Informatics in Medicine Unlocked 9 (2017) 219–229

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Artificial intelligence based clinical data management systems: A review

Gazali, Sandeep Kaur, Inderbir Singh *

Department of Pharmaceutics, Chitkara College of Pharmacy, Chitkara University, Rajpura, 140401, Patiala, Punjab, India

ARTICLE INFO

Keywords:
Clinical data management (CDM)
Case report forms (CRF)
Data management
Clinical data management softwares

ABSTRACT

The clinical data management system is widely used to manage the data that is collected during the clinical trials. This system offers various comfortable methods via which the data can be collected, managed and stored easily for further use. The points regarding the clinical data management that are covered in this review article are the basic introduction covering the general aspects of the clinical data management system and various clinical data management softwares plus the use and the importance of case report forms in CDM and the data management procedures that are being followed up for the proper management of the data so that it is easily accessible by the personnel. Also, the use of various softwares in the clinical data management process has been discussed depicting how the softwares perform various functions to keep the data in a managed, secured and an accessible form. As an endpoint, various future challenges and options are considered which give a detailed idea about the growth of clinical data management in the Pharmaceutical Industry.

1. Introduction

The analysis and authorization of new pharmaceuticals is based upon the trust that clinical trials will intend to find the answers to the investigation problems by providing clinical data which further proves or disapproves a particular hypothesis. The type and quality of the clinical data plays an imperative task in the conclusion of the study performed. Therefore, this clinical data so obtained is appropriately managed to obtain the accurate results of the clinical study thus, a system Clinical Data Management (CDM) is needed for the authentication of the study. Now, a question arises that what is CDM? Clinical Data Management forms a fundamental part in the clinical trial studies. CDM is implied in all the facets of operating computers, dispensation of the clinical data, managing the subject data and database systems to support the collection of the data. Clinical Data Management is precisely defined as the collection, integration and validation of the trial data [1]. When the clinical trials are performed, the prime duty of the investigators is to collect the data of the patient's wellbeing after a specific interval of time. Further, this data is given to the trial sponsor who quantifies and qualifies the given data by statistical means. When the approval of new drugs is to be made by the regulatory agencies it is reliant upon the clinical trial data presented. The trust on the clinical data is usually adhered to the quality practices and standards of the clinical trials performed [1,2]. Therefore, the organizations assure that the clinical trials performed and the data obtained are in the hands of well qualified and trained staff and a flow chart representing the staff that is involved in the clinical data management system is represented in Fig. 1. Thus, the key objective of CDM is to offer high quality data by observing the errors and missing data and keeping it as low as possible to congregate maximum data for analysis. Various practices have been developed to ensure that the data obtained is complete, processed correctly and reliable. This has been easily achieved by the use of applications of the software that presents effortless detection and motion of data discrepancies and is used to maintain audit trials. In clinical data management, softwares are generally required to address the electronic data capture, preparation of the electronic FDA submission, acceleration of the clinical trial management processes. A tabular form representing the use of software systems in various management stages of the clinical data has been depicted in Table 1 [2,3]. These sophisticated innovations have helped in maintaining data quality in complex trials and handle large trials. Another question that now arises is "what is a high quality data?" A high quality data is defined as the data which is suitable and accurate for statistical quantification. The data should satisfy the protocol specific parameters should be compliable with the protocol necessities. This tells us that when the data is not compliable with the protocol specifications we can exclude the tolerant from the ultimate database. It should be acknowledged that in some cases regulatory agencies can be engrossed in viewing such data. Probing further, usually misplaced data is also a matter of apprehension for the scientists, talking about high quality data there should be no errors and misses. Most prominently, high eminence data should accept only the arbitrary

* Corresponding author.
E-mail address: inderbir.singh@chitkara.edu.in (I. Singh).

https://doi.org/10.1016/j.imu.2017.09.003
Received 2 August 2017; Received in revised form 13 September 2017; Accepted 13 September 2017
Available online 15 September 2017

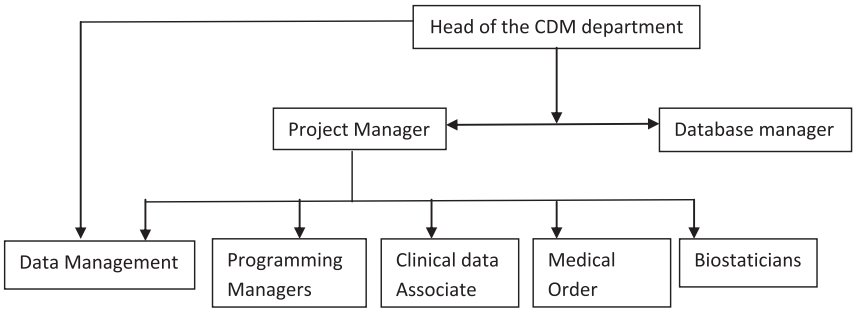

Fig. 1. Manpower involved in clinical data management.

variation in the data that would not influence the result of the study in the future. The data should convene with the regulatory necessities that are provided for high quality of the data [2,4].

The novel developments in the technology of computer hardware and software have contributed in making clinical trials effective, reliable and timely which act as the centerpiece for conducting a successful clinical trial. With the advancement in the computer world and the availability of design tools, software vendors, commercial databases and security applications the management of clinical trial data is widely attainable, less time consuming, easier, more scalable and secure than the past. Undoubtedly, these attributes give the assurance in the confidence of the results but novel challenges have evolved with the use of these attributes which also must be kept in mind. These usually include learning curve, cost, trading with unforeseeable events and changing errands [4]. There are good clinical data management practices that deal with data acquisition, privacy, electronic data capturing, Case Report Form (CRF) printing, preservation of CRF, data storage, validations and many more.

Standard operating procedures are the processes that are followed to accomplish data management activities and to support the responsibility of obeying the guidelines as per ICH GCP and 21 CFR part 11. Standard operating procedures or SOP's are usually applied in pharmaceutical processing and are also related for clinical studies. In clinical studies, the main focus is on recurring application of unaffected processes and their documentation therefore, it supports the isolation of origins, effects and causes. Further, applications made are with respect to the priority of patient treatments when the restricted sources get utilized according to estimation on urgency, staffing possibilities and ranking. The quality assurance team is responsible for monitoring that the study and test meet the SOP. SOPs also act as a reference to new employees by answering questions without interrupting supervisors [3,4].

Usually procedures are engaged with safe working and they are sometimes referred to as safe work method statements. They are preceded by methods of analyzing tasks in administrative center which includes a loom called job safety analysis in which hazards are identified and deterrent measures are described. The flowchart depicting the workflow and the travel of data in a Clinical Data Management process is given in Fig. 2.

Once the CRF and the protocol are approved, the process moves further towards the annotation of the CRF and the preparation of the documents also, approval of blank CRF can be done. Further, the database design is set up and the test data is entered into the database. Moving forward, the database is quality checked and quality control is done. Next step involves the entry of data which can be done via maintaining the log papers and then data is then sent for batch validation. After performing batch validation the discrepancies are managed and after having a proper quality check of data, the statistical report is generated [4].

2. Case report forms

A case report form can be electronic or a paper based system and is generally abbreviated as CRF. It is widely used tool by the sponsor for the collection of the data from the patients participating in a clinical trial. The data regarding the participation of a patient in a clinical trial is documented in the case report forms which also includes the adverse events. The CRF is usually developed by the sponsor to collect the data that would be required for testing the hypothesis in a clinical trial or it can also be used for answering a plethora of questions regarding the clinical testing [5]. The size of the CRF varies from a snapshot of handwritten pages regarding the physical condition of the patients or hundreds of pages of electronically captured data that has been collected as a result of the trial performed for over a period of weeks or months. Also, the CRF accurately depicts or provides detailed information about the protocol referred for the clinical trial. A sponsor also holds the duty of managing the production of the results gained by the trial and monitoring the collected data [5].

CRF records the data obtained during patient's partaking in the clinical trial. Usually this data is primarily de-identified that is, any information regarding patient's name, medical record number etc. is removed and an inimitable study number allotted to the patient. The supervision of this process is done by the Institutional Review Board. From a sponsor's point of view, the main aspect of a clinical trial is to obtain an accurate CRF but, due to errors caused by human or machine this is not absolutely achievable. To counteract these errors the sponsor usually

Table 1
Evolution of clinical trial management using various software systems.

S. No.	Data Management Stage	Software System	Applications
1.	Clinical Trial Development Planning and Trial Initiation	BRAAN, DataLabs, Fast Track Systems, IRBWISE, ProIRB, IRBNet	Protocol design and execution, Institutional Review Board, Investigator relationship management
2.	Clinical Trial Management System	Siebel Clinical, Oracle, ClinSource	Integrated CTMS Providers, International Management Package for Administration of Clinical Trials (IMPACT)
3.	Patient subject and study enrollment	Oracle, Phase Forward, DataTrack, Parexel, eResearchTechnology, DataLabs, Nextrials, ClinPhone, CRF, invivodata	Investigator relationship management, electronic data capture, case report forms, electronic patient diaries
4.	Study Monitoring and Reporting	Oracle Clinical, Phase Forward, NetRegulus, Aris Global	Clinical Trial Supply Management, cost tracking, document management, adverse event reporting
5.	Study Completion and Regulatory Filing	SyTech, Wimmer systems	Data analysis and reporting, regulatory submission assembly, communication and review, Custom solutions