CHAPTER 2

REVIEW OF LITERATURE
Clinical Data Management (CDM) has existed since the inception of science of healing, though devoid of documented evidence with respect to its individual steps in the early phase of its evolution. It is an inevitable fact that all researchers perform part of the CDM task consciously or unconsciously without paying much attention to the technical aspects involved [17]. Clinical data collection processes, what data to be collected and how to collect it and its management traverse the evolution of clinical research studies/trials crossing many hurdles including scientific, ethical, and regulatory apart from religious, political, and legal grounds. Data and its management is impacted by the characteristics of the studies/trials (design of the study: non-interventional studies, clinical trials, statistical processes, randomized, multicenter, geographic level: international/national/regional, a clinical trial phase, number of patients, purpose of the study: diagnostic, observational, therapeutic) [18], [19], [20], [21], [22], and continuously evolving ethical [18], [23], [24] and stricter regulations progressively demanding greater accountability. More and more various types of data are to be acquired to address regulatory compliance, it is important to ensure all relevant data is collected and reported in the manner prescribed by regulatory agencies. Thus data is to be managed and maintained effectively to keep up with dynamic regulatory environment. Clinical trials/drugs regulatory market approval is becoming one of the most regulated sectors worldwide. The pharmaceutical industry has started taking regulatory compliance seriously. Today CDM has evolved as a separate science and has key role to play in the overall process of drug development but limited reference material is available in the literature. CDM involves a spectrum of activities/responsibilities that encompass the range of operational, technical and regulatory affairs of drug development process, and these are affected by CDM practices. Regulatory compliance can be ensured by effective data management. The global face of drug development and consolidation of the industry, coupled with stricter regulatory environment has necessitated that the practice of CDM reflects the framework/harmonization of global practices at all stages of drug development. Regulations governing clinical trials in India have
come a long way in the last decade. Indian GCP guidelines and the regulatory framework of clinical trials regulations are continuously advancing to address inadequacies and along lines of WHO, ICH, US-FDA and European GCP guidelines as well as the Ethical Guidelines for Biomedical research on Human Subjects issued by ICMR [25]. The global initiatives, in particular- CDISC (Clinical Data Interchange Standards Consortium), SCDM (Society for Clinical Data Management), ICH-GCP are having significant impact on how clinical data collected and managed to mitigate prevalent heterogeneity in CDM procedures attributed to different data sources and CDM procedures used by different organization, and to recognize the need of common standards for uniform practice. The concerted efforts stamp from a common denominator to reduce development cost and regulatory approval time, thereby time to market.

2.1 Pre-1940 Brief History of CDM

Ayurveda, one of the world’s oldest healing systems for the purpose of conquering various ailments and improving human health, was prevalent in India around 3,000 BCE [26]. Though detailed guidelines for the management of diseases mentioned in Ayurveda texts are not supported by recorded documentation of data collection procedures [27], but it is reasonable to presume that some data observations must have been made to arrive at the various treatment options. In 1593, Sir Richard Hawkins cured scurvy with oranges and lemons. In 1747, James Lind determined that scurvy can be prevented by use of citrus, an unimplemented scientific breakthrough. Nearly 50 years later, British Navy inducted citrus fruits as compulsory part of diet to prevent scurvy [27]. This is not the only example of a lost translated opportunity. Antisepsis and antibiotic treatment apart from scurvy were other opportunities lost because gained knowledge was forgotten, disbelieved, neglected or unexploited. One of the most likely possibilities for such lapses between the observations and translational implementation may be lack of effective data collection, merger and analysis. This can be attributed to the absence of standard methods of organizing and representing data from the diverse data sources or observations. To be useful, data need to be organized, processed and interpreted in a standard manner that permits seamless data integration, translation of the data into information, and thereby the flow of knowledge.
The Figure 2.1 sourced and reproduced from “A treatise of scurvy by Lind, 1753” available at http://www.jameslindlibrary.org/lind-j-1753/ [28] depicts the face of data management limited to the textual documentations.

Figure 2.1: A Treatise of Scurvy by Lind, 1746 [28]
By the end of 18th century, Daniel Sutton published Suttonian system of inoculation with reference to smallpox vaccine [29]. This published literature of human experiment has a small note which cited about data processing by checking and rechecking the same manually (Figure 2.2).

Figure 2.2: The Inoculator or Suttonian System of Inoculation (1796), reproduced from http://www.jameslindlibrary.org/sutton-d-1796/ [29]
Frederick Akbar Mahomed (d. 1884) of London is said to have laid the foundation of collaborative clinical trials. This is because he emphasized on data as he established Collective Investigation Record by channelizing data collection from physicians practicing outside the hospital setting [30]. With the advent of newly enacted U.S. Food, Drug, and Cosmetic Act in 1938, all new drugs were expected to have pre-market safety evaluation data for the first time [31]. With this the glimpses of prototype existence of modern data management came into picture.

The first randomized control trial of streptomycin in pulmonary tuberculosis conducted in 1946 was a model of systematic enrolment criteria and data collection compared with a contemporary approach for clinical observations [27 and references there in]. It emphasized that tasks of a clinician warrant for strict adherence to the admission criteria, correct random allocation, structured organization of the standard investigations and diligent completion of the case records [32]. Most of the steps were carried out with an objective to have clean data from the trial and were handled by a physician. Needless to say that these tasks now falls under the preview of CDM [33].

Figure 2.3 sourced from Journal of the National Cancer Institute, 91(1), 1999 [33], depicts the publication details of the report.

Figure 2.3: BMJ, Oct. 30, 1948 “Streptomycin Treatment of Pulmonary Tuberculosis.” [33]
Studies were going on in the United States on streptomycin treatment against tuberculosis at the same time when it was being investigated in the United Kingdom (U.K.). Although Americans had adequate supplies of the product which led to more effective cure for subjects but the trial produced less conclusive clinical trial data [31], [34]. It may be indicative of flaws in the data collection methodology and its effective management.

2.2 Brief History of CDM:1950’s - 1980’s

Clinical trials as they concern health care investigations, apart from adhering to scientifically vigorous standards (professional ethics), structuring research and data management within the framework of health care ethics and regulatory compliant is an important and complex task, with a multitude of subtleties related to the safety and the protection of human subjects.

Violations of ethical principles relevant to clinical research have been documented [35]. The ethical advances in the protection of human subjects are a result of several revelations of abuse of human rights, public disasters and serious scientific misconducts, in particular, Trials of war criminals- Nuremberg Code 1948, Thalidomide-Declaration of Helsinki in 1964, and Syphilis Study at Tuskegee - Belmont Report 1978. These landmark developments have evolved into the various professional conduct standards and principles that guide the practice of medicine, and reinforced the need for best practices to collect trial data. For example, in 1932 male prisoners with syphilis were misinformed as to their treatment and recruited without consent. These men were not informed when penicillin became available for the treatment of syphilis. In another study (the Jewish Chronic Disease Hospital Study), live cancer cells were injected in patients with various chronic debilitating diseases. Consent was claimed, but was never documented due to the investigator’s contention that informing the patients of the procedure would frighten them unnecessarily. These and other abuses served to tighten both legal regulations and ethical guidelines for clinical research [35], [36]. In parallel to health care ethical codes, clinical studies began to become embodied in regulation as government authorities started recognizing a need for controlling medical therapies in the early twentieth century. [27]

The issues of health care ethics and regulatory challenges are becoming more numerous and complex with the scientific advances requiring dynamic updates in ethical and legal framework of clinical trials.
1950’s and 1960’s were era of increasingly advanced trial methodologies for drugs of different therapeutic segments [31]. These studies involved larger sample size demanding considerable scalability in data management process.

In early 1950’s AMA (American Medical Association) initiated a register for reporting of adverse drug reactions, though it had no mechanism to enforce data collection [31], [34].

In 1962 U.S. FDA did major amendments to its previous act. This modification demanded demonstrations of evaluation of efficacy of the investigational product. The additional data necessitated the conduct of controlled trials to support efficacy claims [31]. The whole change was observed as the dynamic crusade towards establishment of separate CDM units within pharmaceutical organizations.

Robert Temple, Deputy Center Director for Clinical Science, Center for Drug Evaluation and Research (CDER), at the Food and Drug Administration (FDA), has commented that as late as the 1960’s and early 1970’s, most of the studies were “inadequate beyond belief” and "You would be horrified [at the clinical trial data] submitted to the agency” [31],[37]. To address these inadequacies in the late 1970’s, for academic institutions, the use of mainframe computers at the clinical trial sites became prevalent. They permitted direct data entry into the system. Ironically, the restrictive methodology as well as high operating cost associated for data capture made them impractical for use [38]. But this was a prominent step towards structured methodology for CDM practice. In the 1980’s and 1990’s, remote data entry (RDE), use of personal computers, was introduced into the pharmaceutical industry. Technology became more connected; a signal sent by telephone line made the transfer of data easier. All together this made easier to scale and distribute, thereby resulting into overall upsurge in CDM methodology. Floppy disks and a smaller hard diskette helped to collect more amounts of data, supporting its easy portability [38], but had data integrity issues.

2.3 Brief History of CDM: late 1990’s

ICH-GCP guidelines 1996 have become the universal standard for ethical conduct of clinical trials (ICH reference). Sections E2A, E2B(R2), E2D E3, E6, E8 and E9 include specific references to CDM highlighting system validation, harmonization of adverse experience terminology and the reformatting of key tables and listings for reporting purposes as key procedure areas [39]. Data management staff role and contribution throughout the drug development process including design, documenting, data listing activities and study
reporting is highlighted [40]. There is reference to EU GCP Directive and a number of FDA guidance documents focusing on the submission of electronic CRFs and data listings [40], [41]. It is highlighted that the involvement of regulatory agency staff in the development of ICH and other guidelines, particularly with regard to electronic submissions should ultimately facilitate the review process, reduce the review period and further encourage consistency across regulatory agencies. ICH has started shaping how data management departments are structured, execution of data management tasks and data exchange modalities between sponsors and regulatory authorities [39], [41], [42], [43], [44]. There is widespread adoption of ICH helping develop common frameworks against which various activities involved in drug development can be evaluated, new Clinical Data Management System (CDMS) implemented and training and education imparted [45]. ICH-GCP guidelines are enforced by regulatory laws for trial conduct and provide assurance that the data and reported results are not only credible and accurate but also the integrity and confidentiality of trial subjects are respected and protected [46].

2.3.1 CDISC (Clinical Data Interchange Standards Consortium)

Around the time ICH GCP 1996 came into existence, a consortium, CDISC [47], by various companies from the pharmaceutical industry was formed with an aim to streamline clinical research. This was the turning point for ever evolving science of CDM [38]. Industry has acknowledged the importance of global, platform (CDMS)-independent data standards enabling information system interoperability and serving as a template for format of standard clinical trial tabulation datasets for submission to a regulatory authority such as the US-FDA [48], [49]. Subsequent to the standard for submission data, time taken for NDA (new drug application) review has reduced by approximately 45% (comments by J. Woodcock CDISC Conference November 2012) [38]. The CDISC standards provide the method for the data and metadata to be standardized, made more accessible [50] but lacks the steps how to manage clinical trial data.

With the advent of CDASH (Clinical Data Acquisition Standards Harmonization) in 2011, data managers have information about what has to be captured for a particular dataset in the CRF i.e. it provides basic recommendations about the data collection fields of approximately 18 domains [51].
2.3.2 SCDM (Society for Clinical Data Management)

SCDM, a non-profit, international organization, was founded with an aim to advance the field of CDM [52]. GCDMP (Good Clinical Data Management Practices) are guidelines by SCDM (last version November, 2013). The document gives the information about the CDM process by describing the minimum standards and best operational practices [17]. Though, the document has CDM steps mentioned but they are not specific to a particular therapeutic area. Also, the document is not in the public domain; accessible only to SCDM members and available on substantial payment [53]. There is urgent need for such guidelines in India within the ambit of Indian regulations and subsequent process standardization.

2.4 CDM methodology practice: prior to early 1990’s

In the past, both in India and internationally, similar face of CDM in practice was prevalent across most of the organizations [54]. In general, the CDM units in pharmaceutical companies were comprised of small groups of people carrying out a spectrum of activities associated to CDM tasks. There were no separate roles as project managers, monitors, Clinical Research Associate (CRA), medical monitors, medical writers, and pharmacovigilance or data management manager. The professional qualification(s) of data mangers/entry operators were not very desirous except data entry skills; data management task was limited to keying-in of data into the database. Ethical oversight was not practiced to applicable standards, trials approvals, informed consent processing, and audits lacked transparency. There was near complete lack of standardization, and quality issues such as GCP, data documentation, data quality, data protection, and data integrity were not adequately addressed. Training of personal in GCP, regulatory issues, clinical trial design, implementation, operational monitoring, clinical data management, and quality assurance were not primary concerns. Different quality standards were prevalent for local/ global trials studies. ‘No-Carbon-Required’ (NCR) paper was not used for paper CRFs. CRFs were generated by taking a printout followed by creating sufficient number of photocopies of the same and used as data collection instrument at the site. Database usually consisted of invalidated excel spreadsheets. Once data entry was complete, the raw data was subjected to manual review and consistency
check, performed primarily visually through comparison with the hard copy of the CRF pages. Data entry errors or outliers or missing values, which needed corrections, were updated by re-keying in the values. With progress of time, the use of NCR paper became prevalent and positively impacted data processing.

2.5 Current practice of CDM: International Trials

2.5.1 Overview

International CDM guidelines, best practices and standards that must be followed are as listed below [17] but are not mandated under Indian guidelines and regulations:

- Electronic records have to comply with a CFR (Code of Federal Regulations), 21 CFR Part 11.
- GCDMP document by SCDM.
- Clinical Data Interchange Standards Consortium

CDM is a group task and must have the roles listed below, which may be considered as minimum requirements for a CDM team [17]:

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

Team members of CDM actively contribute in all stages of clinical trial right from inception to completion therefore should have satisfactory procedural knowledge to achieve needed quality by applying correct CDM processes resulting into extreme reduction in overall time needed for drug development to its marketing. Various processes in CDM include [17]:

- CRF designing,
- CRF annotation, database designing,
data-entry, data validation,
discrepancy management,
medical coding,
data extraction, and
database locking

All above tasks should be assessed for quality at specified intervals during the clinical study. In the current international scenario, there is an augmented mandate to develop the CDM standards to meet the regulatory requirements and stay ahead of the competition. CDM team can meet these demands with the implementation of regulatory compliant data management tools. Additionally, it is becoming compulsory for global pharmaceutical industry to submit the data electronically. CDM specialists have to fulfil all expectations and set quality standards for trial data and also a drive to adapt to the rapidly changing technology. Technological developments have positively affected the CDM schemes and procedures facilitating generation of high quality data within stipulated time. CDM should be assessed by the systems and processes adopted and the standards implemented for its execution. The major challenge from the regulatory standpoint would be the standardization of CDM processes across various organizations, development of regulations and data standards. In spite of these hurdles, CDM is evolving into a standard based branch of clinical research. This could be successfully achieved by striking a balance between the expectations, constraints, existing systems, technological developments and business demands [17], [55]. Commonly used Clinical Data Management Systems (CDMS), software tools for data management are: CLINTRIAL, ORACLE CLINICAL, MACRO, eClinical Suite and RAVE [17]. Considering functionality these software are almost similar and with no significant advantage of one system over the other. These software tools need sophisticated Information Technology platform apart from being cost prohibitive. Moreover, some multinational pharmaceutical giants use custom-made CDMS tools to suit their operational needs and procedures. Some of the open source CDMS are, TrialDB OpenClinica, PhOSCo, and openCDMS. These CDM tools have the capability of capturing audit trail for data management activities required in studies for regulatory submissions. According to the roles, functions and responsibilities, multiple user IDs can be created with limited access to the respective functionalities.
2.5.2 Open-Source Clinical Trial Data Management: CDM Software

Legal requirement in Europe demands the conduct of clinical trials in agreement with ICH GCP [56]. Current scenario suggests that the interest of independent academic groups has considerably declined in conducting trials [57]. The reason for the same may be attributed to the extensive reporting and documentation needs making the task unnecessarily complicated [58]. Instead the situation demands that a well-designed clinical study must be amenable enough and should have simplicity in terms of adopted data handling and analysis techniques [59]. This idea of simplicity was also supported by statement of CONSORT (Consolidated Standards of Reporting Trials) which just had four steps in its flowchart [60]. There have been discussions between associates of following institutions: Oxford University, London School of Hygiene and Tropical Medicine, International Aids Vaccine Initiative, and Medical Research Councils of Uganda, United Kingdom & South Africa. Major problems identified post discussions were that the clinical trial teams do not have the resources or skills to establish and use software systems needed to manage data from trials in compliance with the ICH GCP. For developing nations the condition gets much more exacerbated particularly for non-commercial research centers with limited information systems infrastructure and support teams [61].

In particular, there is complete lack of guidance from regulatory agencies (US FDA and EMEA) about

- What should be the method deployed to evaluate the available various competing systems
- What should be the requirements for the clinical studies, where the data generation is required for a regulatory submission

These systems should work in a similar manner for both single-centered investigator-led small trials as well as large regulatory standard multi-centered randomized controlled trials. In addition to that these systems should be inexpensive, easily amenable (flexible), adaptable and compatible with current software already implemented, especially statistical and reporting software. Thus it seems prudent for international health research groups to associate their energies and finances to help with the advances of open systems needed for the task. This will not only address the challenges of managing (capturing, cleaning, extracting, and storing) of trial data but also the added desirable benefits of improving data reliability and quality [56].
It is a known fact that the clinical trial software is extremely costly [62]. However, the software which are very prevalent across the industry are Oracle Clinical and Clintrial, both being commercial database systems [63]. As these and other commercial software has high price constraints due to costly licenses, they also increase cost of trial depending upon the study size, thus making unaffordable for a typical non-commercial low budget trial. Hence organizations of developing country are left with no choice, thus making their trial(s) either non complimt with international regulations or the have to outsource the work of processing of case report to off-site locations [56].

2.5.3 Glimpse of CDM: European scenario-Heterogeneity prevails in CDM

The deployment of Clinical Data Management Systems (CDMS) has become important in clinical studies to handle the growing need of data upsurge i.e. its collection and analysis. There is overall increase in the use of these electronic data management systems. The ECRIN (European Clinical Research Infrastructure Network) data management working group [13] conducted a two-part standardized survey that evaluated:

- quality management for clinical trials
- data management
- software tools

Questionnaires data received through the survey of about 90% of centers having a CDMS in routine use revealed [13]:

- CDMS used for clinical data management are very heterogeneous
- for academic clinical trials, data management and quality management resources are missing in Europe i.e. achieving efficient data management for GCP compliant global clinical studies is still a venture full of challenges for academic research centers.
- Although quality management systems for data management are in place in most centers/units, there exist some deficits in the area of system validation.
- Implementation of standards like CDISC (Clinical Data Interchange Standard Consortium) is difficult due to the heterogeneity of data management software solutions.
- disadvantages of use of commercial software for academic research:
- high costs
- risks regarding delivery
- future software maintenance
- high software licenses
- Restricted integration with other systems
- High cost and difficulty in upgrading the IT (Information and Technology) infrastructure

- EDC (Electronic Data Capture) systems require extensive training on problems encountered and quality management; best practices need to be established to handle following difficulties with its different versions:
  - CRF pages are not displayed correctly (e.g. missing data field boxes).
  - EDC systems often may generate unnecessary queries.

- Due to limited human resources, deficits were observed in:
  - quality management,
  - system validated CDMS
  - Documented evidence for audits of data management.

Streamlining of clinical research is needed for improving healthcare for patients. Also, heterogeneity can be avoided by use of CDISC standards [13].

Above survey has showed several problems with data management systems in clinical trials conducted at academic centers. With an aim to bring ECRIN and other data centers to the same level of quality and standardization and to make them evolve towards a common quality level a report was developed using an approach as mentioned below [64].

Use of CDMS has increased in clinical research due to escalation in the amount of data collection which requires scalable processing and sizable analysis. Following two data collection methods prevalent were deployed for CDM:

- eCRF (electronic Case Report Form) : Remote Data Capture (RDC) where the data is remotely captured in electronic CRF
- pCRF (paper Case Report Form) : Using traditional paper CRF [65].

It was found that there are still many challenges in development and maintenance of a suitable data management environment [64]:

- Various software products as one of the reasons for heterogeneity observed
- Open source or commercial products supported by industry are less preferred as compared to the different proprietary software solutions [64]

- Deficiency was observed in
  - quality of data management
  - computer system validation

- Adequate levels of data management could not be observed due to limited human and financial resources

- For international clinical trials, complexities of running a local IT/data management center were underestimated.

- There was no widely recognized, acceptable, specific, practicable, simple and open (freely available) standard for CDM activates compliant to GCP with the accompanying IT platform.

ECRIN standard were developed as per the requirements of ICH GCP and European directives. In addition, other relevant documents were also consulted [64]. The statements mentioned in these documents are often too general to the requirements of specific clinical trial processes and structures. Standard requirements (Europe wide document) were developed as generic as possible (not considering specific national standards but most national specifications covered) but also specific enough to be useful for the conduct of a data center to support clinical research in multinational clinical trials.

A computer system used for clinical studies must be seen as part of the IT infrastructure of the trial center using the same, specifically designed to support clinical trials [64]. It is not only important to examine not just trial specific components of data management systems (in particular the clinical data management application by the DM group but also the general aspects of data management systems that are used by all trials [64].

Certification policy was developed and based on the same accreditation as an ‘ECRIN approved data centers’ was granted to satisfy implementation of GCP-compliant standard quality level for CDM performance in pan European trials for academic research. The pilot phase of the ECRIN certification program conducted in late 2011 for European data centers has led to a considerable revision (version June 2012) of the original ECRIN standards [53].

A list of standard requirements was published for data and information technology (IT) management in trials units with the following purpose:
- Based audit findings units ability was assessed for compliance and effective management of data
- Clear understanding should emerge from these standards for good practices and regulatory needs, so that they may be regards as general guidance for managing and establishing services for high-quality data management for non-commercial trials units in Europe.
- Standardized and structured approach was adopted to generate these new standards, with several rounds of discussions, feedback, dissemination, telephone conferences and face-to-face meetings to steadily iterate for common consensus for data center requirements.

### 2.6 Indian Regulations for conduct of Clinical Research

The clinical research market in India is likely to grow at a constant rate of 20-25 percent [66]. Although India shows steady increase in clinical trials registered since the introduction of clinical trial registry of India (CTRI), still it lags behind established countries in the field of clinical research [67]. Evolution of modern clinical research in India has occurred with the major breakthrough by ICMR New Delhi, with its guidelines for biomedical research on human subjects in the year 2000. This was updated in the year 2006 ‘Ethical Guidelines for Biomedical Research on Human Participants’ [9].

Schedule Y of the Drugs and Cosmetics Act 1988 has recognized the regulatory guidelines for clinical trial authorization in India. The schedule compelled the industry to carry out Phase III clinical trials for registration of a new drug and supported growth of a primarily generic Indian pharmaceutical industry. However, this schedule had insisted that the clinical studies to be conducted in the country must be done at a phase below (lower) than its global rank. This had obstructed overall amalgamation of Indian clinical industry to align with world-wide clinical development [27].

Revision of Schedule Y in January 2005 was an important turning point as it has widen the narrow and restrictive definitions of clinical trial phases, as compared to its previous versions, by providing pragmatic definitions for clinical trial phases I to IV, acceptance of concurrent Phase II-III, as part of global clinical trials [68]. The earlier constraints on number of subjects
and study sites in early phases, as specified in Schedule Y 1988, were revoked- allowed to grant the freedom to the sponsor company to decide these in relation to protocol requirements [27].

Indian GCP guidelines of 2001 were legalized in Schedule Y 2005. This schedule detailed GCP obligation of investigator, EC (ethics committee) and sponsor and suggested templates/forms/structure for critical documents e.g. informed consent, protocol (report), EC approval, reporting of serious adverse event [27], but of CDM processes to be followed. These laudable amendments in Schedule Y have been a major step to protect the integrity of Indian clinical trial industry, in conducting ethically compliant trials and providing the much-needed regulatory support to GCP guidelines. CDM as a critical activity of a clinical trial also need to evolve into a standardized procedure.

However, Indian clinical research is at a perilous phase as the execution of local regulations and global guidelines does not seem to be in congruence [69]. These guidelines do not provide explicit recommendations as to how each CDM step should be executed to achieve standardization across industry. There is also no supplementary or detailed documentary assistance in ICMR (in public domain at the time of writing of this thesis, 2015) to my knowledge [9]. With more number of qualified clinical researchers, data managers and overall scientific developments, there is an immediate need for regulatory processes to be streamlined [70].

Advancement of clinical research promises new and innovative therapies and vaccines with upgraded technologies. This calls for ongoing need for updation of regulations, guidelines and recommendations for better safety, efficacy and quality of Intellectual Property with less time to hit the market in a cost effective manner. These continuous scientific developments will come true with headway progression in instream branches of clinical research like ‘Clinical Data Management’. Overall conduct has to be done within the upgraded robust regulatory and sound ethical framework, necessitating major apprises in current procedures or best practices, followed by subsequent standardization [70].

Indian regulations are continuously advancing and becoming more and more stringent in order to safeguard the interest of trial subjects and ensure drug safety. One such example is the letter by DCGI (Drug Controller General of India) which says that if there is any delay of
more than 6 months in the start of the trial then the same should be brought to the notice of the authorities for necessary action (Figure 2.4).

![Figure 2.4: DCGI Notice](image)

This notice illustrate that the Indian regulatory authorizes are receptive to the ideas and committed to create a transparent environment which is in the interest of common public.

### 2.7 Indian Regulatory Requirements: CDM for Vaccines

Each country has its own sets of guidelines to conduct clinical research. The focus of this thesis is definition, implementation and standardization of CDM activities in the context of vaccine trials in an Indian pharmaceutical company. Indian GCP, the major guidelines available in India, and its application in vaccine trials is discussed below. The details are described (Table 2.1) of the mandatory regulatory obligations, mentioned by CDSCO in
Indian GCP for streamlining the clinical studies, with reference to CDM, and endorsed for adoption by the DTAB (Drug Technical Advisory Board), the highest technical body under D&C (Drugs and Cosmetics Act) [8].

Inferences drawn from these guidelines depend upon the individual perceptions as there is no standard format, checklist or illustration given in the guidelines, depicting how to fulfil these requirement(s).

Of note, Indian GCP includes segments not covered by ICH GCP e.g. vaccines, herbal preparations, devices etc. [25], but there is an urgent need to define in details the CDM steps for each of the segments mentioned. The universal face of drug development stresses on the fact that both Indian drug regulators and industry pay more attention to world-wide acceptable CDM best practises. Utmost attention is needed for streamlining of guidelines with respect to global standards to evolve into the era of standardize best practises for CDM procedures.

Table 2.1: Requirements of Indian GCP [8], relevant for CDM

<table>
<thead>
<tr>
<th>Information as mentioned in Indian GCP: relevant for CDM [8]</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Sponsor is responsible for securing agreement with all involved parties on the allocation of Protocol related and other responsibilities like Data processing.</td>
</tr>
<tr>
<td>It shall be the responsibility of sponsor to make arrangements for safe and secure custody of all study related documents and material for a period of three years after the completion of the study or submission of the data to the regulatory authority(ies) whichever is later.</td>
</tr>
<tr>
<td><strong>Record Keeping and Data Handling</strong></td>
</tr>
<tr>
<td>The basic concept of record-keeping and handling of data is to record, store, transfer, and where necessary convert efficiently and accurately the information collected on the trial subject(s) into data that can be used to compile the Study Report.</td>
</tr>
<tr>
<td><strong>Documentation</strong></td>
</tr>
<tr>
<td>All steps involved in data management should be documented in order to allow step-by-step retrospective assessment of data quality and study performance for the purpose of audit. Following the SOPs facilitates documentation. Documentation SOPs should include details of checklists and forms giving details of actions taken, dates and the individuals responsible etc.</td>
</tr>
<tr>
<td><strong>Correction</strong></td>
</tr>
<tr>
<td>All corrections in the CRFs or any other study related documents should be made in a way that does not obscure the original entry. The correct data should be inserted with the reason</td>
</tr>
</tbody>
</table>
Information as mentioned in Indian GCP: relevant for CDM [8]

for the correction if such a reason is not obvious. The corrections should carry the date and initials of the Investigator or the authorised person.

**Electronic Data Processing**

For electronic data processing only authorised person should be allowed to enter or modify the data in the computer and there should be a recorded trail of the changes and deletions made. A security system should be set-up to prevent unauthorised access to the data. If data is altered during processing the alteration must be documented and the system should be validated. The systems should be designed to permit data changes in such a way that the data changes are documented and there is no deletion of data once it has been entered. A list of authorised persons who can make changes in the computer system should be maintained. Adequate backup of the data should be maintained.

**Definitions Validation**

*Validation of Study*: The process of proving, in accordance with the principles of Good Clinical Practice, that any procedure, process equipment, material, activity or system actually leads to the expected results.

**Definitions**

*Validation of Data*: The procedures carried out to ensure and prove that the data contained in the final report match the original observations. The procedure is applied to Raw Data, CRFs, computer software, printouts, statistical analyses and consumption of Study Product / Comparator Product.

**Validation of Electronic Data Processing Systems**

If trial data are entered directly into the computer there must always be an adequate safeguard to ensure validation including a signed and dated printout and backup records. Computerised systems – hardware as well as software - should be validated and a detailed description of their use be produced and kept up-to-date.

**Responsibilities of the Investigator**

Investigator should ensure that the observations and findings are recorded correctly and completely in the CRFs and signed by the responsible person(s) designated in the Protocol. Laboratory values with normal reference ranges should always be recorded on a CRF or enclosed with the CRF. Values outside the clinically accepted reference range or values that differ importantly from previous values must be evaluated and commented upon by the Investigator. Data other than that requested by the Protocol may appear on the CRF clearly
2.8 Why CDM Standardization for Vaccine Trials

India is expected to play a pivot role for vaccine research and development. Though country is fast becoming the hub for the same, efficient and quality clinical data management still remains a challenge.

Invention of vaccines has been a major milestone in medical science. It has saved countless lives as compared to other drugs. But even today only very few companies manufacture vaccines despite that global demand exceed supply. This is generally attributed to vaccines safety concerns. In 2009, 23000 girls in India were vaccinated under a human papilloma virus vaccination program, a clinical trial by PATH (Program for Appropriate Technology in Health), a US-based NGO (non-governmental organization), in collaboration with the ICMR and the State Governments of Andhra Pradesh and Gujarat [71]. The trial was suspended/ stopped in April 2010 after the deaths of some of the vaccinated girls. Inquiry committee set-up by the Government pointed out major deficiencies in informed consent process and monitoring; these imperfections were attributed to the flaws in planning and conduct of the

Information as mentioned in Indian GCP: relevant for CDM [8]

<table>
<thead>
<tr>
<th><strong>Responsibilities of the Sponsor and the Monitor</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>The sponsor must ensure that electronic data processing system conforms to the certain documented requirements for completeness, accuracy, reliability and consistent intended performance (i.e. validation). The Sponsor must maintain SOPs for using these systems. The Monitor should take adequate measures to ensure that no data is overlooked. If the computer system automatically assigns any missing values – the fact should be clearly documented. Sponsor should safeguard the blinding, if any, particularly during data entry and processing. The Sponsor should use an explicit Subject identification code that allows identification of all the data reported for each Subject. Ownership of the data and any transfer of the ownership of data should be documented and intimated to the concerned party(ies).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Information as mentioned in Indian GCP: relevant for CDM [8]</th>
</tr>
</thead>
<tbody>
<tr>
<td>marked as the additional findings and their significance described by the investigator. Units of measurement must always be stated and transformation of units must always be indicated and documented.</td>
</tr>
</tbody>
</table>

---

2.8 Why CDM Standardization for Vaccine Trials

India is expected to play a pivot role for vaccine research and development. Though country is fast becoming the hub for the same, efficient and quality clinical data management still remains a challenge.

Invention of vaccines has been a major milestone in medical science. It has saved countless lives as compared to other drugs. But even today only very few companies manufacture vaccines despite that global demand exceed supply. This is generally attributed to vaccines safety concerns. In 2009, 23000 girls in India were vaccinated under a human papilloma virus vaccination program, a clinical trial by PATH (Program for Appropriate Technology in Health), a US-based NGO (non-governmental organization), in collaboration with the ICMR and the State Governments of Andhra Pradesh and Gujarat [71]. The trial was suspended/ stopped in April 2010 after the deaths of some of the vaccinated girls. Inquiry committee set-up by the Government pointed out major deficiencies in informed consent process and monitoring; these imperfections were attributed to the flaws in planning and conduct of the
study [71], [72], [73]. A well planned trial shall focus on the conduct all aspects of a clinical trial execution including CDM.

Thus ‘Safety and Efficacy’ of new drug product or vaccine are the prime focus of regulatory agency, prior to approval for marketing. Well planned pre-clinical and clinical studies need to be done before applying for approvals for marketing. Management of clinical data and its documentation are critical to the outcome of a study.

The interaction between the vaccine industries in the country with the regulator drug controller general of India on January, 2015, declared as the year of streamlined regulations, has initiated the process of charting out a new regulatory pathways for vaccines that is to promote a legal, regulatory and administrative framework for the safety of vaccines at national and international levels [74]. This supports the focused agenda to take India ahead and make it a global leader in vaccines R & D (Research and Development)[75].

Even with strong technical competences Indian biotech companies have not been able to position themselves as potential top global players in the emerging bio-similar arcade. Moreover, the Indian guidelines relating to biologics is nebulous, as in most countries of the world. Indian regulators and industry has to take a team approach to articulate common vision suited to deliver the outcome to the strategy [70]. This is possible by giving equal weightage for all streams of clinical research including CDM. For example, having sound strategies only for patient safety, with no detailed CDM guidelines, will not take the Indian biologics fraternity to the leadership position. The outcome of a trial is as good as the quality of the data collected as part of a trial and inferences drawn.

Data, ‘the clinical information gathered from each patient enrolled in study’ is the most valuable information and its handling and effective management is the most critical step of a clinical study. CDM group keeps the database ready so that the clinical information collected at the site could be entered into the database [2].

The clinical trials business continues to look for prospects to successfully and professionally manage data to bring meaningful outcome of a study. Some of the common inadequacies observed with execution of traditional non-standardized CDM in trials were [37]:

- It was required to develop paper CRFs (case report forms) valid for all countries in their local languages inviting problems associated with logical transcription and overall logistics.
- Concerns related with data legibility needed utmost attention
- The time required for review and subsequently to generate study forms like CRF, DCFs (Data Clarification Forms) were usually delayed or over stretched.
- Tracking of CRF, DCFs was a big issue
- Use of different version of same CRF caused augmented possibility of conflicts and inconsistencies in the records
- Subject enrollment, recruitment and randomization in the study were delayed, as there was no transparency in the medical status, causing major delays in screening [76].

Regulatory guideline demands credible, dependable, trustworthy, authentic and correct data. All changes to the data must be traceable through associated audit trail and with no potential for unauthorized access. All this can be achieved through precise adoption of industry prevalent acceptable practices and subsequent procedural standardization.

Heterogeneity prevails [13] as no standardized &/or validated industry-wide CDM procedural implementation platform exists, and in particular for vaccine trials. Such a framework may serve to be a clinical scale to evaluate and assist in regulatory review as per GCP mandate. Standardized CDM procedural definitions are needed to:
- optimize data management task,
- easy exchange of data between all stakeholders (partners, providers and regulators)
- to facilitate comparisons between results of different trials i.e. to facilitate meta-analysis

Excel® spread-sheets or Access database were the only tool (both the tools were invalidated) for data management in the past. Apart from decreasing the overall work efficiency, each and every step of the method was manually accomplished thereby was prone to more number of errors. Unless these tools are used after suitable validations, the data integrity may be questioned and the results may be deemed imprecise or possibly even professed as fraudulent [77].

Some of the major differences as observed in most of the protocol of the vaccine trials in comparison of other therapeutic products are:
- **Focus of outcome is mostly prophylactic**: protocol may include baseline test (physical examination /laboratory) establishing that the diseases under consideration is not pre-existing
- **Healthy subjects are involved**: screening for patient enrollment is done if there is no pre-existing aliment.
- *End points* may be established based on at least one or combination of the parameters not limited to the following: immunogenicity, tolerability, safety, reactogenicity, seroconversion, seroprotection and efficacy. [78].

- *Antibody titers are part of laboratory data*, as applicable: evaluation for antibody titer level at least twice i.e. at baseline and at the end is done to check for immunoconversion and seroprotection.

Though broadly the CDM processes may remain same as for other trials but fine tuning is required of individual steps for data processing for vaccine studies. Selection of industry prevalent common/best practises must be done in order to meet the GCP requirement of authentic and credible data.

Insight into the history illustrates that it is not easy to adopt a different methodology or technology into the clinical development life cycle. Implementation of best CDMS will not promise regulatory compliance unless it is supported by appropriate adoption of right CDM processes. Once the correct procedures are adopted, validated through audits, then this can be taken to the next level which is ‘procedural standardization’. Standardization of CDM procedures will reduce the operation cost by easy implementation of most of the task like data collection, its processing and transformation due to effortless re-deployment and ease of adoption. Standardization is an assurance for reasonably less number of procedural errors; it supports unbiased data analysis followed by timely regulatory submissions.