ISSN 2581-6217

World Journal of Pharmaceutical Science & Technology

Journal homepage: www.wjpst.com

Review Article

CLINICAL DATA MANAGEMENT: FACTORS INFLUENCING QUALITY DATA; REGULATIONS, GUIDELINES AND STANDARDS; CDM PROCESS.

Komal Sapkal[1], Ms. Archana Gawade[2]

1. B.Pharm Scholar, Rajarambapu college of Pharmacy, Kasegaon and Advance Diploma in Pharmacovigilance and clinical research Scholar from Elite institute of Pharma Skills Pune.
2. Managing Director, Elite Institute of Pharma Skills, Pune

Address for correspondence:

Komal Sapkal, Rajarambapu college of Pharmacy, Kasegaon and Advance Diploma in Pharmacovigilance and clinical research Scholar from Elite institute of Pharma Skills Pune.
E-mail- komalsapkal20@gmail.com

Received: 15-3-2022, Revised: 29-4-2022, Accepted: 5-6-2022

ABSTRACT

Clinical data management (CDM) is an indispensable part of clinical research. CDM activities lead to the collection of reliable, high-quality and statistically sound data generating from the clinical trials. Clinical trial data are important for the drug and medical device development processing pharmaceutical companies to examine and evaluate the efficacy and safety of the new medical product in human volunteers. The results of the clinical trial studies generate the most product. Several practices in CDM including CRF annotation, case report form (CRF) designing, data extraction, data entry, data validation, database designing, database locking, discrepancy management and medical coding are evaluated for quality checks at regular intervals during clinical trial.

KEYWORDS: Clinical trial, Clinical data management, Data capture, Clinical Management Process, Factors influencing quality data,

INTRODUCTION

Clinical studies are planned to augment the current medical knowledge inter-related with the treatment, diagnosis, and prevention of diseases or conditions. Drug discovery is a too lengthy, expensive, and most

difficult process. Drug discovery requires identification of drug compound, screening, and assays for therapeutic safety and efficacy. Clinical trials are research studies that explore whether a medical drug treatment or device is safe and effective for humans. Before the clinical trial, the sponsor research may include pharmacodynamics, pharmacokinetics, absorption, distribution, metabolism and excretion studies, and toxicity testing of the test article in animals (pre-clinical studies). During pre-clinical studies, in vitro and in vivo testing is performed and based on the result of pre-clinical studies; the sponsor research starts their clinical trials in human volunteers. Hence, the studies follow strict scientific standards and these standards protect patients and help produce reliable study results.

It has become impossible to perform clinical research without the utilization of CDMS. The use of CDMS has become essential in clinical trials to manage the increasing amount of clinical data. [1] In fact, the quality of the data validation process has a direct influence on the quality of data presented as part of new drug application (NDA) submission. High quality data is characterized by maintaining the number of errors and missing data to the lowest possible limit and obtaining maximum data for analysis. [2]

FACTORS INFLUENCING THE QUALITY OF DATA

There are several factors that have a serious impact on the overall quality of data collected.

1. Case report form (CRF) design

CRFs needs to be thoroughly prepared in order to collect complete and accurate data. Both the protocol and the CRF need to be designed in parallel to establish consistency between the two.

2. Field monitoring guidelines

The quality of field monitoring guidelines has a direct association with the quality of data presented to the CDM.

3. Source data verification (SDV)

Source data verification is one of the significant phases of the data validation process, without which the integrity and quality of data may get affected.

4. Data conventions

In the multicentre clinical trials, varying date conventions being followed by several investigators may pose difficulties and problems when it comes to entering the data on the database.

REGULATIONS, GUIDELINES AND STANDARDS IN CDM

Since the pharmaceutical industry relies on the electronically captured data for the evaluation of medicines, there is a need to follow good practices in CDM and maintain standards in electronic data capture. These

electronic records have to comply with a Code of Federal Regulations (CFR), 21 CFR Part 11. This regulation is applicable to records in electronic format that are created, modified, maintained, archived, retrieved, or transmitted. [3] Adequate procedures and controls should be put in place to ensure the integrity, authenticity, and confidentiality of data. If data have to be submitted to regulatory authorities, it should be entered and processed in 21 CFR part 11-compliant systems. Most of the CDM systems available are like this and pharmaceutical companies as well as contract research organizations ensure this compliance. Clinical data interchange standards consortium (CDISC) is global, open, multidisciplinary, non-profit organization that has established standards to support the acquisition, exchange, submission and archival of clinical research data. The two most important standards established by CDISC are study data tabulation model implementation guide for human clinical trials (SDTMIG) and the clinical data acquisition standards harmonization (CDASH) standards. The SDTMIG standard functions as a guide to the organization and summarizes the details of model and standard terminologies for the data.6 Whereas, CDASH enlists the basic data information required from a scientific, regulatory clinical and perspective. It also outlines the standards for the collection of data in a clinical trial. [4]

CDM Process:

The CDM process, like a clinical trial, begins with the end in mind. This means that the whole process is designed keeping the deliverable in view. 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.

Review and finalization of study documents

During this review, the CDM personnel will identify the data items to be collected and the frequency of collection with respect to the visit schedule. A Case Report Form (CRF) is designed by the CDM team, as this is the first step in translating the protocol-specific activities into data being generated. The data fields should be clearly defined and be consistent throughout. The type of data to be entered should be evident from the CRF. The CRF should be concise, self-explanatory, and user-friendly (unless you are the one entering data into the CRF). Along with the CRF, the filling instructions (called CRF Completion Guidelines) should also be provided to study investigators for error-free data acquisition. CRF annotation is done wherein the variable is named according to the SDTMIG or the conventions followed

List of clinical data management activities

a) Database design

b) Data collection

c) CRF tracking

d) Data entry

e) Data validation

f) Discrepancy management

g) Medical coding

DATABASE DESIGNING

Databases are the clinical software applications, which are built to facilitate the CDM tasks to carry out multiple studies. [5] Generally, these tools have built-in compliance with regulatory requirements and are easy to use. "System validation" is conducted to ensure data security, during which system specifications, [1] user requirements, and regulatory compliance are evaluated before implementation. Database designing and design of CRF show positive correlation. The eCRF facilitates the entry of data into an underlying relational database. For a clinical trial employing a paper CRF, the relational database is built separately. In both cases, the relational database allows entry of all data captured on the CRF.

DATA COLLECTION

Data collection is done using the CRF that may exist in the form of a paper or an electronic version. The traditional method is to employ paper CRFs to collect the data responses, which are translated to the database by means of data entry done in-house. Data collection, also called as data acquisition can be accomplished by the use of paper or electronic medical records, interactive voice response systems, local electronic data capture systems or central web-based systems. These paper CRFs are filled up by the investigator according to the completion guidelines. The ICH guidelines on GCP utilize the term case report form‟ or „CRF‟. A CRF is a data collection tool used in clinical trials to support investigators and coordinators in capturing all protocol-required information. In the e-CRF-based CDM, the investigator or a designee will be logging into the CDM system and entering the data directly at the site. A well-designed CRF facilitates data collection and entry, and directly benefits other facets of data management and statistical analysis. An informative and structured CRF simplifies database design and data validation processes as well as manipulation of data during statistical analysis.

CRF TRACKING

A CRF tracking system is essential to ensure that all the required CRFs are collected. The CDM team will track the retrieved CRFs and maintain their record. CRFs can also be tracked for missing pages and illegible data manually to assure that the data are not lost. In case of missing or illegible data, a clarification is obtained from the investigator and the issue is resolved. The entries documented in the CRF will be examined by the

clinical research associate (CRA) for completeness. These CRFs are then handed over to the CDM team. The CDM team will track the retrieved CRFs and maintain their record.

DATA ENTRY

Data entry takes place according to the guidelines prepared along with the DMP. This is applicable only in the case of paper CRF retrieved from the sites. Usually, double data entry is performed wherein the data is entered by two operators separately.[6] These are two types of data entry, single data entry and double data entry. In earlier days, the data entry is in double data entry due to the usage of pCRF. Nowadays, electronic CRFs are increasingly chosen by investigators and sponsors of clinical research instead of the traditional pCRFs [7] due to the data entry process is easy and less duration for data entry. The second pass entry (entry made by the second person) helps in verification and reconciliation by identifying the transcription errors and discrepancies caused by illegible data. Moreover, double data entry helps in getting a cleaner database compared to a single data entry. Earlier studies have shown that double data entry ensures better consistency with paper CRF as denoted by a lesser error rate. [8]

DATA VALIDATION

Data validation helps to develop clinical trial databases that are comprehensive and superior in quality to meet the study's objectives and comply with regulatory standards. One of the common procedures adopted to validate a clinical trial database is, batch validation. In batch validation, one executes a series of checks that are developed to validate the clinical trial database. Discrepancy is defined as a data point that fails to pass a validation check. Discrepancy may be due to inconsistent data, missing data, range checks, and deviations from the protocol. In e-CRF based studies, data validation process will be run frequently for identifying discrepancies. Edit check programs are written to identify the discrepancies. Edit checks are consisting manual and computer checks. Edit checks detect the data entry issues and errors. The main errors are typography errors, copying errors, coding errors, and range errors. Data cleaning is the process to fixed invalid data and missing data and send queries to the investigator. The main work of the discrepancy management team is to identify the errors and cleaning data. Data cleaning as a three-stage process, involving repeated cycles of screening, diagnosing, and editing of suspected data abnormalities [9].

DISCREPANCY MANAGEMENT

This is also called query resolution. Discrepancy management is a process of cleaning subject data in the clinical data management system (CDMS). Discrepancy management includes reviewing discrepancies, investigating the reason, and resolving them with documentary proof or declaring them as irresolvable. In some cases, it also includes reviewing the discrepancies, scrutinizing the reason, and concluding them along with the documentary evidence or stating them as unworkable. Majority of the CDMS contain a discrepancy database including all the discrepancies that will be recorded and stored with audit trail. The CDM team

reviews all discrepancies at regular intervals to ensure that they have been resolved. The resolved data discrepancies are recorded as 'closed'. This means that those validation failures are no longer considered to be active, and future data validation attempts on the same data will not create a discrepancy for same data point. But closure of discrepancies is not always possible. In some cases, the investigator will not be able to provide a resolution for the discrepancy. Such discrepancies will be considered as 'irresolvable' and will be updated in the discrepancy database. Discrepancy management is the most critical activity in the CDM process. Being the vital activity in cleaning up the data, utmost attention must be observed while handling the discrepancies.

MEDICAL CODING

Medical coding helps in identifying and properly classifying the medical terminologies associated with the clinical trial. In multi-centric clinical trials, several investigator or medically qualified experts are from different sites/centres and therefore recording the medical term(s) in a uniform manner is a big challenge. Therefore, medical coders from CDM team process these terms and perform medical coding. The concomitant medications, medical history adverse event, serious adverse event data are generally coded using medical dictionaries. These medical dictionaries are available in online and used in all the phases of clinical trial. There are several standardized medical coding dictionaries are available for coding. Coding symbols for a thesaurus of adverse reaction terms, International Classification of Diseases 9 Revision Clinical Modification, medical dictionary for regulatory activities (MedDRA), World Health Organization (WHO) Adverse Reactions Terminology, WHO Drug Dictionary Enhanced (WHO-DDE). [10]

There are several standardized medical coding dictionaries; however the following five dictionaries are frequently used for medical coding: COSTART- coding symbols for thesaurus of adverse reaction terms, Mudra- medical dictionary for regulatory activities, WHO-DDE- world health organization drug dictionary enhanced, ICD9CM International classification of diseases 9 revision clinical modification and WHO-ART- World health organization adverse reactions terminology.[11] The available medical dictionaries assist in coding the adverse events that may occur during the study, prior to and concomitantly administered medications and pre-or co-existing illnesses. Medical coding helps to achieve the data consistency and avoid unnecessary duplication.

DATABASE LOCKING

Database lock is the process that ensures no changes or manipulation of data during the analysis. Database locking is done after review, query resolution, and a determination has been made that the database is ready for analysis. After a proper edit check, no discrepancies the database and cleaning data, the database will be locked. Database lock has two types, Interim lock and final lock. Interim lock is generally done at any time in between the study, and the final lock is done when all data entry is completed. To ensure this, a pre-lock

checklist is used and completion of all activities is confirmed. Once the database is locked no modification, corrections, updates in the database. Data extraction is done from the database after locking. Only in case of a critical issue or any crucial operational reasons, privileged clients can access and amend the data even after the database is locked. However, this entails absolute documentation and an audit trail to be maintained with satisfactory explanation for updating the previously locked database. Database locking is followed by data archival.

DATA ARCHIVING

It is the long-term storage of all essential documents which individually and collectively permit the evaluation of the conduct of a clinical trial and the quality of the data produced. Archiving of clinical trial data must be carried out in compliance with the EU clinical trials directive (2001/20/EC), Volume 10 of Eudralex-the rules governing medicinal products in the European Union, International conference for harmonization-good clinical practice (ICH-GCP) guidelines (CPMP/ICH/135/95) and GCP directive. Essential documents must be archived for sufficient periods to allow for audit and inspection by regulatory authorities.

Data privacy collection, processing and transfer of personal data, in particular individuals‟ patient records, in compliance with applicable laws are vital to the success of clinical studies. According to Article 3(2) c of the EC clinical trials directive3, a trial may be undertaken only if "the rights of the subject to physical and mental integrity, to privacy and to the protection of the data concerning him in accordance with directive 95/46/EC are safeguarded. [12]

DATA QUALITY CONTROL AND QUALITY ASSURANCE

Both the ICH GCP and EU GCP guidelines state that the "quality control must be applied to each stage of data handling". The CDM process is quite complicated and can involve many people and multiple systems. It is important, therefore, to have an effective, quality

Controlled system so that the process runs smoothly and efficiently. Audits help to ensure that the CDM process operate effectively and conducted to GCP. [12]

CONCLUSION

CDM has evolved in response to the ever-increasing demand from pharmaceutical companies to fast-track the drug development process and from the regulatory authorities to put the quality systems in place to ensure generation of high-quality data for accurate drug evaluation. CDM is a multidisciplinary endeavour that involves the handling of data or information. Clinical trials industry also faces the mounting pressure of planning, monitoring and implementing data management systems in a vibrant clinical trials research arena where the swift strides towards state-of-the-art technology makes the existing infrastructure redundant. Regardless of the challenges, data management in clinical research is rapidly developing into a standardized

system. Developments on the technological front have positively impacted the CDM process and systems, thereby leading to encouraging results on speed and quality of data being generated. At the same time, CDM professionals should ensure the standards for improving data quality. [14] CDM, being a speciality in itself, should be evaluated by means of the systems and processes being implemented the standards being followed. The biggest hurdle would be the planning and implementation of data management systems in a changing operational environment where the rapid pace of technology development outdates the existing infrastructure. Regardless of the challenges, data management in clinical research is rapidly developing into a standardized system. [15]

REFERENCES

1. Kuchinke W, Ohmann C, Yang Q, Salas N, Lauritsen J, Gueyffier F, et al. Heterogeneity prevails: The State of Clinical Trial Data Management in Europe - results of a survey of ECRIN. Trials. 2010; 11:79.
2. Gerritsen MG, Sartorius OE, vd Veen FM, Meester GT. Data management in multi-center clinical trials and the role of a nation-wide computer network. A 5 year evaluation. Proc Annu Symp Comput Appl Med Care.1993:659-62.
3. CFR - Code of Federal Regulations Title 21 [Internet]. Maryland: Food and Drug Administration; Available from: http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/ cfcfr/CFRSearch.cfm?fr=11.10. [Updated 2010 Apr 4; Cited 2011 Mar 1].
4. CDASH. Texas: Clinical Data Interchange Standards Consortium. c2011. Available at http://www.cdisc.org/cdash. Accessed on 8 October 2016.
5. Fegan GW, Lang TA. Could an open-source clinical trial data-management system be what we have all been looking for? PLoS Med 2008; 5:e6.
6. Cummings J, Masten J. Customized dual data entry for computerized data analysis. Qual Assur 1994; 3:300-3.
7. Le Jeannic A, Quelen C, Alberti C, Durand-Zaleski I; CompaRec Investigators. Comparison of two data collection processes in clinical studies: Electronic and paper case report forms. BMC Med Res Methodol 2014; 14(1):7.
8. Reynolds-Haertle RA, McBride R. Single vs. double data entry in CAST. Control Clin Trials 1992; 13:487-94.
9. Van den Broeck J, Cunningham SA, Eeckels R, Herbst K. Data cleaning: Detecting, diagnosing, and editing data abnormalities. PLoS Med 2005; 2(10):e267.
10. Babre D. Medical coding in clinical trials. Perspect Clin Res 2010; 1(1):29.
12. Baber D. Medical Coding in Clinical Trials. PerspexClan Res. 2010; 1(1):29–32.
13. Elizabeth Ness, Clinical data management Centre for Cancer Research National Cancer Institute. Available at https://ccrod.cancer.gov/confluence/ display/CCRCRO/CCR+Orientation. Accessed on 15 October 2016.
14. Rondel RK, Varley SA, Webbs CF. Clinical Data Management. 2nd edition. Wiley Publications; 1993.
15. Haux R, Knaup P, Leiner F. On educating about medical data management - the other side of the electronic health record. Methods INF Med 2007; 46:74-9.