**Tool Summary Sheet**

**Tool:** Clinical Data Management Plan Template – External Facing

**Purpose:** This Clinical Data Management Plan (CDMP) template may be employed for studies using an Electronic Data Capture System (EDC), unless another template has been agreed upon.

**Audience/User:** Lead Data Managers and Principal Investigators of studies using Electronic Data Capture Systems

**Details:** The template should be customized to the protocol, the study’s special needs / circumstances, and the requirements of the data capture system. Sections may be edited or deleted as needed.

As an alternative to this detailed template, individuals may prefer to reference the ‘Data Management Considerations (Condensed)’ tool, which is available in NIDCR’s Toolkit for Clinical Researchers ([http://nidcr.nih.gov/research/toolkit](https://www.nidcr.nih.gov/research/human-subjects-research/toolkit-and-education-materials)). This considerations tool lists important issues to consider when evaluating a group’s ability to electronically capture and validate clinical research data, and summarizes the wide array of data management tasks required. The tool could also be used to suggest headings and content for a simplified clinical data management plan.

**Best Practice   
Recommendations:**

* Items in blue italics and enclosed in braces { } are instructional text that should be deleted prior to approval.
* Items enclosed in single <> are placeholders. Replace as clarified in the enclosed text.
* References to “Sponsor” may refer to the funding sponsor or to the study PI, depending on the context of the sentence and on the decisions at the time of CDMP development. Update the template at the time of implementation to refer to the appropriate party.
* Please retain the CDMP Template identifier in the lower left hand section of the footer.
* Remove this Tool Summary Sheet prior to use of this template.

**Tool Revision History:**

**Version**

|  |  |  |
| --- | --- | --- |
| **Number** | **Date** | **Summary of Revisions Made:** |
| 1.0 | 12Jan2012 | First approved version. |
| 2.0 | 06Jan2017 | Revised Approval Signatures page to reflect PI as primary signatory. |

Clinical Data Management Plan

Protocol Title: <Protocol Title>

Protocol Number: <#>

|  |  |  |
| --- | --- | --- |
| **Version Number and Date:** | Version <x.x>, <DDMMMYYYY> | |
| **Funding Sponsor:** | National Institute of Dental and Craniofacial Research | |
| **Study Principal Investigator:** | <PI Name and credentials> | |
| **Data Coordinating Center Data Management Contact(s):** |  |  |

**Summary of Changes:**

**Version**

| **Number** | **Date** | **Affected Section(s)** | **Summary of Revisions Made:** |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Clinical Data Management Plan

Protocol Title: <Protocol Title>

Protocol Number: <#>

## APPROVAL SIGNATURES

{Ensure that the CDMP version number and date are correct before providing the document to the primary signatory.}

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Principal Investigator: <Name, Affiliation> Date

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Project Data Manager: <Name, Affiliation> Date

## TABLE OF CONTENTS

[APPROVAL SIGNATURES 4](#_Toc471471355)

[TABLE OF CONTENTS 5](#_Toc471471356)

[ABBREVIATIONS AND DEFINITIONS 6](#_Toc471471357)

[DYNAMIC REFERENCES 7](#_Toc471471358)

[1 Introduction 10](#_Toc471471359)

[2 The Electronic Data Capture (EDC) System and the Underlying Clinical Database 11](#_Toc471471360)

[3 Network Directories 12](#_Toc471471361)

[4 Data Validation Process 13](#_Toc471471362)

[4.1 Univariate Alerts 13](#_Toc471471363)

[4.2 Multivariate and Cross-module Alerts 13](#_Toc471471364)

[5 Verification of EDC Setup and Implementation 15](#_Toc471471365)

[5.1 System Specifications 15](#_Toc471471366)

[5.2 User Acceptance Testing 15](#_Toc471471367)

[6 Data Entry and Data Cleaning 17](#_Toc471471368)

[6.1 Pre-requisites for Site Data Entry 17](#_Toc471471369)

[6.2 Granting Access to the Production Version of the EDC 17](#_Toc471471370)

[6.3 Entering Data 17](#_Toc471471371)

[6.4 Data Security 17](#_Toc471471372)

[6.5 Quality Control Procedures 17](#_Toc471471373)

[6.6 Query Generation 18](#_Toc471471374)

[7 Loading Electronic Files 19](#_Toc471471375)

[7.1 Laboratory Normal Ranges 21](#_Toc471471376)

[8 Medical Coding 22](#_Toc471471377)

[8.1 Adverse Event/Medical History Coding 22](#_Toc471471378)

[8.2 Medication Coding 22](#_Toc471471379)

[9 EDC Exported Database 24](#_Toc471471380)

[10 Reports 25](#_Toc471471381)

[11 SAE Reconciliation 26](#_Toc471471382)

[12 Changes to a Production EDC 27](#_Toc471471383)

[13 Database Closure 28](#_Toc471471384)

[13.1 Closure Checks 28](#_Toc471471385)

[13.2 Quality Assurance Audit and Database Lock 28](#_Toc471471386)

[13.3 Database Unlock 28](#_Toc471471387)

[14 Data Archiving and Provision of Final Materials to Sponsor 29](#_Toc471471388)

## ABBREVIATIONS AND DEFINITIONS

|  |  |
| --- | --- |
| AE | Adverse event |
| CDMP | Clinical Data Management Plan |
| CRF | Case report form (may be paper or electronic representation of the data collection tool) |
| Cross-check | An edit check that compares variables from different CRF pages or subjects variables from different CRF pages to an algorithmic evaluation (term of convenience used to differentiate from multivariate edit checks) |
| CTM | Clinical trial material |
| DCC | Data coordinating center |
| EDC | Electronic data capture |
| MedDRA | Medical Dictionary of Regulatory Affairs. Dictionary for coding adverse events, medical history, and physical exam findings |
| Multivariate Edit Check | An edit check (above and beyond a range check, valid value check, or required criterion) on a variable or set of variables on the same CRF page (term of convenience used to differentiate from cross-checks) |
| Project Manager | Individual who manages the project at the data coordinating center. This person’s job title might project manager, study coordinator, or study team lead. |
| SAE | Serious adverse event |
| SOP | Standard Operating Procedure |
| WHODRUG | World Health Organization coding system for medications |

*{Adjust this list according to the abbreviations actually used in the final document.}*

## DYNAMIC REFERENCES

*{Dynamic references are documents that are relevant to clinical data management activities but are not considered part of the CDMP; they are stored separately. The reasons that these documents are separate from the CDMP are threefold. First, the dynamic references may require individuals other than the CDMP primary signatory to approve them, when approval is required. Second, they may be updated on a frequent basis, which would create an unnecessary administrative burden for the CDMP signatories. Third, they may have a document owner that is someone other than the project data manager. This section should discuss the above issues as they pertain to this study and the external documents listed in Table 1. Table 1 should list the full set of dynamic references to be created for this study, including a description of their content and the local storage location of the item.*

Some items may be handled differently than listed (e.g., some may be combined). Customize the list to the needs of the study.

Not all dynamic references will be prepared at the time of the initial CDMP signing. If the dynamic reference is expected, include it in the list below and create a placeholder file in the local storage location indicating that it is not yet available and the approximate time it is expected to be prepared along with the study title of the document owner (e.g., project manager).}

### Table 1. Dynamic References

| **Item** | **Content** | **Document Owner** | **Current Version Storage Location at DCC** |
| --- | --- | --- | --- |
| Study Contact List | Names, roles, and contact information for key sponsor and DCC study team members | Project lead | {E.g., the study website, the communication plan (including network location)} |
| Annotated Case Report Form | The full set of CRFs, including dataset and variable names, sorted in protocol schedule order | Project data manager |  |
| Data Dictionary | List of dataset names. List of variable names, with corresponding valid values, data types, labels. | Project data manager |  |
| Data Validation Plan | Univariate ranges, description of each electronic edit check and custom function edit check, and any additional manual checks that will be performed | Project data manager |  |
| Clinical Data Review Plan | Description of process for providing data to one or more clinicians who will complete a clinical review.  {Only include if a clinical review will be conducted. If a clinical review will be conducted, this plan would also specify the fields that will be reviewed and the timing of the review. This document may be created well after the initiation of the study in the EDC.} | Project data manager |  |
| EDC Specifications Review Checklist | Describes and tracks the process for review and approval of the study-specific EDC requirements (e.g., blank and annotated CRF, edit check plan) prior to the initiation of User Acceptance Testing. | Project data manager |  |
| Study Startup Checklist | List of items to be completed during study startup. | Project data manager |  |
| Study Timeline | Projected timeline of study events and deliverables | Project manager |  |
| Sponsor Converted Database Specifications | Describes the requirements of the clinical database that is delivered back to the sponsor after database lock.  {This conversion may not be required. Delete item if not applicable. Alternatively, this may be the specifications for the creation of de-identified public use data} | Project SAS programmer |  |
| Development and Testing Responsibility Matrix | Identifies tasks and deliverables for this project and indicates who is responsible for writing, executing, reviewing, and approving each item. | Project data manager |  |
| Clinical Monitoring Plan | Description of clinical monitoring activities  {This often includes plans for site initiation visits and site user training on the EDC, which is why you may want to reference it in the CDMP.} | Lead clinical research associate |  |

## Introduction

{Briefly describe the study, including the clinical phase and high-level protocol design components that are not included in the protocol title (e.g., “a phase 2-3, randomized, double-blind, placebo controlled study of subjects with hypoparathyroidism. Each subject will be treated for up to 3 years.” Consider the nature of the study (e.g., minimal risk study, clinical trial requiring an IND) and indicate. “Activities for this study must comply with all relevant regulations.”

Provide general introductory text as to the purpose of the clinical data management plan (CDMP) including its central role in making explicit to all stakeholders specific information regarding the data management practices needed to ensure appropriate handling of data at all steps of the project to assure a high-quality database at the end of the study, ready for analysis. Describe how the document will be reviewed, approved, and finalized (i.e., signed by whom), how it will be modified (minor or substantial) if needed during the project, and any SOPs in place governing its use and modification.}

## The Electronic Data Capture (EDC) System and the Underlying Clinical Database

{Describe the EDC system proposed for use in this study, including its version number or other identifying information and validation status with respect to 21 CFR Part 11 regulations. More generally, describe the system’s primary information security, disaster mitigation, audit trail, and electronic signature features. Identify ancillary software packages associated with the EDC, such as underlying database stacks and statistical analysis software. Describe the physical relationship of the server(s) used to house the database and statistical or other support software packages vis-à-vis the computers used by the project team.}

## Network Directories

{Describe the security, disaster mitigation, and structure of the network directories (physical and directory/subdirectory) used by the project team to store study-related information.}

## Data Validation Process

{Subsections here should describe the computerized portion of the edit check validation employed by the EDC system. Note: often, full data validation specifications are documented in a standalone Data Validation Plan, cited in the dynamic references table above.}

### Univariate Alerts

{Describe the kinds of univariate alerts afforded by the EDC system. Examples might include valid-value, valid-range, and missing-value alerts and may be separately specified for each electronically captured field. For each subsection 4.1.x, provide additional details on the nature of the automated checks. An example is provided for Section 4.1.1 for valid-value checks; other subsections 4.1.x should be added as appropriate for each category of univariate alerts.}

#### Valid Value Alerts for Multiple-choice Fields (Select only 1)

*{Assuming the EDC provides for valid-value edit checks, provide text to explain the available kinds of checks. Some examples include:*

1. *The user is required to choose at least one checkbox from a series that represents different data choices*
2. *The user is required to choose only “yes” or only “no”*
3. *The user must record a coded value taken from a list that is usually presented on the page or in an instruction guide. Incorrect codes must be detected and rejected by the edit check scripts.}*

### Multivariate and Cross-module Alerts

*{In addition to univariate alerts that may be described in Section 4.1, describe here any relational alerts among groups of variables and among the same variable for protocol-specific assessment times. Such alerts are referred to as multivariate alerts (within one module, same assessment time) and cross-module alerts (across two or more modules or assessment times). For example, a cross-module alert could be specified to generate a query for subjects whose CRFs indicated that they were male (Demography module) and pregnant (Pregnancy module).*

*Common types of multivariate alerts include the following:*

1. *Confirming that only valid options are selected in a “choose all that apply” multiple choice field, where the range of options deemed valid depends on some other parameter.*
2. *Confirming that diastolic blood pressure reading is less than the associated systolic blood pressure reading.*
3. *Confirming that “Other, specify” is completed when “Other” is marked.*

*Common types of cross-module alerts include the following:*

1. *Comparing the dates and times of all assessment time points to confirm that they occur in an appropriate sequence. For example, a Study Day 1 assessment should occur before a Study Day 5 assessment.*
2. *Confirming that all visit dates occur before the date of the investigator’s signature.*
3. *If applicable, confirming that all screening and baseline assessments occur before the first dose of clinical trial material (CTM) is administered.*
4. *Confirming that the recorded onsets of all treatment emergent adverse events are after the first dose of CTM is administered.}*

## Verification of EDC Setup and Implementation

*{Describe here the procedures and documents used to create and validate a protocol-specific implementation of the EDC.}*

### System Specifications

{Detail the functional and other system specifications for the EDC system. The project lead, project manager/study coordinator, clinical research associate, biostatistician, clinical site team, and sponsor can be consulted for guidance on the development of these specifications.

System specifications for the EDC system typically include: data entry screens, annotated CRFs, and the data validation plan. A process should be developed to guide and document review and approval of EDC system functioning with respect to these items. This process should be completed before undertaking user acceptance testing.

Note: it is acceptable for the developed data entry screens to act as the “specifications” for the data entry screens and for the annotated CRF generated for the project to act as the specifications for the database.}

### User Acceptance Testing

{Each study-specific implementation of the EDC system should be tested by candidate users, with the results being documented in a Test Summary Report. User Acceptance Testing commonly includes confirmation of the proper functioning of the items listed in the table below.

Add or remove items as relevant and update details of testing based on study-specific needs.

It may be helpful to include in the testing an exercise involving entry of data for (a) a complete mock subject from start to finish and (b) a mock subject who discontinues early. This can be one way of testing several of the items below.}

| **Item** | **Details of Testing** |
| --- | --- |
| 1. Univariate and valid value checks | Confirm that checks have been properly imported from specifications document. Manually test a subset of checks (e.g., 2 per page). |
| 1. Multivariate and cross-checks | Confirm each <or a specified subset of> checks via test data designed to trigger a query. |
| 1. Custom functions | Confirm each custom function with test data designed to trigger an event.  Second programmer code review, if possible. |
| 1. Generation of subject numbers | Confirm that new subject IDs are generated according to specifications  {(e.g., site number+number of subject enrolled at that site+check digit—subject 16028 is from site 16, the second subject enrolled (02), and 8 reflects the check digit generated by an algorithm used to confirm that the number is valid)} |
| 1. Customized form or variable delivery (e.g., dynamic fields) | Confirm against schedule of events and against other specifications. |
| 1. Data completion guidelines | Confirm that the completion guidelines are properly associated with each form/field. |
| 1. Derived variable computation | Confirm against specifications using test data. |
| 1. Email alerts | Confirm with test data designed to trigger email alerts. |
| 1. Role assignment | Review system. Confirm using list of role functionality, have testers assigned to each role, and ensure that they are only able to do/see what they are entitled to per their assigned role. |
| 1. Data Extracts | Review extracted data, ensure that it matches specifications (e.g., annotated CRF). |
| 1. Data Imports into EDC | Create test data, import, and review imported data. |
| 1. System Reports | Review system reports and ensure that they are functioning according to expectations. Run reports on test data. |
| 1. E-Learning Specifications | Ensure that any e-learning items are delivered based on stated requirements and roles. |

## Data Entry and Data Cleaning

### Pre-requisites for Site Data Entry

{Each site user must be trained on the EDC system prior to being granted permission to work in the production version of the EDC system. Describe site specific training requirements such as “the clinical research coordinator must have completed the data entry and query resolution modules of the EDC and/or must enter test data for the demographics, AE, and study termination pages. These data must be reviewed and approved by the lead data manager.”}

### Granting Access to the Production Version of the EDC

{Describe the process for granting access to the production version of the EDC system for site and other sponsor staff. For example, “Once the test data have been reviewed and approved, the lead data manager will send a written request to the helpdesk or study specific user administrator to grant production system access.}

### Entering Data

{Clarify what data will be entered by trained site personnel as opposed to by external individuals such as DCC personnel. Specify which fields will be double-entered in the EDC system.

For the most part, data entered into the EDC system are based on source documentation maintained at the clinical site. For cases where the EDC system will be used as the initial data record (e.g., direct data entry of a dental exam) and there is no other source documentation, GCP requires that these items be identified in the protocol. For completeness, identify those items in this section as well.}

#### Data Entry Completion Guidelines

{Describe any documents or in-system resources that will be developed to help users during the data-entry process. Describe where these resources may be found (i.e., network drive location, URL, or functional area within the EDC, etc.)}

### Data Security

{Detail here the procedures and methods to be used to ensure data security. Examples include employing individual user accounts with role-based data management and access privileges, website security technologies, database access security, server physical plant security features, and backup or restore processes that constitute the EDC system’s disaster-mitigation plan.}

### Quality Control Procedures

{Describe the quality control (QC) processes to be employed. Such procedures may include built-in procedures of the EDC system such as automated checks to ensure manually entered subject numbers conform to study-defined site/subject # format rules and real-time data value edit checks as referenced above. More-manual checks may include strategies such as developing SAS code to ensure the right number of entries is present for a given data domain.

### Query Generation

#### EDC-generated Queries

{Summarize how queries are generated from within the EDC system, such as those that should result from automated edit checks. Describe how the user is alerted and then allowed, if applicable, to make a correction or alternately to override the alert (along with rationale for overriding the alert). Some data fields may require correction (i.e., should not allow users to override a correction alert); specify these here. Also describe who/what role is tasked with reviewing overrides.}

#### Manual Queries

{Queries issued by DCC study staff (e.g., data managers, DCC study coordinators, clinical research associates) are referred to as manual queries. In this section, explain how manual queries are generated and tracked to resolution, including whether such queries are tracked in the EDC system. Manual queries are used in cases where the data issue is a) sufficiently complex as to be impractical to program as an automatic system check and/or b) requires human judgment. Examples of these cases include misspellings that hinder medical coding and checks that require interpretation of meaning in order to ascertain whether an entry should be queried. For example, a programmed logic check between free-text fields would not be practical because of the human interpretation needed to understand the relationship between “high-blood pressure” on one form and “hypertension” on a corresponding form.}

## Loading Electronic Files

{Describe here how electronic data files generated outside the EDC system will be integrated with the EDC database. Examples of such files include clinical laboratory data or pharmacokinetic data generated at offsite facilities. Detail how often such transfers will take place, whether the transfers will be cumulative or incremental, in what electronic format the transfers will be provided, and how transfers will be logged or documented. Also detail what security procedures will be implemented to protect against malware in the data files. Indicate whether the data are integrated directly into the database, stored on a network share for later processing, or some combination of the two. If data files are to be integrated into an existing database, describe procedures in place for verifying proper database structure and for resolving problems when the data are not in the proper data structure or file format.

It may be helpful to summarize this information in tabular format as shown in the sample Table 2 below.}

### Table 2. Electronic Data to be Integrated in the EDC

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Data Type** | **CRF Page** | **Data Source/ Format** | **Frequency of Transfer/ Location** | **Cumulative or New Only** | **Special Instructions** |
| Hematology | HEMA | NIH lab/ .xls | Nightly/ Secure FTP | Cumulative | Store in masked directory S:\project\data\fromnih. Automated conversion program runs nightly to map medical record number to subject ID and to import into EDC. If errors occur, an email will be sent to the lead data manager and EDC setup expert. |

### Laboratory Normal Ranges

{Clarify the process for receiving and using laboratory-issued normal range data. For instance, “Laboratory normal data will be collected from each site or central laboratory, if necessary for the study. Those data will be included in multivariate edit checks to identify clinically significant values.” Or “Laboratory normal data will be collected from each site as part of each subject’s CRF. Hence, there will not be a separate laboratory normals database for this study.}

## Medical Coding

*{Summarize the tools, methods, and conventions established for this study regarding medical coding. Medical coding is the process of deducing from literal or free-form text a set of terms that conform to a pre-defined dictionary. Coding is commonly used to derive standardized terminology for adverse event descriptions and medications apart from the investigational product that are used in a study. Standardized terminology is essential for analyzing study data programmatically.*

*Specify the coding dictionary version and the level down to which coding will be performed. A tabular summary by data type (i.e., adverse event or medication) may be helpful.*

*Describe how often medical coding for new and updated events/medications will occur during the course of the study. It may be helpful to schedule interim reviews of medical coding prior to DSMB meetings or formal interim analyses. The review schedule must be detailed in the CDMP or elsewhere. A final review of coding must be performed before database lock.}*

### Adverse Event/Medical History Coding

*{Each reported adverse event/medical condition will be assigned both a symptom coding symbol (term) and a MedDRATM body system value. If an automated software package is used to facilitate coding, describe the package and how it will be used in conjunction with any manual coding procedures.*

*This is particularly relevant for coding adverse events that fail to provide a direct match between verbatim and preferred (coded) terms. For example, a coding program may be employed to map each verbatim term in a specified CRF field to a standard term or its synonym* *listed in an existing dictionary definitions or described in a set of pre-determined, study-specific rules. A coding rule is a specification that maps a verbatim term to a standard dictionary term. When dictionary entries and study-specific coding rules fail to provide matches, a medical coding specialist under the direction of the lead data manager may review candidate dictionary entries as well as other related information such as comment fields to identify an appropriate match. Specify whether queries to the site will be issued to facilitate this process.*

*This section also should detail how compound verbatim events such as “cough induced seizures” are to be handled. Many, but not all, such events must be split into more than one record for purposes of medical coding. Indicate whether the data manager will query compound events (requesting data updates from the site) or whether the events will be split into separate entries by the medical coder during coding.}*

### Medication Coding

*{Describe here the processes planned for coding medications and the dictionary that will be used (e.g., WHODRUG 2011, Q4). As with AE coding, describe any software that will be used to facilitate coding, how human review or final adjudication will be managed when automated coding fails, and under what circumstances queries may be issued to the site for clarification or additional details. When multiple drugs are listed in a single field, each should be coded into separate entries to facilitate separate tracking of each medication.}*

## EDC Exported Database

{Detail the scope, frequency, and format of data exports from the EDC system. Explain how predefined vs. custom exports will be created and handled.}

Note: it is sometimes necessary to transfer clinical data in a pre-specified dataset format that may not be supported by the EDC system, for example when transferring to an external entity such as a sponsor. If this is the case, it will be necessary to write code to convert the data. The resulting datasets are distinct from analysis datasets whose creation is supervised by a project statistician.

If conversion and/or transfer are not required, please indicate “Not applicable for this study.” Otherwise, describe the processes for converting data according to the required specifications, and for validating the data conversion. Also describe any pre-established data transfers.

If the data are being converted to a Public Use dataset, per federal requirements, identify that dynamic reference document in this section.}

## Reports

{List reports anticipated to be generated for this study in Table 3; examples are shown below. Additional reports that are agreed upon during the course of the study do not require an immediate amendment to the CDMP. Simply ensure that the final version of the CDMP (usually signed immediately prior to database lock) reflects all delivered reports.}

### Table 3. List of Data Management Reports

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Report Name** | **Description/Purpose** | **Frequency** | **Recipients/**  **Users** | **Method for Provision** |
| Medical Coding Report | Summaries of verbatim and corresponding coded AEs and ConMeds. For approval of medical coding, including split events, preferred terms, and System Organ Class. | 1 month prior to DB freeze for DSMB reports and prior to database lock. | Medical Officer or designee | Emailed zip file |
| Query override report for EDC | Lists all queries where the data entry user overrode the query with no change, includes the explanation for the change/DM review for acceptability of overrides. | Monthly | Data Managers and other DM staff | Automated SAS report stored on local network. |

## SAE Reconciliation

{This section should detail the processes and specific details to be used in reconciling the safety and clinical databases. The safety database may be a separate database used to hold emerging SAE data as they arise during conduct of the study and is subject to different data management timelines and QC processes than the EDC database.

At intervals during a study and/or at the end of the study, the two databases must be reconciled to ensure all information has been appropriately captured and reported to cognizant IRBs and regulatory authorities as applicable. Specify here the milestone events (e.g., DSMB meetings) and/or basic frequency during study conduct at which reconciliation will be undertaken. Identify which study team roles will be involved in the reconciliation process.

The specific data fields to be compared between the two databases must be listed; these commonly include verbatim and preferred terms, onset and stop dates, severity, causality, action taken, etc. Procedures for handling discrepancies between the databases, including identifying which database needs modification and tracking/documentation of discrepancies, should also be clearly detailed.

Note: If SAE reconciliation is not required, please indicate “Not applicable for this study” and remove the text in this section. If Product Safety is being handled by a 3rd-party, update this section accordingly to reflect the SAE reconciliation process being used.}

## Changes to a Production EDC

{As with a study protocol, changes are sometimes needed to the data management plan. When such changes result in modifications to the EDC system, special considerations apply and should be detailed in the CDMP. In this light, describe here the process of proposing, reviewing, and documenting changes to a production EDC system. The impact on the system and the user should be evaluated by the lead data manager, and the team should establish a process for communicating the changes and the impact of the changes to site personnel. Further, the methods for evaluating the changes and determining the level of testing required before the changes are applied to the production server should be described. Testing needs to include not just performance assurance, but also user acceptance testing, conducted in the development environment, with all testing being documented. Indicate where and how testing documentation will be stored. Finally, indicate the study personnel/role authorized to approve the proposed changes to the production EDC system.}

## Database Closure

{This section should describe all processes associated with final database closure. Subsections proposed here include final pre-closure data checks; QA review and actual database lock; and procedures for unlocking a database in the event that data must be corrected after lock.}

### Closure Checks

{Describe the set of closure procedures that will be performed prior to database lock to verify the integrity and completion of the database. In some cases these may be the same data checks described in above sections, repeated until all queries are resolved and the data are determined to be clean. Other appropriate checks at this stage may include:

* Check that all expected CRFs have been entered;
* Check that the project database is consistent with the specifications in the project data dictionary;
* Determine the status of each subject entered (i.e., excluded, ongoing, completed, withdrawn, lost to follow-up, etc.);
* Check for value formatting problems in database exports;
* Check for consistency between whole-date fields and associated part-date fields;
* Confirm that all expected site signatures have been applied.}

### Quality Assurance Audit and Database Lock

{Describe QA processes to be employed during study conduct to assure data quality, such as clinical site monitoring. Specify QA procedures after study conduct to be performed to ensure all possible errors are detected and corrected before final database lock. For example, an audit may be performed of a sample of completed CRFs in the EDC system against exported SAS datasets to ensure the integrity of the final study data.

Explain how approval for final database lock will be sought and documented, and how the lock will be achieved at the database level. If a separate soft lock will take place prior to final lock, describe how the two events will take place and what may distinguish soft from hard lock in terms of user access and write privileges.}

### Database Unlock

{Database unlock should only occur in extraordinary circumstances. As such, special controls are needed to limit the scope of unlock. Explain here the conditions necessary to justify database unlock, the procedures for documenting justification and approval by relevant personnel (e.g., the Principal Investigator, Medical Monitor, Program Officer, etc.), personnel authorized to unlock the database, and procedures for making and documenting the approved changes.}

## Data Archiving and Provision of Final Materials to Sponsor

{Identify what paper or electronic datasets, programs, and documents will be maintained in archive status after study completion.

Include the scope, format, and process for delivery of final electronic data to other entities if the data are not to be retained solely within the study EDC. If subject-specific pdf renderings of the CRFs are to be provided, include that information herein.

If Public Use datasets will be delivered after study completion, indicate that information in this section. The detailed plan for the development of the Public Use datasets need not be included in this document, but the deliverable requirement should be documented here.}