D. Cholesterol testing on a smartphone. *Lab Chip*. 2014;14:759-763.

13. Sheetz N, Wilson BSP, Benedict J, et al. Evaluating source data verification as a quality control measure in clinical trials. *Therapeutic Innovation & Regulatory Science*. 2014;48:671-680.

14. FDA. Guidance for industry: oversight of clinical investigations: a risk-based approach to monitoring. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM269919.pdf. Published August 2013.

15. EMEA. ICH topic E6 (R1): guideline for good clinical practice. Note for guidance on good clinical practice (CPMP/ICH/135/95). http://ethikkommission.meduniwien.ac.at/fileadmin/ethik/media/dokumente/rechtsgrundlagen/GCP.pdf. Published. January 1997.