KenKyu

Clinical Case Reports, Research & Trials 1:8-14 (2016)

Review Article

Role of Clinical data management system in multi-center clinical trial

Jindal A[1], Mishra R[1], Kumari R[2]*, Dubey GP[3]

[1]Junior Research fellow in Centre for Inter disciplinary Biomedical Research of Adesh University, Barnala Road, Bathina, India.
[2]Junior Research Fellow under DST project and Ph.D Scholar, Department of Kriya Sharir, Faculty of Ayurveda, Institute of Medical Science, Banaras Hindu University, Varanasi, Uttar Pradesh-221005, India.
[3]Distinguished Professor, Advanced Centre for Traditional and Genomic Medicine, Institute of Medical Science, Banaras Hindu University, Varanasi - 221 005, Uttar Pradesh, India.

Received November 02, 2016; **Accepted** November 16, 2016; **Published** November 18, 2016

Abstract

Multidisciplinary cooperation in multi trial centers needs an organized data management, configuration to ensure true progress monitoring and high quality research data. So, that Clinical Data Management (CDM) is an essential tool in the medical study, leads to produce high-quality, reliable, and statistically significant data from multiple clinical trials and diminish time phase of drug development to marketing. CDM's members are extremely active during all stages of clinical trial or drug development, from inception to completion and maintain the quality standards of CDM processes On the other hand, multiple procedures in CDM, including Case Report Form (CRF) designing, CRF annotation, database designing, data-entry, data validation, discrepancy management, medical coding, data extraction, and database locking are assessed for quality at regular intervals during a multi centric clinical trial. Presently, CDM is becoming compulsory for drug development companies to submit the data electronically. This review article spotlight on the processes involved and provides the reader an overview of the tools and protocols adopted as well as the roles and responsibilities in CDM.

Keywords: Clinical data management; data management; Multi centric trial; Good clinical data management practices.

Introduction

Multi-center clinical trial (MCCT) is a group of several discipliner, deliberated to find answers to the research question by means of generating data for proving or disproving a hypothesis and performed at many clinic center. Most large clinical trials, particularly Phase III trials, are conducted at different clinical research centers. MCCT include a larger number of participants, different geographic locations, the possibility of inclusion of a wider range of population groups, and the ability to compare results among centers, all of which increase the generalizability of the study. In various cases, efficacy will vary significantly between population groups with different genetic, environmental, and ethnic. It is very important to generate quality data to evaluate the benefit of therapies for the treatment of disease, therefore, MCCT needs a formalized data management structure to ensure true progress monitoring and high quality research data, consequently, MCCT require such a computerized tools i.e. Clinical Data Management (CDM) involve in organizing and managing of the different clinical data, generated from multidisciplinary collaboration. clinical data gathered at the investigator site in the case report form (CRF) are stored in the CDMS and to reduce the possibility of errors due to human entry, the systems employ various means to verify the data. CDM can be self-contained or part of the functionality of a CTMS (clinical trial management system). A CTMS with clinical data management functionality can help with the validation of clinical data as well as the help the site employs the data for other important activities (building patient registries and assist in patient recruitment efforts) [1-2].

CDM is a very applicable and essential part of an any type of clinical trial (fig-1). It is a multidisciplinary activity, including

***Corresponding Authors:** Rinki Kumari, Junior Research fellow in Centre for Inter disciplinary Biomedical Research of Adesh University, Barnala Road, Bathina, India E-mail: rinkiv3@gmail.com

Dubey GP, Distinguished Professor, Advanced Centre for Traditional and Genomic Medicine, Institute of Medical Science, Banaras Hindu University, Varanasi - 221 005, Uttar Pradesh, India. E-mail: gpdubey13@gmail.com

research nurses, clinical data managers, investigators, support personnel, biostatisticians and database programmers. Presently, CDM is highly accepted that the design, implementation, coordination and analysis of clinical trials require a multidisciplinary specialist approach [3].

CDM have several responsibilities from the handling of data to finish with data management or information. During course of work CDM have followed given protocol - to be collected or analyzed along with database design, design of electronic case report forms (CRFs), recruiting patients, collection of data, and statistical analysis [3].

There are many steps involved in the data managing such as-collection, cleaning, and organize or management of all subject data in compliance with regulatory standards. During the process its activity also in data abstraction/extraction; data processing or coding; data transmission; data storage; data privacy and data QA. Therefore, it is relevant to provide high quality of data by keeping the number of errors and missing data as low as possible and gather maximum data for analysis. CDM also provides data confidentiality, safety reporting, records and reporting and monitoring clinical research. A few customary applications rarely used for this purpose, such as Access®, Excel® or FileMaker®, rapidly reach their limits when confronted with non-trivial requirements [4].

In this system, some basic research team involved in the completion of the clinical research like - the principal investigator, sub-investigators, clinical monitors or clinical research associates ("CRAs"), the data management team and statisticians: the clinical research coordinator (CRC) [5-11]. CDM team members have some essential responsibilities which are well defined in the Spanish regulation in clinical trials, FDA regulations and international guidelines on Good Clinical Practice (GCP) [12].

GCP has some important utility like data handling and confidentiality, trial management, subject and safety reporting, quality control, records,reporting and monitoring. Therefore, CDM is valuable assets for a pharmaceutical company because it provide vital confirmation of a drug's efficacy and safety with the collection, integration and validation of clinical data. A large number of data are collected during the complete life cycle of medical study. So,serve as the basis for collection, storage, validation, analysis, submission, approval, labeling and marketing of a compound [13].

Several statistical studies have classified CDM into two class-1. paper-based (PB) and 2. electronic data capturing systems. In first class of CDM, case report forms are manually filled at site and mailed to the company for which trial is being performed. The data on forms is transferred to the CDMS tool through data entry. Paper-based (PB) CDM is most popular method being double data entry where two different data entry operators enter the data in the system independently and both the entries are compared by the system. In case the entry of a value conflict, system alerts and a verification can be done manually therefore, it is the double data Entry method whereas in electronic data capturing systems the investigators directly uploads the data on CDMS and the data can then be viewed by the data validation staff. Once the data are uploaded by site, data validation team can send the electronic alerts to sites if there are any problems. Such systems eliminate paper usage in clinical trial validation of data and single data entry method [14].

CDM ensures the data qualities are captured effectively and ethically by spot staff through paper case report form (CRF) or electronic case report form (eCRF). These data should be absolutely accurate and suitable for statistical analysis and assemble the protocol specified parameters and fulfill with the protocol requirements. This implies that in case of a deviation, not meeting the protocol specifications, consequently, excluding the patient from the final database. It should be borne in mind

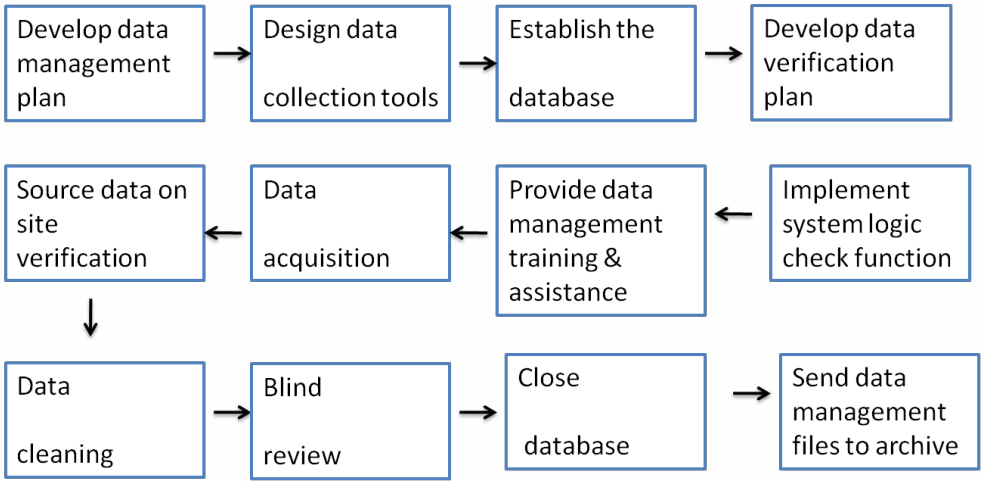

Figure 1: General flowchart for clinical research data management

that in some situations, regulatory authorities may be interested in looking at such data. In the same way, misplaced clinical trial data is also an issue of concern for researchers. These significance data should possess only an arbitrarily 'acceptable level of variation' that would not affect the conclusion of the study as well as also meet the applicable regulatory requirements specified for data quality [14]. During the phase, CDMS data are then transferred for the data validation and during validation the data clarification from sites are done through paper forms, which are printed with the problem description and sent to the investigator site and the site responds by answering on forms and mailing them back.

Electronic data capturing systems

Data collection has been done through different phases that start from the phase zero clinical drug trials. These collected data are designed and made accessible for early review. It will be utilized to characterize the patient population and also to evaluate the safety and efficacy of the trial drug. Structured databases are used to collect, analyze and organize clinical data. Sponsor is responsible for the reliability and quality of data and this could be ensured by collecting and transferring data from study subjects to a clinical data management system. Therefore, data can come from a variety of sources such as Investigator sites, Laboratories and Directly from patients. So that, during multi centric clinical trial, the investigators collect data on the patients' health for a defined time period. This data is sent to the trial sponsor, who then analyzes the pooled data using statistical analysis.

Some other research has cleared that CDM will make accessibility of an overview of clinical data management and introduce the CCR's clinical research database. At the last of this phase, the contributor will be able to define and describe some following important point-

- What constitutes data management activities in clinical research?
- Regulations and guidelines related to data management practices.
- What a case report form is and how it is developed?
- The traditional data capture process.
- How protocols are developed in Cancer Central Clinical Database (C3D)?

Tools for CDM

CDM is a computerized based tool, so that many software tools are available for data management in this system (CDMS). In MCCT, a CDMS has become indispensable to handle the massive amount of data. A large number of CDMS software tools are used in pharmaceutical companies which are expensive and need complicated Information Technology infrastructure to function, however a few open source tools are available as well. Commonly, RACLE CLINICAL, CLINTRIAL, MACRO, RAVE, and eClinical Suite are used as CDM tools while various multi-national and international pharmaceutical companies huge use custom made CDMS software tools according to their process needs. A few open source of CDMS software tools are freely available such as OpenClinica, openCDMS, TrialDB, and PhOSCo and as good as their commercial counterparts in terms of functionality and can be downloaded from their relevant websites.

The CDM Process

The CDM process (starts before the finalization of the study protocol) is designed to deliver an accurate, high quality, valid and statistically significant data and to answer the research question. Some are important steps in the process of CDM (Fig -2) whereas that steps are not exactingly ordered. For example, it is common in longer studies to generate intermediate discrepancies and listings periodically to identify problems that need correction before study completion.

Clinical data management have done various task like –Clinical development plan (CDMP)- should be developed early during the setup of the study for each and every clinical trail; Including describe all the components of the CDMP; Each component in the CDMP should specify; Responsible staff for the work ; Followed Guidelines ,SOPs and high quality output will be produced; The entire team members should review and agree with the CDMP to make sure a consistent approach to the process and guidelines a living document throughout the life cycle of a study; to address any updates or changes made during conduct of the study along with study setup in table -1.

• Case report form (CRF) design
• CRF completion guidelines
• Trial database (DB) setup
• Validation checks

Table 1: showing the study setup steps.

Case report form (CRF) design, is designed to collect the entire data from the patient in MCCT and a paper or electronic questionnaire relevant to clinical trial research (fig-3). Its development represents a significant part of the clinical drug trial and can affect study success [15].

It acts as tool, used by the team members of the clinical trial to collect data from multicentre site and all data on each patient participating in a clinical trial are held and/or documented in the CRF, including adverse events [16].

CRF designing requires giant planning and attention to minute detail and it will aid in assessing the safety and efficacy of the medicinal product accurately; should be designed for optimal collection of data according to protocol compliance and regulatory. Good CRF aid to minimize errors; reduce processing time; minimize free text, used code format.

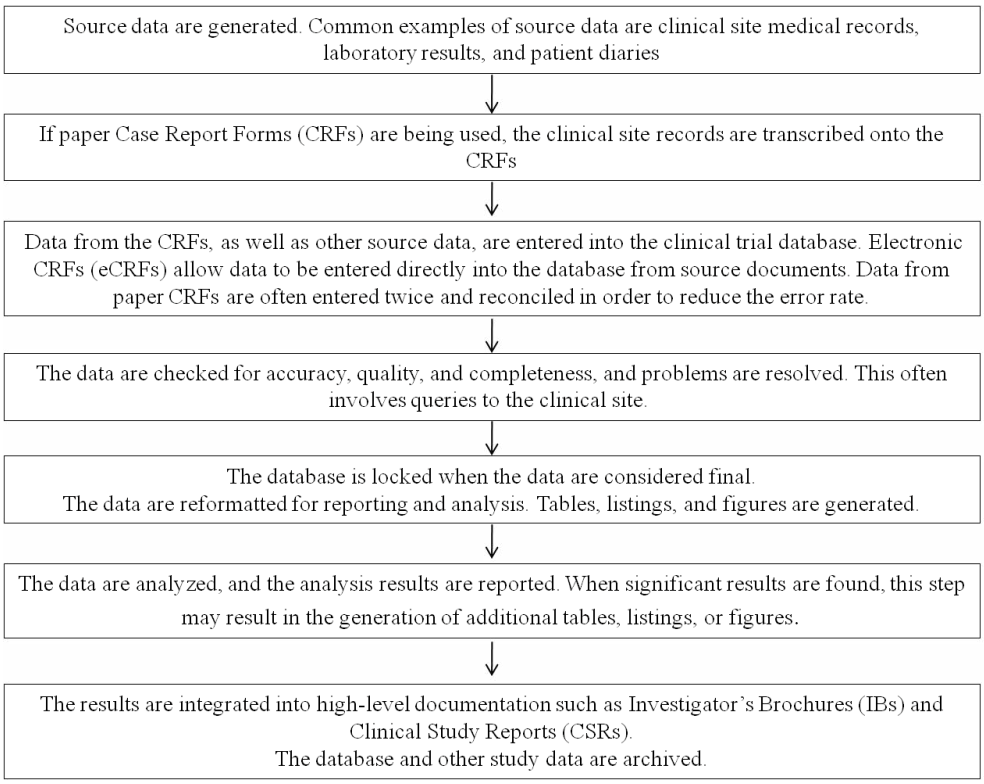

Figure 2: flow chart showing the process steps in the data management.

ADVERSE EVENT		
Did any adverse event occur after the last visit?—*Indicator Question*	Yes ☐ 1 No ☐ 2	
If " Yes", please provide the information in the adverse events (AE) page (page no. XX) and give details below:–Skip		
AE page number	AE serial number	Did any unscheduled visit happen after the last visit? *(Please provide details on page no. YY)–Skip*
☐☐	☐☐	Yes ☐ 1 No ☐ 2
☐☐	☐☐	Yes ☐ 1 No ☐ 2
☐☐	☐☐	Yes ☐ 1 No ☐ 2

Figure 3: CRF format and an adverse event page of CRF is depicted showing codes, and skips questions

CRF completion guidelines- Full, accurate completion of CRFs; Quality of data captured; Fewer queries; Quicker validation of data; Complete, concise and logical guidelines for CRF completion ensure; All required fields are completed; Data recorded in the CRFs are logical; Free text entries are spelled correctly and clinically appropriate; Definitions for items that are

not directly measurable ; procedures for making corrections to data; Handling completed CRFs; Shipping the CRFs from sites to the DM center;

Update CRF completion guidelines.

Trial database setup- The entire clinical data must be entered and stored in a computer system (in Excel spreadsheet, a Microsoft Access application, SAS tables, or a set of tables built in one of the applications such as Oracle, Open Clinica); Success of a Clinical trial depends on the quality and integrity of its Data. DATA BASE structure should be eased and speed of Data entry prevention of errors in data creation and modification; efficient creation of data sets for analysis formats of data file requirements. Output from DB design is a specification of the DB (information for each variable)- Name and label ; Type (e.g., numeric, character, integer, date…) • Length (number of characters) ;Definitions for all coded values are so important.

Validation investigation - All computer systems are involved in the management and processing of collected data must undergo a validation investigation to ensure that they perform as intended and that results are reproducible. It is updated to CRF revolutions or errors need correcting So, this is crucial tool for each study DB and created for all study endpoints and safety data ; Increases data quality ; Greater efficiency for data cleaning ; Identify data inconsistencies and potential errors ; Involved in missing values checks: apply to critical variables similar to center number, subject numbers, primary safety and efficacy variables; Range checks – common checks, identify errors outside of the expected range example- expected weight for the study subject was between 40 – 80 kg).

Data Entry- Data entry takes place according to the guidelines prepared along with the DMP and is applicable only in the case of paper CRF from the sites. As a law, double data entry is performed, for a first pass data entry to be completed, followed by a second

Verification step (entry made by the second person) (15) dual data entry for computerized data analysis. If any discrepancies between the first and second verification may be resolved such that the data entered is a true reflection of that recorded on the paper CRF. Where the operator is unable to read the data entry so, the clinical data manager should be reported and then entry may be clarified with the person who completed the CRF. However, double data entry ensures better reliability with paper CRF as denoted by a lesser error rate (15-16.) Elements of data quality- The fundamental elements of data quality for both paper and electronic records by "ALCOA is the FDA Guidance for Industry and Computerized Systems, used in the clinical data management and investigation.

- Attributable: The source of the data is known.
- Legible: The data are readable and comprehensible to humans.
- Contemporaneous: The data are recorded when they are generated.
- Original: The data are the first recording from the primary source.
- Accurate: The data are correct.

Basically, specific quality data are also required three important other aspects:

- Data is readily available, transmissible, and storable.
- Data is complete and unbiased.
- Data is in a format that is internally consistent and compliant with or readily transformable to an accepted standard.

Data validation

Data validation is the very important application of validation investigation tools to the high quality data. Data validation may be applied in real time at the point of entry in case of electronic CRFs and offline validation may still be required e.g. for cross checks between data types. So, this is the process of testing the validity of data in accordance with the protocol specifications.

Data queries

Data query (DQ) may be issued to the investigative site when entered data does not pass validation rules so that the clinical trials is conducted to request clarification of the data entry whereas it must not suggest the correction that should be made. In case of electronic CRFs, the site staff with appropriate entrée may modify data entries whereas in case of paper CRFs, the clinical data manager applies the data query response to the DB. A copy of the data query is maintained at the investigative site. When a variable has again a data error raised, is known as a "discrepancy" or "query". However, DQ is an error generated (by typographical, copying -coding) when a validation check detects a complexity with the data. Automatically, validation checks a saved "submitted" page and can identify problems with a single variable, between two or more variables on the same eCRF page.

Errors can be resolved in several ways: by correcting the error – entering a new value for example, or when the data point is updated and by marking the variable as correct – some EDC systems required additional response or you can raise a further query if you are not satisfied with the response.

Database finalization and extraction

All predictable data is accounted for, all data queries closed, all external data received and reconciled and all other data management activities complete the database may be finalized.

Metrics and tracking

By using of computerize tool archetypal reports generated and

used by the clinical data manager includes:

• Status of page completion / missing pages

• Status of data queries

• Data queries not resolved within specified time limit

• Commonly raised data queries

Quality control- it is applied at various stages in the CDM process and is normally

Role and responsibilities of the clinical data manager with other team member in CDM, all team member have different roles and responsibilities according to their qualification. For this the minimum educational requirement is graduation in life science and familiarity of computer applications. Medical coders should be medical graduates. Though, the list of roles given below can be considered as minimum

Requirements for a CDM team members-

• Clinical Data Manager (CDM)

• Database Programmer/Designer

• Medical Coder

• Clinical Data Coordinator

• Quality Control Associate

• Data Entry Associate

The clinical data manager plays a key role in the setup and conduction of a MCCT. During a clinical trial the clinical data collected from the multidisciplinary collaboration site for safety and efficacy analysis, which in turn drive decision making on product development

In the pharmaceutical industry. According to the clinical trial protocol, CDM is involved in discussions about data collection options, and then oversees development of data collection tools. During the phase of clinical trial, when a subject enrollment at the initial level, the CDM ensures the data is collected, validated, complete, and consistent. Then, CDM communicates with other data providers like- a central laboratory processing blood samples collected. CDM has also made certain the data is transmitted securely and is consistent with other data collected in the clinical trial. At the end point of the clinical research, the CDM ensures the entire data has been reported and declared final for terminology varies, whereas ordinary descriptions are "Database Lock" and "Database Freeze", and the CDM transfers data for statistical analysis. However, clinical data management team produces results which are source-verified, reproducible and cost-efficient.

Importance of Clinical Data Management System (CDMS)- CDM system plays a essential role in MCCT is to ensure high-quality data are captured effectively and ethically by sites organization through both paper case report form (CRF) and electronic case report form (eCRF). When once data have been monitored for typographical errors, the data can be validated to make sure to logical errors. Example- Verify of the subject's date of birth to ensure that they are within the inclusion criteria for the study.

CDM can make coding of data and the coding is generally centered around two areas 1.adverse event terms and 2. Medication names. The database containing the adverse event terms or medication names can be connected to one of the dictionaries. Example- As an example, ASA (acetylsalicylic acid) could be mapped to aspirin, a common notation. Some important adverse event dictionaries are MedDRA and WHOART and popular Medication dictionaries are COSTART and WHO [17].

Drug Dictionary

At the end of the MCCT the database in the CDMS is extracted and provided to statisticians for further analysis and then, analyzed data are compiled into a clinical study report and sent to the regulatory authorities for approval. Importance of CDM in the clinical field -provides data and database in an exploitable format in a timely manner and ensures clean data and a 'ready to lock' database [18].

Conclusion

MCCT play a key role in the drug development and discovery in clinical areas and delivered a huge quantity of data. The data should be managed so, that CDM has involved in response to the collect, organize and managed each & every data and reproduce error free high quality data. Consequently, CDM is a vital vehicle in MCCT to ensure- The reliability & quality of data being transferred from trial subjects to a database system and then, the collected data is complete and accurate so that results are correct to support statistical analysis, and its subsequent presentation and interpretation, So that, CDM is evolving to become a standard based clinical research unit, by conspicuous a stability between the expectations from and constraints in the accessible systems, driven by technological developments and business demands.

References

1. Lu Zheng, Marvin Zelen (2008) Multi-Center Clinical Trials: Randomization And Ancillary Statistics. The Annals of Applied Statistics.2: 582–600.

2. Raptis, DA, Mettler T, Fischer MA, Patak M, Lesurtel M, et al. (2014) "Managing multicentre clinical trials with open source†". Informatics for Health and Social Care. 2: 39-67.

3. Riley D, Ward L, Young T (1994) For the British Oncology Data Managers Association; (BODMA): Oncology data management in the UK – BODMA's view. Br J Cancer. 70: 391-394.

4. Gerritsen MG, Sartorius OE, vd Veen FM, Meester GT (1993) 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. 659-62.

5. Warlow C (1990) Organise a multicentre trial. BMJ. 300:180-183.

6. Cassidy J, Macfarlane DK (1991) the role of the nurse in clinical cancer research. Cancer Nurs. 14: 124-131.

7. The DCCT Research Group (1993) The impact of the trial coordinator in the Diabetes Control and Complications Trial (DCCT). Diabetes Educ. 19: 509-512.

8. Arrigo C, Gall H, Delogne A, Molin C (1994) The involvement of nurses in clinical trials. Results of the EORTC Oncology Nurses Study Group survey. Cancer Nurs. 17: 429-433.

9. Brown JM, Haining SA, Hale JM (1997) Views on local data management in cancer clinical trials. Clin Oncol (R Coll Radlol). 9: 403-406.

10. REAL DECRETO 561/1993 (1993) Establishing requirements concerning clinical trials on drugs. Spain.

11. Department of Health and Human Services. (2003) Investigational New Drug Application. Responsibilities of Sponsors and Investigators, 21 Code of Federal Regulations 312.50–312.70.

12. The European Agency for the Evaluation of Medicinal Products, ICH Topic E 6. Guideline for Good Clinical Practice. Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95)

13. Krishnankutty B, Bellary S, Kumar NB, Moodahadu LS (2012) Data management in clinical research: An overview". Indian J Pharmacol. 44: 168–172.

14. Raptis, DA, Mettler T, Fischer MA, Patak M, Lesurtel M, et al. (2014) "Managing multicentre clinical trials with open source†". Informatics for Health and Social Care. 39: 67.

15. Cummings J, Masten J (1994) Customized dual data entry for computerized data analysis. Qual Assur. 3: 300-303.

16. Reynolds-Haertle RA, McBride R (1992) Single vs. double data entry in CAST. Control Clin Trials. 13: 487-494.

17. Nahm M, Shepherd J, Buzenberg A, Rostami R, Corcoran A, et al. (2011) Design and implementation of an institutional case report form library. Clin Trials. 8: 94–102

18. ICH Guidance E6: Good Clinical Practice: Consolidated guideline. US HHS, US FDA, CDER, CBER. 1996. [Last accessed on 2013 Jun 11].