INTERNATIONAL JOURNAL OF RESEARCH IN PHARMACY AND CHEMISTRY

Available online at www.ijrpc.com

Research Article

AUDITING OF CLINICAL DATABASE – AN INDISPENSIBLE ACT TO ENSURE HIGH-QUALITY, RELIABLE, AND STATISTICALLY SOUND DATA FROM CLINICAL TRIALS: REVIEW OF PLANNING, METHODOLOGY, COMMON FINDINGS AND STANDARDS FOR GOOD CLINICAL PRACTICE COMPLIANCE

Raju Sama*, N. Vijaya Lakshmi and M. Sathish

Axis Clinicals Limited, 1-121/1, Miyapur, Hyderabad -500 049, Telangana, India.

ABSTRACT

Clinical trials for establishing the safety and efficacy of a drug can be among the most costly and time-consuming elements of product development. The success of any clinical trial is dependent on assuring that the data collected is of good quality. Clinical Data Management is the process of collection, cleaning, and management of subject data in compliance with regulatory standards. It is a critical phase in clinical research, which leads to generation of high-quality, reliable, and statistically sound data from clinical trials. Audit forms an important part of a quality system to assure the data generated is of high-quality, reliable, and statistically sound and to provide verification of data integrity. Clinical data base audits may be conducted at any facility or institution where clinical trials are conducted on human volunteers or subjects. Auditors should be selected based on the suitable qualifications, experience and training. Adherence to specific audit principles is a prerequisite for providing a reliable and relevant audit outcome. Like in all other audits, database audits require careful preparation. Audits of the database are conducted between 'soft lock' and 'hard lock' ('freeze' and 'lock') of the database. Data transcribed from a Case Report Form or other source into the database is usually checked for accuracy through a database audit. Used effectively data base audit can reduce costs and ensure regulatory compliance. There exists no widely recognized, specific, practicable and open standard for Good Clinical Practice compliant data management and the accompanying Information Technology Infrastructure.

Keywords: Clinical trials, Data management, Quality system, data base audit.

INTRODUCTION

Clinical trials for establishing the safety and efficacy of a drug can be among the most costly and time-consuming elements of product development. They are intended to find answers to the research question by means of generating data for proving or disproving a hypothesis. The success of any clinical trial is dependent on assuring that the data collected is of good quality.[1]

Clinical Data Management (CDM) is the process of collection, cleaning, and management of subject data in compliance with regulatory standards. It is a critical phase in clinical research, which leads to generation of high-quality, reliable, and statistically sound data from clinical trials. This helps to produce a drastic reduction in time from drug development to marketing.

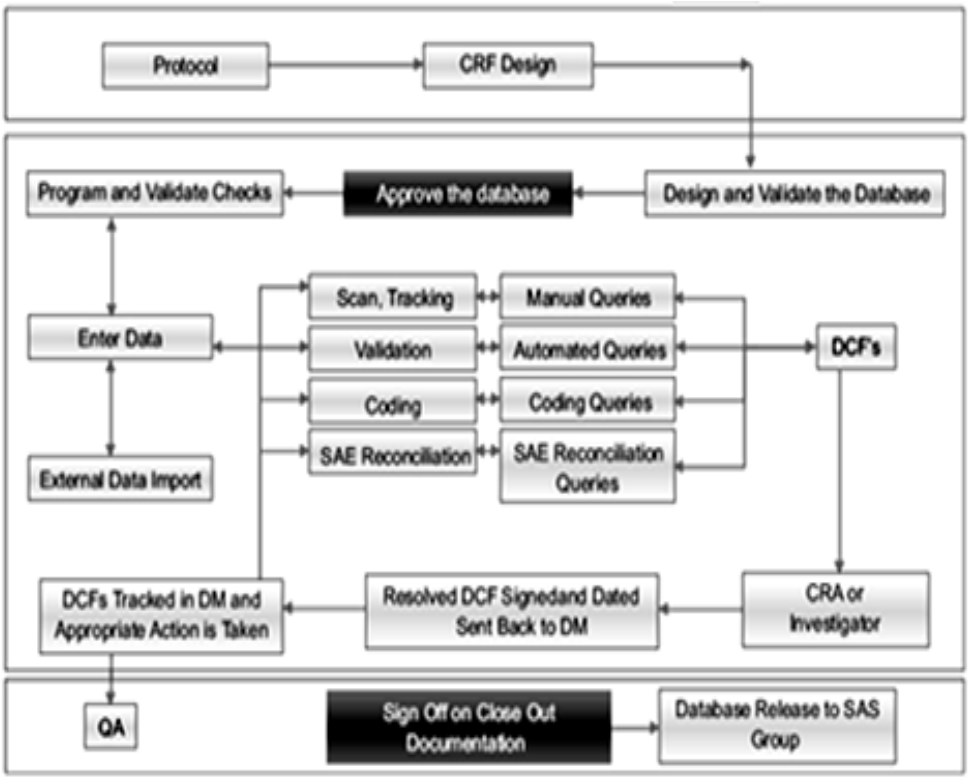
Clinical data base audits may be conducted at any facility or institution where clinical trials are conducted on human volunteers or subjects. Due to the increasing complexity of clinical trials and regulatory scrutiny, the components of a data base audit program and the approaches taken towards designing and managing audits are constantly evolving. In the present scenario, there is an increased demand to improve the CDM standards to meet the regulatory requirements and stay ahead of the competition by means of faster commercialization of product. Audit forms an important part of a quality system to assure the data generated is of high-quality, reliable, and statistically sound and to provide verification of data integrity.

Although data base audit alone cannot transform a poorly conducted or analyzed trial into a credible one, but data base audit program will point out potential problem areas early, so solutions can be found before the data submission to regulatory authorities. Used effectively data base audit can reduce costs and ensure regulatory compliance. To make sure that these benefits will be realized, however, sponsors must develop a comprehensive auditing strategy. This article highlights the processes involved, audit planning and methodology, common findings and standards for good clinical practice compliance.

ICH GCP defines audit as 'a systematic and independent examination of trial related activities and documents to determine whether the evaluated trial related activities were conducted, and the data were recorded, analysed and accurately reported according to the protocol, sponsor's standard operating procedures (SOPs), Good Clinical Practice (GCP), and applicable regulatory requirement(s).

Clinical Data Management Process overview

On closer examination, the process of data management comprises a variety of individual sub-processes. As a clinical trial is designed to answer the research question, the CDM process is designed to deliver an error-free, valid, and statistically sound database. To meet this objective, the CDM process starts early, even before the finalization of the study protocol.[2]



AUDITING OF CLINICAL DATABASE

Auditing Principles

Adherence to these principles is a prerequisite for providing a reliable and relevant audit outcome. These principles relate to auditors:

1. **Ethical Conduct**
   Trust, integrity, confidentiality and discretion are essential to auditing. Actions that may influence the results of an audit should be avoided.

2. **Impartial reporting**
   The obligation to describe truthfully and accurately the audit activities.

3. **Due professional care**
   The application of diligence and judgement in auditing. Reasonable care in all matters and the completeness of the audit report avoiding errors that may jeopardize any of these auditing principles.

Two further principles relate to the audit process:

4. **Independence**
   Auditors cannot audit work where a conflict of interest would arise. They must maintain an objective state of mind throughout the audit process to ensure that the findings and conclusions will be based only on the evidence.

5. **Evidence**
   The rational basis for reaching reliable audit conclusions based on audit criteria.

Choosing an Auditor

One of the first, and most important, decisions the Sponsor faces is choosing an auditor. Because they are sometimes not part of the sponsors' organizations, contract auditors can promote a spirit of objectivity and encourage investigators to communicate problems openly. Auditors can also be chosen from the sponsor's quality assurance organization or from any group not associated with direct management of the clinical trial.

Auditors should be selected based on the following qualifications/experience and training:

1. Suitable experience and education
2. Independence
3. Formal regular appropriate training
4. Understanding of the clinical trial and clinical data management process.
5. Up-to-date knowledge of International Conference on Harmonisation -Good Clinical Practice (ICH GCP) guidelines, Code of Federal Regulations (CFR), 21 CFR Part 11, Society for Clinical Data Management (SCDM) - Good Clinical Data Management Practices (GCDMP) guidelines, Clinical Data Interchange Standards Consortium (CDISC) standards and any country-specific guidelines or regulations, national laws and requirements related to Clinical Data Management.
6. Skills required: Communication, Writing, Language etc.
7. Nature: Tenacity, Power of observation, analytical capability, decision, sense of ethics and maturity.

Audit Planning

Elements of planning for an audit can be incorporated into an audit plan. An audit plan should include:

1. **Scope**
   To identify the intent, purpose, location, date (if known) of the audit activities and any relevant study identifiers.

2. **Contacts**
   To identify the key personnel involved in conducting the audit (both auditors and auditees)

3. **Agenda**
   Outline of detailed activities.

4. **Documentation/systems to be reviewed**
   To identify the documents/systems to be available for review.

5. **Audit History**
   To outline the audit history as relevant to the auditor- e.g., describes past interactions.

6. **Letter/Communication**
   Auditees should receive a letter of introduction with a confirmation of the audit dates and brief synopsis of activities to be conducted.

7. **Provision for Responses**
   Description of how responses are to be made (e.g. inclusion of action plan) and the expectant timeframe.

Like in all other audits, database audits require careful preparation. First, the audit must be arranged with the responsible database manager. Ideally, the person responsible for the data editing and/or programming should also be available for queries during the audit. In most cases, the date of database lock following the audit is a milestone in the project plan. The audit start date is calculated using this milestone date, taking into consideration the time required for conducting the database audit plus time for any required follow-up activities, to ensure that the database can be locked at the projected point in time. Database locking is usually a two-step process. The first step is often referred to as 'soft lock' or 'database freeze' and occurs after all data cleaning, validation and QC activities have been finalized. The second step is called 'hard lock' or 'database lock'. At this stage, the database is handed over to statistics for data analysis and the data can be unblinded (in case of a blinded study).

Thorough planning − paired with flexibility − is essential. Prior to the database audit, the auditor should receive the following documents related to the trial and the database:

- Study protocol including all protocol amendments, CRF, data management plan, statistical analysis plan. Annotated CRF, indicating the designation and names of CRF fields within the database to help the auditor identify the variables correctly.
- List of coding dictionaries employed in the trial (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Affairs (MedDRA), World Health Organisation Adverse Reaction Terminology (WHO-ART), WHO Drug Dictionary (WHO-DD)).
- List of the laboratory units used in the trial and possible conversions of units. List of 'self-evident' corrections performed by data management personnel. The list should include all corrections of obvious errors in the CRF that may be made by data management personnel without prior authorization by the investigator.
- List of all electronic and manual plausibility checks. It is recommended that this list be compared to the CRF as part of the audit to get a better understanding of the nature and adequacy of the checks.
- Standard Operating Procedures (SOPs) of all procedures related to data management. Prior to the audit, the auditor, data manager and the responsible member of the clinical team should agree on an acceptable error rate for comparison of subject data listings with original Case Report Form (CRF) entries.
- It is recommended to weigh the variables, e.g. to distinguish between primary and secondary variables to determine the sample size of the variables to be audited) and to define error rates based on this distinction. However, the definition of acceptable error rates does only make sense, if, at the same time, the consequences and corrective actions that apply in case acceptable error rates are exceeded are specified. It should be recognized that acceptable error rates and audit sample size are interconnected and cannot be treated independently. Details are discussed in subsequent sections of this chapter.

Audit conduct

Audits of the database are conducted between 'soft lock' and 'hard lock' ('freeze' and 'lock') of the database. Traditionally, database audits are performed under the maximum time pressure. However, there is value in conducting an audit earlier in the process, or, in fact, during the entire process. Small audits conducted throughout the study may identify problems with training staff at the site, problems at the sponsor in transcribing the data, or problems with the entry application. Since locking the study is usually a key time point with tight deadlines, doing some of the work earlier on may also improve the time to close. Whenever an audit is conducted, the listings, with errors highlighted and any summary or review should be filed in the study files.

a) Accuracy audits

Data transcribed from a CRF or other source into the database is usually checked for accuracy through a database audit. Data managers compare data in the database against the CRF and any associated correction forms. Many companies still perform the audit near the close of the study to determine an error rate. Unfortunately, if a late audit does detect problems, correcting them will prove challenging and time consuming. A more efficient approach is to perform audits early and, if appropriate, repeatedly, to catch systematic problems. Then,

at the close of the study, the data to be audited would be either the final data only or fields identified as critical or in some way risky to the study. This can be especially useful when the audit plan calls for a 100% audit of key fields. In this case, the audit of the key fields can begin as patients are considered "clean," that is, without any outstanding discrepancies.

The number most frequently used in selecting data for an audit is 10%. This is often supplemented by a 100% audit of safety fields such as those for AEs. Some companies also audit 100% of a selection of key efficacy fields. It is very important to note that a "10% audit" still does not tell us exactly what was or will be audited. Is this 10% of the patients, pages, or data? Ten percent of the patients may be easy to select but does not guarantee good coverage of investigator sites. Ten percent of CRFs is better as long as all pages are represented. Ten percent of the data by data set is a very good sample but can be hard to program and hard to select the pages associated with that data.

Many companies say that their acceptable error rate is 1% to 5% (one to five percent). However, articles regarding this topic maintain that data as well controlled as clinical trial data should have errors only in the range of 10 to 50 per 10,000. This translates into .1% to .5%. This latter figure also is in line with numbers for high quality double entry and should be a good and reasonable target for most organizations.

What is to be done if the rate is unacceptable as a result of the audit? If the audit is early in the data management process, it may be possible to improve upon the process or systems to improve the rate. If the audit is performed at the end of the study, it would be advisable to increase the number of fields audited to confirm the rate. Some companies immediately plan a 100% audit of all of the data. Other companies perform another 10% sample. Still others examine the result first and try to determine if any particular type of data or specific data modules are the source of the problem. They then conduct 100% audit of just that data.

One of the big advantages of electronic data capture (EDC) systems does not have to audit the database looking for transcription or entry errors. However, even for electronic data capture (EDC) studies, sponsors should consider a check of all changes to data that were made in response to queries. This check can be performed on an ongoing basis or as part of Quality Control checks at study lock. Experience has shown us that edits made to correct errors often introduce new errors. This is likely to be as true for site staff as it is for data entry or data management staff. Not all EDC systems (and the processes associated with using those systems) would support such a review, but it can be considered.

b) Summary review

There are certain kinds of cleaning or discrepancy checks that are better performed near the close of a study when the data is more complete. These include listing reviews, summary reports, and simple analyses of the data as a whole. The goal is to detect unusual values that stand out in the context of a set of data but that might otherwise pass cleaning rules or other discrepancy identification methods.

A listing review of text fields is a good example of how trained humans pick up inconsistencies that cannot be programmed into edit checks. Data managers may review listings of text fields to check for nonsensical words that are introduced because entry operators are focusing on what they see rather than the meaning of a phrase. A separate listing review by Clinical Research Associates (CRAs) is often required for study lock. The CRAs may notice nonsensical phrases and the like, but, more importantly, they may find problems with protocol compliance.

For example, they may review medications and find some listed that are not permitted by the protocol. Or, they may find medications listed in the medical history section. They may even find serious safety problems listed in comments associated with lab results or Adverse Event (AE) reports.

Humans are very good at detecting patterns or unusual values. Listing reviews of numeric values may also work for smaller studies to detect unusual values or outliers. For large studies, summary reports created from ad hoc queries or simple statistics performed on the data can identify unusual patterns or outliers by looking at the:

- Number of records or values per patient
- Highest, lowest, and mean for numeric values
- Distribution of values for coded fields (e.g., how many of each code)
- Amount of missing data

These summary reviews can be run by data management staff but in some companies, clinical programmers will look at the data using Statistical Analysis System (SAS). Graphs of lab and efficacy data or other simple displays or analyses can also identify possible problems

with units, decimal places, and different methods of data collection that might not otherwise be caught by simple cleaning checks. These will probably come out of the programming or statistical group. In the end, the best review of the data is to run the planned analysis programs on the data even before it.

c) Reconciling

In the best case, clinical data is stored in a single location and extracted for review or analysis as needed. However, in the setting of a drug or device trial, it is not unusual for the data to be stored in more than one location, and for very good reasons. When this is true, reconciliation may be necessary to ensure consistency between the systems.

The most common reconciliation with external systems is for serious adverse events (SAEs). Data on SAEs are typically stored in both the clinical data management system (CDM) and also in a separate SAE system. When reconciling at study close, data management staff look for:

- Cases found in the SAE system but not in the CDM system
- Events found in the CDM system but not in the SAE system
- Deaths reported in one but not the other — perhaps because of updates to the SAE report
- Instances where the basic data matched up but where there are differences, such as in onset date.[3]

COMMON AUDIT FINDINGS

Training

- Staff appropriately not trained [especially in relating to how GCP impacts on their role].

General Data Management /Analysis /Reporting

- Lack of formalized process to control management, analysis and reporting of trial data.

Data Collection CRF/Ecrf

- Data collected in the CRF (eCRF) doesn't meet the requirements of the protocol?
- CRF not reviewed appropriately.
- Functionality of the eCRF cannot be assured.

eCRF

- Data entered by unauthorized person.

Database Design and Maintenance

- the data base(s) (or simple spreadsheet) used for assimilation of the data capable of collecting all the CRF/trial data are not appropriate.

Data Entry and Verification

- Inaccurate electronic data (with respect to the paper CRF and other databases).
- Inaccurate transfer of other data (e.g. laboratory) into clinical database/stats analysis package.
- Changes to the data in the database after initial entry not controlled.
- A Person made changes to the data in the database/CRF who is not authorized by the investigator?
- Subject confidentiality is not maintained.

Data Validation

- Inappropriate persons reviews and approves the validation specification.
- Specification subject to change control not defined.
- Validation programming (where used) not validated and how is this documented.
- Validation programming (where used) validated but not documented.

Data Coding

- Queries have not been raised to change keywords to affect MedRA coding.

Data Transfer/Release

- Data base locked before completion of all data management activities.
- Final data made available (e.g. passed to the statistician) for analysis before data base lock.
- Database errors identified post lock resolved but not documented.

Data Quality

- No confidence that cross checks from the data listing in the Clinical Study Report or the Case Report Form will be verifiable with source documents?
- No evidence that the source data verification was undertaken by monitors.
- Electronic source documents available are not reliable.
- Data in the clinical study report (CSR) /Common Technical Document is not accurate.
- Protocol and GCP deviations are not captured in the CSR adequately.

- Accuracy of data to be used for dose escalation decisions is not ensured appropriately.

Statistics & Data Analysis

- No statistical inputs into the protocol (i.e. trial design) and there is no QC check of the sample size.
- Randomisation not produced appropriately.
- Uncontrolled distribution of Randomisation produced.
- No definite procedure for checking analysis and programming
- Analysis populations decided with bias.
- Lack of audit trail in the statistical analysis.
- Proper documentation and an audit trail haven't maintained with sufficient justification for updating the locked database.

Interim Analysis

- Protocol not followed.
- Data monitoring committee not established prior to trial commencement.

Documentation and Trial Master File

- Data management and statistical processes used in a trial cannot be reconstructed from documentation.
- Documentation not available for Data management and statistical processes used.

Data Security

- According to the roles and responsibilities multiple user IDs are not created with access limitation to data entry, medical coding, database designing, or quality check/Assurance.
- Database tools haven't built-in compliance with regulatory requirements.
- Adequate procedures and controls are not in place to ensure the confidentiality of data.
- Adequate Change controls and System Control procedures are not in place.
- Non-maintainace of data integrity when changes made to the computer system, such as software upgrades, security and performance patches, equipment repairs, etc.

DATA BASE AUDIT CHECK LIST

DESCRIPTION	YES	NO	NA	REMARKS
TRAINING AND COMPETENCE				
Have standard SOPs that must be followed.				
Training Records exists for all team members				
Training Records are Upto date				
COMPUTER SYSTEMS	**YES**	**NO**	**NA**	**REMARKS**
Internal Security				
According to the roles and responsibilities multiple user IDs are created with access limitation to data entry, medical coding, database designing, or quality check/Assurance.				
Each user can access only the respective functionalities allotted to that user ID and cannot make any other change in the database.				
All user accounts are Password-protected individual accounts.				
Provided with Automatically limit number of failed login attempts.				
Provided with Automatically record unauthorized login attempts.				
Provided that Electronically require users to change their passwords at regular intervals.				
Provided that Automatically passwords protect computer systems when idle for short periods.				
Provided that Automatically log users off computer systems when idle for long periods.				
External Security				
Restricted access to computer system and data via external software applications by encrypting data as it is transferred and/or using a firewall.				
Ensured to Prevent, detect and mitigate effects of viruses and other harmful software code.				
Adequate procedures and controls are in place to ensure the confidentiality of data.				
Change controls				

Maintain ing data integrity when making changes to the computer system, such as software upgrades, security and performance patches, equipment repairs, etc.				
Carefully evaluate effects of any changes before and after making them.				
Validated changes that exceed previous operational limits.				
Documented all computer system changes.				
System Controls				
Full backup and recovery system s are in place to protect against data loss if records are maintained only in electronic form.				
Ensured that a backup system maintains data integrity.				
Stores backup records at a secure offsite facility.				
Maintains backup and recovery logs.				
TOOLS	**YES**	**NO**	**NA**	**REMARKS**
Database Tools have built-in compliance with regulatory requirements and are easy to use				
User requirements and regulatory compliance are evaluated before implementation.				
Documented types of Softwares /database tools used for each project				
Documented Hardware used for Data Capture/Input				
If data have to be submitted to US regulatory authorities, it should be entered and processed in 21 CFR part 11-compliant systems.				
CLINICAL DATABASE DESIGN	**YES**	**NO**	**NA**	**REMARKS**
Study details like objectives, intervals, visits, investigators, sites, and patients are defined in the				
CRF layouts are designed for data entry and these entry screens are tested with dummy data before moving them to the real data capture				
Writes test plans and test data for data entry screens and data tables to ensure proper data storage.				
Reviews and tests the database designed and entry screens with regard to completeness and Usability.				
Tools used ensure the audit trail and help in the management of discrepancies.				
“System validation” is conducted to ensure data security, during which system specifications, User requirements and regulatory compliance are evaluated before implementation.				
PROCESSING LOCAL LAB DATA	**YES**	**NO**	**NA**	**REMARKS**
Managed local laboratory data including the specification and checking of normal ranges and units and QC of ranges and units against listings from the patient database.				
Generated standard reports of missing local laboratory data from the patient database.				
Defines the types of edit checks against the data.				
Runs edit checks against the standardized data.				
Develops data base specifications, collecting, processing and reporting for lab data.				
Procedures for Processing External Data (loading/merging) Central Lab Data or Other Electronic Data available				
CRF DATA ENTRY PROCESS- DISCREPANCY RESOLUTION	**YES**	**NO**	**NA**	**REMARKS**
Data entry takes place according to the guidelines prepared along with the DMP.				
Double data entry is performed wherein the data is entered by two operators separately				
Ongoing quality control of data processing is undertaken at regular intervals during the course of CDM				
Discrepancy resolution process is available (includes AE-subject data)				
Creates reports to track data entry process.				
DATA QUERY PROCESSING & TRACKING	**YES**	**NO**	**NA**	**REMARKS**
Maintains process, SOPs and standards for Query Resolution and query tracking.				
Maintains tracking/inventory and identify outstanding queries.				
Categorizes queries by age/site and regenerates if necessary.				
Tracks down and retrieves outstanding queries.				

Generates and interprets standard reports of query status, in support of standardized metrics.				
Relates queries per site to additional training requirements at site to attempt to reduce query needs.				
Creates standard reports to efficiently identify outstanding queries, query types per site, etc.				
Procedures available to Relate queries to processes and activities (e.g. CRF design) requiring process improvements.				
DATA VALIDATION	**YES**	**NO**	**NA**	**REMARKS**
Documents and maintains data validation process, SOPs and standards				
Applies standard data validation techniques, software and guidelines to routine data cleaning activities				
Identifies area of manual review where electronic checks are not effective.				
Initiates automated methods to minimize manual review				
Relates elements of protocol to defining data validation checks.				
Writes clear, concise queries.				
Generates queries based on standard data cleaning practices.				
Defines standards for written queries and query process.				
Developed a standard query language.				
COMMUNICATION OF DATA TRENDS	**YES**	**NO**	**NA**	**REMARKS**
Documents and maintains process, SOPs and standards for identifying signals and trends in data				
Uses established guidelines to identify and communicate trends to date.				
Reviews data for safety or efficacy at aggregate and site levels and identifies clear trends or outlier values and summarizes results on time.				
Executes standard reports of trends in clinical data.				
Oversee design and specification of project specific clinical data and status reports.				
DATABASE UPDATES	**YES**	**NO**	**NA**	**REMARKS**
Follows up on query responses, errors identified in data cleaning by performing accurate database updates.				
Documents database changes in the automated system audit trails and paper/electronic documentation.				
Developed and maintains process, standards and SOPs for performing database updates.				
Reviews audit trails, database change rates to assure staff expertise, extraordinary problems with CRF design, investigator site training, database screen design, etc.				
Proper documentation and an audit trail have been maintained with sufficient justification for updating the locked database.				
SAE RECONCILIATION	**YES**	**NO**	**NA**	**REMARKS**
Reconciles clinical databases' adverse events with serious adverse event reporting databases according to guidelines.				
Understands and implementing the SAE reconciliation process as per SOP				
Documents the outcome of the reconciliation process clearly and consistently.				
Relates different or similar medical terms/conditions in order to reconcile information presented in different text/coding terms from different systems.				
Understands the data that SAEs are reconciled against and make decisions on what to query.				
SAFETY REVIEW	**YES**	**NO**	**NA**	**REMARKS**
Identifies safety issues/trends for the study based on clearly defined guidelines and a review of all clinical trial data.				
Understands safety profile of the drug under study and disease state of patients.				
Communicates safety trends to project team.				
CODING (AEs: SIGNS AND SYMPTOMS)	**YES**	**NO**	**NA**	**REMARKS**
Available and familiar with all standard adverse event dictionaries, e.g. MedDRA.				
Utilizes available tools, systems and processes in support of the coding of medical terms				

Can manually code adverse events, when/if required.				
Creates an adhoc listing of coded adverse events for clinical review of pointing/mapping.				
Understands drugs dictionary drug classes and what they mean.				
Creates an ad hoc listing of medications.				
Can manually encode medication data, when/if required.				
Identifies all standard drug dictionaries, e.g. WHODRUG.				
Creates adhoc listings of medications for clinical review of pointings/mappings.				
DATABASE LOCK PROCEDURES	**YES**	**NO**	**NA**	**REMARKS**
Ensures that all steps preparatory to locking are accomplished, e.g.: patients are received, unverified records resolved, all edit checks are resolved, all lab data loaded, all data clarification queries have been returned from sites, all coding complete.i.e after a proper quality check and assurance, the final data validation is run for database locking.				
The SAS datasets are finalized in consultation with the statistician.				
Ensured that Data management activities have been completed prior to database lock.				
Once the approval for locking is obtained from all stakeholders, the database is locked and clean data is extracted for statistical analysis.				
Documents the locking procedure followed and any deviations from it.				
Establishes and coordinates the timely completion of the database lock procedures.				
Understands and follows the process if the database needs to be unlocked.				
In case of a critical issue or for other important operational reasons, privileged users can modify the data even after the database is locked.				
Informs team of database locking timelines/issues.				
Develops and maintains process and SOPs standards relevant to database locking and unlocking				
Data extraction is done from the final database after locking.				
ARCHIVING DATABASE AND ASSOCIATED INFORMATION	**YES**	**NO**	**NA**	**REMARKS**
Understands and applies the processes and standards relevant to database archiving as defined by the SOPs				
Performs archiving of case report forms that follow study or company procedures or any FDA regulations.				
Communicates timelines, retention requirements, archiving process, and access rights to DBA.				
DISASTER RECOVERY	**YES**	**NO**	**NA**	**REMARKS**
Disaster Recovery Plan (DRP) maintained in a separate and secure location.				
Backup and recovery procedure available for computer system components and data				
Verification Plan (all system components are in place, the data was restored, and the system is ready for production use)				
Process includes an analysis of potential lost or corrupted data and the actions to be taken to resolve any issues.				
System hardware inventory exists				
System software/application inventory exists				
Disaster Recovery Plan is tested and documented				
DATABASE QUALITY CONTROL	**YES**	**NO**	**NA**	**REMARKS**
Understands and applies quality acceptance criteria.				
Distinguishes critical from non-critical errors.				
Performs database audit and generates report as per guidelines.				
Documents the details related database QC performed				
Established quality criteria and quality error rates acceptance limits.				

INSIGHT ON STANDARDS FOR GOOD CLINICAL PRACTICE COMPLIANCE

Clinical Data Management Systems (CDMS) are used more and more to handle the increasing amount of data that must be collected, processed and analysed in clinical research, whether that data is initially captured remotely and directly from clinical sites using Remote Data Capture (RDC), or using more traditional paper based methods. There exists no widely recognized, specific, practicable and open standard for GCP-compliant data management and the accompanying IT infrastructure. To expand upon the last point: GCP requirements on data management are mostly unspecific at the technical level[4]. EU Directive 2001/20/EC[5], EU Directive 2005/28/EC[6] and Annex 11[7] define GCP compliance for clinical trials but specify only a few technical requirements for data management (e.g. necessity for data privacy, security system, system descriptions). The FDA Guidance for Computerized System Used in Clinical Trials[8] or 21 CFR Part 11[9] covering electronic records and electronic signatures are legally binding in the US but have less relevance for the EU and other regions. This regulation is applicable to records in electronic format that are created, modified, maintained, archived, retrieved, or transmitted. This demands the use of validated systems to ensure accuracy, reliability, and consistency of data with the 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. Adequate procedures and controls should be put in place to ensure the integrity, authenticity, and confidentiality of data.

Similarly, specific regulations exist in many EU countries and several national guidance documents for IT are available (e.g. UK, Germany, Denmark) but with limited or no relevance for other countries.[10, 11, 12]

For computer system validation purposes a number of additional guidelines are in use for specific aspects of data management, like the PIC/S Guide,[13] which defines requirements from the inspectors' point of view and the Good automated manufacturing practice (GAMP®) guide[14] defining best practices for system validation. On the other hand, ISO standards cover only the general level of IT infrastructure aspects (e.g. ISO 27001, security management system).[15]

Society for Clinical Data Management (SCDM) publishes the Good Clinical Data Management Practices (GCDMP) guidelines, a document providing the standards of good practice within CDM. GCDMP was initially published in September 2000 and has undergone several revisions thereafter. The July 2009 version is the currently followed GCDMP document. GCDMP provides guidance on the accepted practices in CDM that are consistent with regulatory practices. Addressed in 20 chapters, it covers the CDM process by highlighting the minimum standards and best practices. Members of the Society of Clinical Data Management (SCDM) can download the guide, non-members may purchase the copyright protected document.[16]

Clinical Data Interchange Standards Consortium (CDISC), a multidisciplinary non-profit organization, has developed standards to support acquisition, exchange, submission, and archival of clinical research data and metadata. Metadata is the data of the data entered. This includes data about the individual who made the entry or a change in the clinical data, the date and time of entry/change and details of the changes that have been made. Among the standards, two important ones are the Study Data Tabulation Model Implementation Guide for Human Clinical Trials (SDTMIG) and the Clinical Data Acquisition Standards Harmonization (CDASH) standards, available free of cost from the CDISC website (www.cdisc.org). The SDTMIG standard[17] describes the details of model and standard terminologies for the data and serves as a guide to the organization. CDASH v 1.1[18] defines the basic standards for the collection of data in a clinical trial and enlists the basic data information needed from a clinical, regulatory, and scientific perspective.

In summary, there is no standard for GCP-compliant data management and the underlying IT infrastructure available, which is both generally applicable and practical, as well as being open and available free of charge.

REFERENCES

1. Audit of an Investigator Site - A Crucial Task in Clinical Research to ensure a Reliable Clinical Trial: Review of Planning, Methodology and Techniques, Raju Sama, Vijaya Lakshmi N and Sathish M. International Journal of Research and Development in Pharmacy and Life Sciences. 2016;5(4):13-29.
2. Data management in clinical research: An overview, Binny Krishnankutty, Shantala Bellary and Naveen BR. Kumar, and Latha S.

Moodahadu, Indian Journal of Pharmacology. 2012;44(2):168–172.
3. Practical Guide to Clinical Data Management, Second Edition, by Susanne Prokscha, CRC Press, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742:109-110.
4. ICH Topic E 6 (R1): Guideline for Good Clinical Practice. CPMP/ICH/135/95, EMEA London. 1996, /2002
5. Directive 2001/20/EC. Official Journal of the European Communities, L121/34-121/44.
6. Directive 2005/28/EC. Official Journal of the European Communities L 91/13-91/19.
7. Annex 11 - Computerized Systems. EudraLex - Volume 4: Guidelines for good manufacturing practices for medicinal products for human and veterinary use. Bruxelles 31. 2008, October
8. FDA: Guidance for Industry: Computerized Systems Used in Clinical Trials. 1999.
9. FDA: 21 CFR Part 11, Electronic Records; Electronic Signatures. Final Rule, Federal Register. 1997;62(54):13429.
10. Lauritsen JM: Implementation of Good Clinical Practice Software. Draft, University of Southern Denmark, Denmark. 2007.
11. KKS Network: Policy Document. University of Cologne, Germany, 2001.
12. DIMS Project Team: Trials Units Information Systems - System Standards. Data and Information Management Systems Project. UKCRC/NIHR. 2009.
13. Pharmaceutical Inspection Convention: PIC/S Guidance: Good Practices for Computerised Systems in regulated "GXP" environment. PIC/S. 2007, September , Geneva, Switzerland
14. GAMP® 5: A Risk-Based Approach to Compliant GxP Computerized Systems. International Society for Pharmaceutical Engineering. February , Tampa, Fl., USA. 2008.
15. ISO 27001-2005 - Information technology -- Security techniques -- Information security management systems -- Requirements.
16. Good Clinical Data Management Practice. Version 4, Society for Clinical Data Management. 2009.
17. Study Data Tabulation Model [Internet] Texas: Clinical Data Interchange Standards Consortium.c2011. Available from: http://www.cdisc.org/sdtm.
18. CDASH [Internet] Texas: Clinical Data Interchange Standards Consortium. c2011.