Data Management in Clinical Research

Article by Akinremi Nanna
Msc. Clinical Research, Texila American University
Email:- oritseweyimiakinremi@gmail.com

Abstract

Clinical data management helps to produce a drastic reduction in time from drug development to marketing. Team members of CDM are actively involved in all stages of clinical trial right from inception to completion. They should have adequate process knowledge that helps maintain the quality standards of CDM including case report form (CRF) designing, CRF annotation, data designing, data-entry, data validation, discrepancy management, medical coding, data extraction and data locking are assessed for quality at regular intervals during the trial. Presently 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. With the implementation of regulatory complaint data management tools. CDM team can meet these demands. Additionally, it is becoming mandatory for companies to submit data electronically. It is advocated that CDM professionals should meet appropriate expectations and set standards for data quality and also have a drive to adapt to the rapidly changing technology. Pls refer Binny Krishnankutty, Shantala Bellary, Naveen B. R. Kumar et al 2011 Data management in clinical research an overview Indian Journal of pharmacology 44(2):168-172. doi 10. 4103/0253-7613.93842

Definition

Data management is defined as a system for managing large amount of data ref Oxford advance learners dictionary.

While the encyclopedia Wikipedia defines clinical data management in clinical research as defined by the National Institute of Health (NIH) as the ultimate goal of clinical research data management (CRDM) is to ensure that data support the conclusion drawn from research. It also stipulates that clinical data management (CDM) is a critical phase in clinical research, which leads to generation of high – quality, reliable and statistically sound data from clinical trials. Clinical data management assures collection, integration and availability of data at appropriate quality and cost ref https://en.m.wikipedia.org/wiki/clinical-dat assessed on the 20th of June 2015

Introduction

The primary objective of clinical data management (CDM) is to provide high quality, reliable data for reporting randomized controlled trials (RCTs) in line with good clinical practice (GCP) requirements. In treatment trials of neglected Tropical diseases (NTDs) in endemic countries, CDM systems need to be efficient and AFFORDABLE (key issue), the challenges of poor infrastructures, license costs associated with GCP-complaint software and limited human resources to provide the required expertise are daunting, but it is argued that high quality CDM for NTDs can be achieved and the challenges can be overcome through the use of open-access tools ref Raymond Omollo, Michael Ochieny and Tansy Edwards 2014 innovative approaches to clinical Data management in resource limited settings using open source technologies PLOS Neglected Tropical diseases 2014 sen.8 (9). e3134.

Keywords: NEGLECTED TROPICAL DISEASES, Good clinical practice soft complaint software, clinical data management systems, data management e-CRF,

Good clinical data management practices validation, clinical data interchange standard

Discussion

Clinical data is intended to find answers to the research question by means of generating data for proving or disproving hypothesis Binny Krishnankutty 2011 The quality of data generated plays an important role in the outcome of the study Clinical data management is a relevant and important part of a clinical trial. All researchers try their hands on CDM knowingly or unknowingly. Clinical data management is the process of collection, cleaning and management of subjects data in compliance with regulatory standard, the primary objective of CDM processes is to provide high quality data by keeping the number of errors and missing data as low as possible and gather maximum data for analysis to meet the objectives best practices are adopted to ensure that data are complete, reliable and processed correctly. Pls refer Ptizer quality management in clinical trials 2009,030209

Components for generating clinical data

1) Creating, implementing, and upholding standard operating procedures (SOPs) for trial execution
2) A quality scientific and medical design of the protocol
3) Clinical investigation and site pre-assessment and selection
4) Regulatory agency and ethics committee approval
5) Developing and providing appropriate informed consent (Language transparency of benefits and risks) and obtaining ethics committee approval of the informed consent process
6) Investigators meetings and training
7) Adequate recording and reporting data
8) Periodic monitoring
9) Audits

Many soft ware tools are available for data management and these are called clinical data management systems (CDMS) used in pharmaceutical companies are commercial but a few open source tools are available as well as. The commonly used ones in pharmaceutical companies are mostly commercial but a few open source tools are available as well ref lu and Su publisher 2010 (Open access journal of clinical trials 2, 93-105 clinical data management current status, challenges and future directions from industry perspective). Commonly used CDM tools are oracle clinical, clintrial, macro, rare and e clinical suite, In terms of function ability these soft ware tools are expensive and need sophistical information technology infrastructure to function. Additionally some multinational pharmaceutical grants use custom made CDMS Tools to suite their operational needs and procedures. Among the open source sophisticated tools the most prominent ones are open clinical, open CDMS, Trial DB and phosco. These CDM software are available free of cost and are as good as their commercial counterparts in terms of function ability. These open source software can be downloaded from their respective web site –ref Binny Krishnankutty et al 2011

In regulatory submission studies maintaining an audit trail of data management activities is of paramount importance. These CDM tools ensure the audit trail and also help in the management of discrepancies. According to the roles and responsibility, multiple user IDs can be created with access limitation to data entry, medical coding, data designing or quality check. This ensures that each user can access only the respective function ability allotted to that user ID and cannot make any other change in the data base. For responsibility where changes are permitted to be made in the data. The software will record the change made, the user ID that made the change and the time and date of the change. For audit purposes (audit trail) During a regulatory

audit the auditors can verify the discrepancy management process, the changes made and can confirm that no unauthorized or false changes were made.

Regulatory, guidelines and standard in CDM

Akin to other areas in clinical research, CDM has guidelines and standards that must be followed since the pharmaceutical industry relies on the electronically captured data for the evaluation of medicines there is the 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 Regulatory (CFR) 21 CFR part 11. This regulations 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 data and time of operator entries and actions that create, modify, or delete electronic records (3) Adequate procedures and control 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 complaint systems. Most of the CDM systems available are like this in pharmaceutical companies as well as in contract research organizations which ensure this compliance ref Binny Krishnankutty 2011 Good clinical data management practices (GCDMP) provide guidance on the accepted practices in CDM that are consistent with regulatory practices by high lightening the minimum standard and best practices Clinical Data Interchange standard consortium (CDISC) a multidisplinary non-profit organization has developed standards to support acquisition exchange, submission and archival of clinical research data and metadata entered. This include 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 changes that have been made among the standards, two important aspects are the study data tabulation model implementation guide for Human clinical trials (SDTMIG) and the clinical data acquisition standard harmonization (CDASH) standards available free from cost from the CDISC website (www.cdisc.org). The SDTMIG Standards describe the details of model and standard technologies for the data and serves as a guide to the organization CDASH V 1.1 (5) defines the basic standard for the collection of data in a clinical trial and enlists the basic data information needed from a clinical regulatory and scientific perspective (CDISC).

The 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 document

The protocol is reviewed from the database designing perspective, for clarity and consistency.

During the 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 that is 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. For example if weight has to be captured in two decimal places the data entry field should have two data boxes placed after the decimal. Also the units of

measurements e.g. centimeters, meters or kilograms should be mentioned next to the data field The case report form CRF should be self explanatory, and user friendly. Along with the CRF the filling instructions called CRF Completion guidelines) should also be provided to study investigators for error free data acquisition. The CRF is done according to study data tabulation model implementation guide for human clinical trials (SDTMIG) or any other convention followed internally Based on this a data management plan (DMP) is developed.

DMP Document is a road map to handle data under foreseeable circumstances and describes the CDM activity to be followed in the trial. PLs refer CDISC clinical data interchange standards consortium

Database designing

Databases are the software application which are built to facilitate the CDM tasks to carry out multiple studies. Generally these tools have built in compliance mechanisms with regulatory requirements and are easy to use. System validation is conducted to ensure data security, during which system specification, user requirements and regulatory compliance are evaluated before implementation. Study details like objectives, interval visits of investigators, sites and patient are defined in the data base entry. These entry screenings are tested with dummy data before moving them to the real data capture ref Binny Krishnankutty 2011

Data collection and C2

Data collection is done using case report forms (CRF). The entries made in the CRF will be monitored by the clinical research associate CFR for completeness and filled up CFRs are retrieved and handed over to the CDM team. CRF are tracked for missing pages and illegible data are not lost.

In case of missing or illegible data, a clarification as obtained from the investigator and the issue is resolved.

Class C2

Definition –class 2 is a security rating established by the US National computer security centre (NCSC) and granted to products that pass Department of Defence (DOD) Trusted computer system Evaluation criteria (TCSEC) tests A C2 rating ensures the minimum allowable level of confidence demanded for Government agencies and offices and other organizations that process classified or secure information. TCSEC Standards were established in the 1985 DOD document Department of Defense trusted computer systems Evaluation criteria known unofficially as the “Orange Book “ Pls refer Margaret Rouse sep 2008 WhatLs.com and –Searchsecurity.com

Data entry and discrepancy management

Data entry takes place according to guidelines prepared along with the DMP. This is applicable only in the case of paper CRF retrieved from the sites. Usually double entry data is performed where in the data is entered by two operators separately. The second pass entry (entry made by the second person) helps in verification and reconciliation by identifying the transcription errors and discrepancies causes by illegible data. More over 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 DISCREPANCY MANAGEMENT –This is also called query resolution. Discrepancy management includes reviewing discrepancies investigating the reason and resolving them with documentary proof or declaring them as irresolvable. Discrepancy management helps in cleaning the data and gathers enough evidence for

deviations observed in the data. Almost all CDMS (clinical data management systems) have a discrepancy management base where all discrepancies are recorded and stored with audit trail

When discrepancies are found they are referred to investigators for clarification The CDM team reviews all discrepancies at regular intervals to ensure that they have been resolved. The resolved discrepancies are recorded as closed. Some of these may include spelling errors. Ref Binny Krishnankutty

Medical coding and database locking

Medical coding helps in identifying and properly identifying the medical terminology associated with the clinical trial. Classification of events, medical dictionaries available on line are used. Technically, this activity require needs the knowledge of medical terminology, understanding of disease entities, drugs used and a basic knowledge of the pathological processes involved. Functionally it also requires knowledge about the structure of electronic medical dictionaries and the hierarchy of classification available in them.

Commonly used medical dictionary for coding is the medical dictionary for regulatory activities (MEDDRA) is used for the coding of adverse events as well as other illnesses and World Health organization –drug dictionary Enhancement (WHO-DDE) is used for coding the medications, other dictionaries are WHO-ART. Some pharmaceutical companies customize dictionaries to suit their needs and meet their operating procedure

Medical coding helps in classifying reported medical terms on the CRF to standard dictionary terms in order to achieve data consistency and avoid unnecessary duplication. An investigator may use different terms for the same adverse event but it is important to code all of them to a single standard code and maintain uniformity in the process. The right coding and classification of adverse events and medication is crucial as an incorrect coding may lead to masking of safety issues or highlight the wrong safety concern related to the drug

After a proper quality check and assurance, the final data validation is run. If there are no discrepancies, the SAS (STANDARD ANALYSIS SYSTEM) Data sets are finalized in consultation with the statistician. All data management activities should have been completed prior to database lock. To ensure this, a pre-lock checklist is used and completion of all activities is confirmed. This is done as the database cannot be changed in any manner after locking once approval is obtained from all stake holders

SAS(statistical Analysis System) is a soft ware suit developed by SAS Institute for advance analysis, multivariate analysis, business intelligence, data management and predictive analysis Ref SAS Institute North Carolina state university 1966-1976 https://en.m.wikipedia.org/wiki/sas The data base is then locked and clean data is extracted for statistical analysis Generally no modification of the data base is possible but in case a critical issue or for other important operational reasons privileged users can modifier the data However in this contest privileged user has not been properly defined. Any adjustment will require proper documentation and an audit trail has to be maintained with sufficient justification for updating the Locked database. Data extraction is done from the final database after locking. This is followed by archival.

The roles of team members in CDM team

In the CDM team different roles and responsibilities are attributed to the team. The minimum educational requirement for a team member in CDM should be a graduate in life science and knowledge of computer application. Ideally medical coders should be medical graduates, however in the industry paramedical graduates are also

recruited as medical coders. Some key rolls are essential for all CDM Teams. The list of roles stated herein can be considered as minimum requirements for a CDM team.

Data manager
Database programmer/Designer
Medical coders
Clinical data coordinator
Quality control associate
Data entry associate

The data manager is responsible for supervising the entire CDM process. The data manager prepares the DMP, approves the CDM procedures and all internal documents related to CDM activities. Controlling and allocating the data base access to team members is also the responsibility of the data manager

Different professional organizations have outlines on clinical data management. There is the International Network of clinical data management (INCDMA) The INCDNA aims at the promotion of collaboration among clinical data management groups around the world. It is active International forum for discussion of and feedback on current topics of relevance to the discipline of CDM. It is composed of members of the boards of the SCDM, ACDM (UK), DMB (FRANCE), PSDM (THE NETHERLANDS) Who participate in the in the INCDMA proceedings and funding. It also regroup Dm leaders and subject matter experts from Europe, North America, Israel, Japan China and Australia pls refer https://www.acdm.org.uk and https://en.wikipedia.org/wiki/clinical-data-management

Conclusion

Clinical data management has evolved in response to the ever increasing demand from pharmaceutical companies to fast track the drug development process and from regulatory authorities to put quality systems in place to ensure generation of high quality data for accurate drug evaluation. To meet this expectation there is the graduate shift from the paper based to the electronic system of data management

Developments in the technological front have positively impacted on the CDM process and systems there by leading to encouraging results on speed and quality of data been generated

THE biggest challenge from the regulatory perspective would be in standardization of data management process across organizations and development of regulations to define the procedure to be followed and the data standards from industry perspective, the biggest hurdle would be the planning of data management systems in a changing operational environment where the rapid pace of technological development outdates the existing infrastructure. In spite of these CDM is evolving to become the standard based clinical research entity by striking a balance between the expectation from and constraints in the existing systems driven by technological development and business demand

References

[1]. Binny Krishnankutty, Shantata Bellary, Naveen B R Kumar et al 2011 Data management in clinical research : An overview Indian Journal of Pharmacol. 44(2) 168-172

[2]. Clinical Data Interchange standards consortium (CDISC) website-www.cdisc.org

[3]. Haux, R. Knanp P. Leiner F. on educating about medical data management –the other side of the electronic health record. Methods Inf med 2007, 46;74-9 (Pubmed)

[4]. http://www.acdm.org.uk assessed 18th July 2015

[5]. https://en.m.wikipedia.org/wiki/clinical-dat assessed 21st of July 2015

[6]. Kuchinke W, Ohman C, Yang Q, et al. Heterogeneity prevails :The state of clinical data management in Europe results of a survey of ECRW Centres Trials 2010 11.79

[7]. Lu Z., Su J Clinical data management current status, challenges and future directions from industry perspective open access J. Clin Trials 2010, 2:93-105
[8]. Margaret Rouse sep 2008 Whatls.com
[9]. Oxford Advance learners Dictionary assessed 20th July 2015
[10]. Ptizer quality management in clinical trials 2009, 038209
[11]. Raymond Omollo, Michael ochieny and Tansy Edwards 2014. Innovative approaches to clinical Data management in resource limited settings using open source technologies PLOS Neglected Tropical diseases 2014 sen. 8 (9). e3134
[12]. Reynolds –Haertle R A, McBride R Single vs double data entry in CAST control Clin. Trials 1992; 13:487-94 (PUBMED)