The Changing Role of Data Management in Clinical Trials

Fauziyya Shahul Hameed[1], Neeraja Noundla[2], Sameeksha Poldasari[3]
[1]*Corresponding Author, BDS, Student, Clinosol Research Pvt Ltd*
[2]*Co-Author, B. Pharm, Student, Clinosol Research Pvt Ltd*
[3]*Co-Author, M. Pharmacy, Student, Clinosol Research Pvt Ltd*

Abstract-Clinical Data Management (CDM) has played an important role in clinical trials over the years. CDM used to be primarily concerned with the gathering, cleaning, and archiving of clinical trial data. However, with technological advancements and a greater emphasis on data quality and integrity, CDM now plays a much more strategic role in clinical trials. CDM is now involved in clinical trial protocol design and implementation, data management plan development, electronic data capture (EDC) system selection and management, and data integration from multiple sources. By implementing thorough data validation processes and ensuring that trial data is accurate, complete, and timely, CDM also plays an important role in ensuring data quality and patient safety. Overall, the CDM role in clinical trials has grown to include a broader set of responsibilities, with a greater emphasis on data quality, data integrity, and patient safety.

INTRODUCTION

A Clinical trial is designed to solve a research topic by providing data to prove or disprove a hypothesis. The study's results are highly influenced by the quality of the data that was generated. Clinical data management is the key aspect of conducting clinical trials. In the course of their research, all researchers engage in CDM activities, either knowingly or unknowingly. While conducting our research, we engage in some of the CDM processes without specifying the actual phases. The reader is given an overview of CDM processes in clinical trials in this article. [1]

The process of collecting, cleaning, and managing subject data in compliance with regulatory guidelines is known as Clinical Data Management (CDM).[1] The main goal of CDM processes is to provide high-quality data by minimising errors and missing data while collecting as much data as possible for analysis. To achieve this goal, best plans are implemented to ensure that data is complete, reliable, and processed correctly. This has been easily achieved by the use of software applications that keep an audit trail and make it simple to identify and resolve inconsistencies. Advanced innovations have made it possible for CDM to manage sizable trials and guarantee data quality even in challenging trials.[2] High-quality data must be completely accurate and fit for statistical analysis. These should meet the protocol's parameters and requirements.

This means that if there is a deviation, such as failing to meet protocol specifications, we may consider removing the patient from the final database. The data should also meet the applicable regulatory data quality requirements.

TOOLS FOR CDM

There's colourful software available for managing data, and these are pertained to as clinical data operation systems (CDMS).

A CDMS has come essential in multicenter trials to handle the massive quantum of data. ORACLE CLINICAL, CLINTRIAL, MACRO, RAVE, and clinical Suite are exemplifications of constantly used CDM tools. These software tools are functionally analogous and neither system has a significant advantage over the other. Some transnational medicinal companies use custom-made CDMS tools to meet their functional conditions and procedures. The most popular open-source tools are Open Clinica, OpenDNS, Trial DB, and PhOSCo. These are as good as their marketable counterparts in terms of performance. These CDM tools aid in maintaining inspection trails which is vital in nonsupervisory submission studies. Multiple stoner IDs can be created with access restrictions to data entry, medical coding, database design, or quality check grounded on the places and liabilities.

This ensures that each stoner has access to only the functions assigned. However, of the stoner id that made the change and of the time at which the change was made for inspection purposes(inspection trail), If at all data changes are permitted to a stoner also the software will make a record of the change made. (2) Regulations, guidelines, and norms in CDM Since pharmaceutical companies calculate on electronically captured data to estimate a drug, CDM has some guidelines and norms that must be followed. These electronic records must cleave to 21 CFR Part 11 of the law of Federal Regulations (CFR). This requires the use of validated systems to insure data delicacy, trustability, and thickness, as well as the use of secure, computer-generated, time-stamped inspection trails to record the date and time of driver entries and conduct that produce, modify, or cancel electronic records. (3). To guarantee the delicacy, trustability, and confidentiality of data, guidelines, and controls should be in place. However, it must be entered and reused in systems that misbehave with 21 CFR Part 11, If data must be submitted to nonsupervisory authorities.

The Clinical Data Interchange norms Consortium(CDISC), non-profit multidisciplinary association, has developed norms to support the accession, exchange, submission, and archival of clinical exploration data and metadata. Metadata is information about the data that has been entered. This includes information about the person who made the clinical data entry or change, the date and time of entry/ change, and the details of the changes made. Two necessary norms are the Study Data Tabulation Model perpetration companion for mortal Clinical Trials (SDTMIG) and the Clinical Data Acquisition norms Adjustment (CDASH), both of which are free to download from the CDISC website. The SDTMIG standard (4) describes the details of the data model and standard languages and serves as a companion for the organisation. CDASH v1.1 defines the abecedarian norms for data collection in a clinical trial and elaborates on the introductory data. information needed from a clinical, nonsupervisory, and scientific viewpoint. (5)

THE CDM PROCESS

CDM is a process that is initiated with the deliverable in mind. An error-free, valid, and statistically sound database is what the CDM process is meant to deliver, just as a clinical

trial is made to provide an answer to the research question. To achieve this goal, the CDM process begins early, even before the study protocol is finalised.

REVIEW AND FINALISATION OF STUDY DOCUMENTS

The protocol is studied for clarity and consistency from the viewpoint of database design.

During this study, CDM personnel will identify the data items to be collected as well as the frequency with which they will be collected in relation to the visit schedule. The CDM team creates a Case Report Form (CRF) as the first step in creating protocol-specific activities for data generation. The data fields should be well-defined and consistent. The CRF should indicate the type of data to be entered. the units in which measurements must be taken should be mentioned alongside the data field. The CRF should be brief, self-explanatory, and easy to use (unless you are the one entering data into the CRF). Along with the CRF, study investigators should be given filling instructions (called CRF Completion Guidelines) to ensure error-free data collection. The variable is named according to the SDTMIG or the conventions followed internally when CRF annotation is performed. Annotations are coded terms that are used in CDM tools to identify variables in a study.

Data Management Plan (DMP) is created based on these. The DMP document is a road map for dealing with data under probable circumstances and describes the CDM activities that will be carried out during the trial. The database design, data entry, and data tracking guidelines, quality control measures, SAE reconciliation guidelines, discrepancy management, data transfer/extraction, and database locking guidelines are all described in the DMP. A Data Validation Plan (DVP), which includes all edit-checks to be carried out and the calculations for derived variables, is also prepared in addition to the DMP. The edit check programs in the DVP aid in data cleaning by identifying discrepancies.

List of CDM activities

- Database designing
- Data collection
- CRF tracking

- Data entry
- Data confirmation
- Discrepancy management
- Medical coding
- Database locking

Database Designing

Databases are clinical computer software created to make it easier for CDM to perform several studies. (6) In general, these instruments are simple to use and developed to misbehave with nonsupervisory conditions.

To enhance data security," system confirmation" is carried out, during which system specifications, (7) stoner needs, and nonsupervisory compliance are assessed prior to deployment. The database contains information about the study ideal, time intervals, visits, investigators, spots, and cases, and CRF layouts are developed for data entry.

Data collection

Data collection is done using the CRF that may live in the form of a paper or an electronic interpretation. The traditional system is to employ paper CRFs to collect the data responses, which are restated to the database by means of data entry done in-house. In the the-CRF-based CDM, the investigator or a nominee will be logging into the CDM system and entering the data directly at the point. In the the-CRF system, the chances of crimes are less, and the resolution of disagreement happens briskly. (8)

Since pharmaceutical companies try to reduce the time taken for medicine development processes by enhancing the speed of processes involved, numerous pharmaceutical companies are concluding fore-CRF options (also called remote data entry).

CRF Tracking

The entries made in the CRF will be covered by the Clinical Research Associate(CRA) for absoluteness and filled-up CRFs are recaptured and handed over to the CDM platoon. The CDM platoon will track the recaptured CRFs and maintain their record. CRFs are tracked for missing runners and unreadable data manually to assure that the data aren't lost. In case of missing or unreadable data, an explanation is obtained from the investigator and the issue is resolved. (9)

Data entry

Data entry This is applicable only in the case of paper CRF recovered from the spots. (8) The alternate pass entry (entry made by the alternate person) helps in verification and concession by relating the recap crimes and disagreement caused by unreadable data. Also, double data entry helps in getting a cleaner database compared to a single data entry. former studies have shown that double data entry ensures better thickness with paper CRF as denoted by a lower error rate. (9)

Data Confirmation

Data confirmation is the process of testing the validity of data in agreement with the protocol specifications. Edit check programs are written to identify the disagreement in the entered data, which are bedded in the database, to insure data validity. These programs are written according to the sense condition mentioned in the DVP. These edit check programs are originally tested with ersatz data containing disagreement.

The distinction may be due to inconsistent data, missing data, range checks, and diversions from the protocol. In-CRF- grounded- CRF studies, data confirmation will be run constantly to identify disagreement. (10)

These disagreements will be resolved by investigators after logging into the system. Ongoing quality control of data processing is accepted at regular intervals during the course of CDM. For illustration, if the additional criteria specify that the case's age should be between 18 and 65 times (both inclusive), an editing program will be written for two conditions viz. age 65. still, the condition becomes TRUE, a distinction will be generated, If for any case. These disagreements will be stressed in the system and Data explanation Forms (DCFs) can be generated. DCFs are documents containing queries pertaining to the disagreement linked.

Discrepancy management

Distinction operation It also pertains to query resolution. Distinction operation includes reviewing disagreements, probing, and resolving them or declaring them as irresolvable. distinction operation helps in drawing the data and gathering enough substantiation for the diversions. Nearly all CDMS (Clinical data operation systems) have a distinction database where all disagreements will be recorded and stored with an inspection trail. On the basis of the

types linked, disagreements are moreover flagged to the investigator or closed in-house by Self-Evident Corrections(SEC) without transferring DCF to the point. The most common SECs are egregious spelling crimes. For disagreements that bear interpretations from the investigator, DCFs will be transferred to the point. Investigators will either write the resolution or explain the circumstances that led to the distinction in data. When a resolution is handed in by the investigator, the same will be streamlined in the database. In the case of-CRFs, the judgments are handed online by the investigator.

The CDM platoon reviews all disagreements at regular intervals in order to make sure that they've been resolved. The resolved data disagreements are recorded as ‘ unrestricted ’. This means that those confirmation failures are no longer considered to be active, and unborn data confirmation attempts on the same data won't produce a distinction for the same data point. In some cases, the investigator won't be suitable to give a resolution for the distinction. Similar disagreements will be considered as ‘irresolvable ’ and they're streamlined in the distinction database. distinction operation is the most critical exertion in the CDM process. Being the vital exertion in drawing up the data, utmost attention must be observed while handling the disagreement.

Medical coding

Medical coding helps in relating and duly classifying the medical languages associated with the clinical trial. For the bracket of events, medical word books available online are used. Generally, the Medical Dictionary for Regulatory Conditioning(MedDRA) is used for the coding of adverse events and other ails.

The World Health Organization – Drug Dictionary Enhanced(WHO- DDE) is used for rendering the specifics. These workbooks contain separate groups of adverse events and medicines in proper classes.

Medical coding helps in classifying reported medical terms on the CRF to standard dictionary terms in order to achieve data thickness and avoid gratuitous duplication.

Database locking

After a proper quality check and assurance, the final data confirmation is run. However, the SAS datasets are finalised in discussion with the statistician, If there is no disagreement. All data operation conditioning should have been completed previous to database cinch. Once the blessing for locking is obtained from all stakeholders, the database is locked and clean data is uprooted for statistical analysis. Generally, no revision in the database is possible. But in case of a critical issue or for other important functional reasons, privileged druggies can modify the data indeed after the database is locked. This requires proper attestation and an inspection trail has to be maintained with sufficient defence for streamlining the locked database. (10) (12) (13) (14) (15)

ROLES AND RESPONSIBILITIES IN CDM

Different roles and responsibilities are attributed to the team members in a CDM team. The minimum qualification for a team member in CDM should be a graduate in life science and knowledge of computer applications. Ideally, medical coders should be medical graduates. However, in the industry, paramedical graduates are also recruited as medical coders. The list of roles given below can be considered as minimum requirements for a CDM team:

- Data Manager
- Database Programmer/Designer
- Medical Coder
- Clinical Data Coordinator
- Quality Control Associate
- Data Entry Associate

The data director prepares the DMP and approves the CDM procedures and all internal documents related to CDM exertion. The database programmer/ inventor performs the CRF reflection, creates the study database, and programs the edit checks for data evidence. The medical coder will do the picture for adverse events, medical- ails, and attendant medicine administered during the study. The quality control associate checks the delicacy of data entry and conducts. (10) The data entry labour force will be tracking the damage of CRF runners and performing the data entered into the database.

CONCLUSION

CDM has evolved in response to the ever-adding demand from pharmaceutical companies to gormandize- track the medicine development process

and from the non-supervisory authorities to put quality systems in place to insure the generation of high-quality data for accurate medicine evaluation.

To meet these prospects, there's a gradual shift from paper-grounded to electronic systems of data operation. Technological developments have appreciatively impacted the CDM process and systems, thereby leading to encouraging results on the speed and quality of data being generated. At the same time, CDM professionals should ensure the norms for perfecting data quality. (11) CDM, being a specialty in itself, should be estimated by means of the systems and processes being enforced and the norms being followed. The biggest challenge from the nonsupervisory perspective would be the standardisation of the data operation processes across organisations, and the development of regulations to define the procedures to be followed and the data norms.

From the assiduity perspective, the biggest chain would be the planning and perpetration of data operation systems in a changing functional terrain where technological development outdated the being structure. In malignancy of these, CDM is evolving to become a standard-grounded clinical exploration reality, by striking a balance between the prospects from and constraints in the being systems, driven by technological developments and business demands.

REFERENCES

[1] Gerritsen MG, Sartorius OE, vd Veen FM, Meester GT. Data management in multi-center clinical trials and the role of a nationwide computer network. A 5-year evaluation. *Proc Annu Symp Comput Appl Med Care.* 1993:659–62.

[2] Lu Z, Su J. Clinical data management: Current status, challenges, and future directions from industry perspectives. *Open Access J Clin Trials.* 2010;2:93–105.

[3] CFR - Code of Federal Regulations Title 21 [Internet] *Maryland: Food and Drug Administration.* [Updated 2010 Apr 4; Cited 2011 Mar 1].

[4] Study Data Tabulation Model [Internet] *Texas: Clinical Data Interchange Standards Consortium.* c2011. [Updated 2007 Jul; Cited 2011 Mar 1].

[5] CDASH [Internet] *Texas: Clinical Data Interchange Standards Consortium.* c2011. [Updated 2011 Jan; Cited 2011 Mar 1].

[6] 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.

[7] 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 centres. *Trials.* 2010;11:79. 8. Cummings J, Masten J. Customised dual data entry for computerised data analysis. *Qual Assur.* 1994;3:300–3.

[8] Reynolds-Haertle RA, McBride R. Single vs. double data entry in CAST. *Control Clin Trials.* 1992;13:487–94.

[9] Krishnankutty B, Bellary S, Kumar NB, Moodahadu LS. Data management in clinical research: An overview. Indian J Pharmacol. 2012 Mar;44(2):168-72. doi: 10.4103/0253-7613.93842. PMID: 22529469; PMCID: PMC3326906.

[10] Ottevanger PB, Therasse P, van de Velde C, Bernier J, van Krieken H, Grol R, et al. Quality assurance in clinical trials. *Crit Rev Oncol Hematol.* 2003;47:213–35.

[11] 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.

[12] Lu ZW. Technical challenges in designing post-marketing eCRFs to address clinical safety and pharmacovigilance needs. Contemp Clin Trials. 2010;31:108–18.

[13] Myles J. Experiences leveraging Electronic Health Records (EHR) in direct support of drug development. CHI's 2nd Annual Electronic Data in Clinical Trials: Collecting and leveraging data to optimise clinical trials. Philadelphia, PA, March 8–9, 2010.

[14] Hyde A. The changing face of electronic data capture: from remote data entry to direct data capture. Drug Inf J. 1998;32(4):1089–1092.

[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.

[16] Ottevanger PB, Therasse P, van de Velde C, Bernier J, van Krieken H, Grol R, et al. Quality assurance in clinical trials. Crit Rev Oncol Hematol. 2003;47:213–35.
[17] Cummings J, Masten J. Customised dual data entry for computerised data analysis. Qual Assur. 1994;3:300–3.
[18] Lu Z, Su J. Clinical data management: Current status, challenges, and future directions from industry perspectives. Open Access J Clin Trials. 2010;2:93–105.
[19] Lu ZW. Electronic data-capturing technology for clinical trials: Experience with a global postmarketing study. IEEE Eng Med Biol Mag.