CHAPTER 3

RESEARCH METHODOLOGY
CHAPTER 3

RESEARCH METHODOLOGY

3.1 Indian Pharmaceutical Industry Survey

A survey was conducted to glean in industry perspective of the needs and challenges of CDM practice in India. The survey indicates lack of industry wide common CDM data standard. The ECRIN data management centers survey was used as a reference [13] for the questionnaire (Yes/No questions listed in Table 3.1). Anonymous input was used and the names of respondents are not disclosed to maintain confidentiality. The 46 respondents, 28 in number having more than 5-year of clinical research experience, representing 16 companies having their offices in India (Ranbaxy Laboratories Ltd., Biological E limited, CliniRx Research Pvt. Ltd., PATH Clinical Research, Kinapse Clinical Research, Novo Nordisk, Parexel International, Cognizant Technology Solutions, INC Research, Quintiles, Apcer Pharma, Panacea Biotec Ltd, Glenmark Pharmaceuticals, Venus Remedies Limited, Theorem Clinical Research and Tata Consultancy Services), responded to the questionnaire. The survey data was considered together with verbal discussion with experts from the CDM field, including representatives of various cross-functional teams involved in CDM and related activities at Panacea Biotec Ltd. The survey was related only to CDM practice in general; it was not connected to a specific clinical trial. No ethical approval for the survey was required because no patient data were collected [13].
Table 3.1: Questionnaire: Survey of CDM practice in Indian Pharmaceutical Industry

| Question 1: | ‘Will you welcome guidelines similar to the following in Indian regulations for CDM (GCDMP by SC adopting, CDASH/CDISC by SDTM, US FDA’s: Computerized Systems Used in Clinical Trials (05/2007), Electronic Records Electronic Signatures - Part 11, Scope & Application (08/2003), General Principles of Software Validation (01/2002) etc.).’ |
| Question 2: | Heterogeneity of CDM procedures exists in India? |
| Question 3: | Heterogeneity of software products for CDM exists in India? |
| Question 4: | Heterogeneity of CDM process/procedures will lead to deficits in quality management? |
| Question 5: | Heterogeneity of CDM procedures will lead to non-compliance to GCP? |
| Question 6: | Existence of limited human and financial resources for CDM for Indian Pharma Company? |
| Question 7: | In India, for various therapeutic segments, complexity of running an in-house CDM unit exists due to non-standardization of processes and lack of detailed CDM related regulations? |
| Question 8: | In India, no specific, simple, practical, free and open standard explaining the stepwise procedures for CDM activities of vaccine trials exists, for GCP-compliance data management? |
| Question 9: | Will you welcome guidelines stating step wise simple procedures for CDM activities for drug and/or vaccine trials? |
| Question 10: | Please state your total work experience in Clinical Research (CR Operation/CDM/PVG). |

3.2 Inferences drawn from Indian GCP in the context of CDM

Compliance with GCP, an ethical and scientific quality standard for designing, conducting, recording and reporting trials that involve the participation of human subjects, provides public assurance that the rights, safety and well-being of trial subjects are protected [7], [8].

In absence of any guidelines in Indian GCP, in general with respect to specific procedural steps to be adopted in the processing of a clinical trial data, and in specific CDM steps of vaccine as investigational product, inferences were drawn to define CDM procedural steps in the context of Indian GCP. Indian GCP demands that ‘raw data processing’ from clinical trial should be done in a manner so that the data is credible, reliable and accurate. Of particular note, it is discretion of individual organizations to make a choice of data management steps, based on the inferences drawn as per GCP, and how to conduct the CDM task for a particular intervention (therapeutic or prophylactic or diagnostic), sources of CDM procedural diversity necessitating CDM process standardization. Of further note, no internationally accepted and practiced unified guidelines are available in public domain for GCP related to CDM.
Table 3.2 gives the details of inferences in the context of CDM which were deduced from Indian GCP to maintain data integrity. Within the Indian- GCP Guidelines, an inference can be drawn from a number of sections, including sections 4.0 “Record keeping and Data Handling”; 4.1 “Documentation”; 4.2 “Correction”; 4.3 “Electronic Data Processing”; and also in sections 7.3 “Definitions: Validation-Validation of Study, Validation of Data”; 4.4. “Validation of Electronic Data Processing Systems”; 4.6.“Responsibilities of the Investigator”; and section 4.8. “Responsibilities of the Sponsor and the Monitor”.

<table>
<thead>
<tr>
<th>Information as mentioned in Indian GCP [7], [8]</th>
<th>Inferences drawn with respect to CDM [14]</th>
</tr>
</thead>
</table>
| The Sponsor is responsible for securing agreement with all involved parties on the allocation of Protocol related and other responsibilities like Data processing. | - Data Processing is an important step  
- Responsibility lies with the sponsor  
- Satisfactory steps and procedures for handling of data has to be identified by the sponsor  
- Once steps for data processing were identified, sponsor is responsible for its implementation.  
- for processing of data it is needed to have SOPs (standard operation procedures) |
| 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. | - All the documents related with CDM activities must be maintained for the duration of time as mentioned in the GCP.  
- The responsibility of safe archival of study documents lies with the sponsor for the said period. |

**Record Keeping and Data Handling**

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

- Procedures and Techniques should be established for the following:
  - Data Recording
  - Data Storage
  - Data Transfer
  - Data Conversion to specific format
<table>
<thead>
<tr>
<th>Information as mentioned in Indian GCP [7], [8]</th>
<th>Inferences drawn with respect to CDM [14]</th>
</tr>
</thead>
<tbody>
<tr>
<td>compile the Study Report.</td>
<td>- Ideally, companies which have set procedures addressing the above can compile the study data into study report.</td>
</tr>
<tr>
<td></td>
<td>- SOPs and minimum standards to implement the above must be framed internally</td>
</tr>
</tbody>
</table>

**Documentation**

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.

- Steps should be identified for CDM activities.
- SOPs should be framed to implement these steps.
- Documentation of these steps must be carried out in a manner so that retrospective evaluation of data quality and study performance can be easily analysed.
- SOPs must include checklist and forms. These checklist and forms has to be developed in-house.
- Following details must be part of SOPs and implementation procedure:
  - What actions was taken,
  - When the action was taken
  - Who took the action

As far as possible, all the information which may be required for the purpose of audits must be anticipated, and must be part of internal SOPs.
<table>
<thead>
<tr>
<th>Information as mentioned in Indian GCP [7], [8]</th>
<th>Inferences drawn with respect to CDM [14]</th>
</tr>
</thead>
</table>
| **Correction**  
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 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. | - Any changes to the data must be done in a manner so that original information is visible and thus audit trail is maintained.  
- All changes must have:  
  - Date of Correction  
  - Signatures of the person who did the correction  
  - Reason for changes if the reason is not obvious  
Anything which may be obvious to one organisation may not be obvious to another organisation or to regulatory body. So set procedures must be in place to handle the scenario. |
| **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. | - As stated in the guidelines only authorised person should be allowed to enter or modify data in computer  
- Qualifications or related requirement related with the expertise/experience of the authorised person is left to the organization.  
- The design of computer system which shall be used for data management will have what kinds of validations, needs more clarifications.  
- It is left to the organizations to adopt the policies and procedures to establish-  
  - audit trail of the work done  
  - security system to prevent unauthorised access to the data  
  - how to handle data changes/ modification during processing  
  - list of authorised persons who can access data  
  - adequate backup |
### Information as mentioned in Indian GCP [7], [8]

**Definitions**

**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.

**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.

### Inferences drawn with respect to CDM [14]

- Indian GCP do gives definition of ‘Validation’ which includes validation of study and data.
- Validation of study shall encompass the following to have consistent intended performance:
  - procedure,
  - process equipment,
  - material,
  - activity or
  - system
- It is left to the organization to adopt processes to proof that the steps implemented actually leads to the expected results.
- It is very important that the data in the CSR (Clinical Trial/Study Report) should match with original observation.
- It is expected from regulatory authorises that the procedures must be in place to proof that the data is validated.
- In this context, apart from other things, procedures must be identified and implemented to validate the following:
  - Raw Data
  - CRFs
  - computer software
  - printouts
- Raw Data refers to all records or certified copies of the original clinical and laboratory findings or other activities in a clinical study necessary for the reconstruction and evaluation of the trial [8].

---

38
Information as mentioned in Indian GCP [7], [8]

**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.

<table>
<thead>
<tr>
<th>Responsibilities of the Investigator</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
</tr>
<tr>
<td>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 marked as the additional findings and their significance described by the investigator. Units of measurement must always be stated and transformation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inferences drawn with respect to CDM [14]</th>
</tr>
</thead>
<tbody>
<tr>
<td>In case the data is entered directly into the computer, procedures must be identified and implemented by the stakeholders for validation including the following:</td>
</tr>
<tr>
<td>- Method to take printout with dated signatures and how to take the backup of the records.</td>
</tr>
<tr>
<td>- Computerised systems including hardware as well as software</td>
</tr>
<tr>
<td>- Up-to-date procedures must be in identified and implemented for maintaining detailed description of computerised systems use for processing of data.</td>
</tr>
<tr>
<td>It is the responsibility of the investigator to identify and implement processes so that the observations and findings are noted with no errors in the CRFs. Thus it is expected that there should be set procedures for the same.</td>
</tr>
<tr>
<td>- Procedures must be in place to handle laboratory data. Normal reference ranges must be predefined and should be enclosed/recorded in the CRF</td>
</tr>
<tr>
<td>- Procedures must be in place to handle outliers in the laboratory data.</td>
</tr>
<tr>
<td>- Data other than that asked in the study protocol must not be ignored. Set procedures to handle the same must be established.</td>
</tr>
<tr>
<td>- Units of measurement must always be used and where needed, it must be transformed. Procedures should be established and</td>
</tr>
</tbody>
</table>

39
<table>
<thead>
<tr>
<th>Information as mentioned in Indian GCP [7], [8]</th>
<th>Inferences drawn with respect to CDM [14]</th>
</tr>
</thead>
<tbody>
<tr>
<td>of units must always be indicated and documented.</td>
<td>documented, stating how to do the same.</td>
</tr>
</tbody>
</table>

**Responsibilities of the Sponsor and the Monitor**

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).

- It is the overall responsibility of the sponsor to identify, establish, document and implement the procedures for electronic data processing for its:
  - completeness
  - accuracy
  - reliability and
  - validation

- SOPs for the same must be in place.
- Study monitor must ensure that all data is entered and nothing is skipped.
- Procedures must be in place to handle missing data, especially when it is auto-filled by the computer.
- Blinding or Masking is the method of “control experimentation” in which one or more parties involved are not informed of the treatment being given [8].
- Adequate procedures must be established so that even during data processing the blinding is not unmasked.
- Procedures must be adopted in such a way that the confidentiality of the study subject is maintained, even during data handing.
- Unique Subject identification code must be used for data processing/reporting
- Procedures to establish ownership of the data and its transfer must be maintained /documented


**Regulatory Compliance Risk**

As per the requirements given in the Indian GCP [8], following should be explicitly the part of the software and processes selected for electronic data processing:

- Only authorized person should have access to enter or modify data.
- All the updating in the data should be marked by underline audit trail.
- Any illicit and unauthorized access to the data should be refrained by sufficient system level securities.
- If data is changed during handling the modification must be documented and the system should be validated.
- The CDMS should be designed to document data modifications in such a way that all changes are documented with no possibility of data deletion once it has been entered.
- A list of details of authorized personnel should be maintained who can make updation in the CDMS.
- Data backup must be adequate enough and should be properly retained.

Though eCRFs were not used in the current project, but the major part of data handling and processing was done electronically, therefore the above points were considered as regulatory compliance risk for the research work.
3.3 CDM procedural steps

CDM procedural steps were conceptualized to align with industry prevalent best practices. The steps to develop CDM processes for vaccine clinical trials were adopted after evaluating literature, considering expert’s opinions (regular telephone conversations and face-to-face discussions with experts both at Panacea Biotec Ltd, and across industry, and conclusion drawn from the outcome the above described industry survey), and calibrated to international guidelines or practices described below, avoiding copyright infringement, to suffice the requirements as per Indian GCP, and also international specifications; none of the below listed documents used as a reference are specific to vaccine trial CDM:

- GCDMP by a SCDM [79]
- ACDM (Association for Clinical Data Management) public website [80]
- ECRIN data management task [13], [53], [64]
- SDTM (Study Data Tabulation Model) [48], CDASH [51] and other similar CDISC documents [47].

The survey described in the section 3.1 highlighted the need for simple, practicable, practical and standardized CDM procedures apart from that there exist limited human and financial resources for CDM, to meet requirements of GCP and thereby GCP compliance. Interestingly to question “Heterogeneity of CDM procedures will lead to non-compliance to GCP”? the response was ambiguous. Approximately 53.33% respondents answered ‘Yes’ and 46.67% answered ‘No’, reflecting that CDM practices at variance does not equates to non-compliance to GCP in the absence of CDM specific procedural guidelines.

The CDM steps applied in the Myfive™ vaccine trial were audited by QA department. If there were no major or critical auditing findings, the working model established for Myfive™ trial was replicated in NUCOVAC® vaccine study to achieve process standardization. All these activities were carried out at Panacea Biotec Ltd., New Delhi, an Indian Pharmaceutical company, within the scope of Indian regulations [81] after the needed approval as mandated by DCGI, the Indian clinical trials regulatory authority and the Institutional Ethics Committee. Certain information in the following sections and results (Chapter 4) with respect
The study protocol for Myfive° and NUCOVAC® Vaccine Studies are briefly outlined in Table 3.3.

**Table 3.3**: Brief Information about the vaccine trials [81]

<table>
<thead>
<tr>
<th>Summary</th>
<th>DTwp-HepB-Hib (Myfive°) Vaccine Study</th>
<th>Pneumococcal (NUCOVAC®) Vaccine Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study Title</td>
<td>A Randomized, Multicenter, Open Label, Comparative Study to Evaluate the Immunogenicity and Reactogenicity of a Fully Liquid Pentavalent DTwp-HepB-Hib Vaccine (Myfive°, Panacea Biotech Ltd.) with Pentavalent DTwp-rHepB-Hib vaccine (Pentavac SD/MD, Serum Institute of India Ltd.) in Healthy Infants.</td>
<td>A Randomized, Open Label, Comparative, Single Dose Phase I/II Study to Evaluate the Safety, Tolerability and Immunogenicity of two Formulations (with and without preservative) of 10-valent Pneumococcal Polysaccharide Conjugate Vaccine (Adsorbed) NUCOVAC® in Healthy Adults.</td>
</tr>
<tr>
<td>Phase of Development</td>
<td>Phase II/III</td>
<td>Phase I/II</td>
</tr>
<tr>
<td>Number of subjects</td>
<td>600 (300 in each arm). Initially, the study was conducted in 48 healthy infants.</td>
<td>A total of 48 eligible subjects were enrolled in the study, 24 in each study arm.</td>
</tr>
<tr>
<td>Dose administration</td>
<td>0.5ml per dose by deep intramuscular injection</td>
<td>0.5 ml by deep intramuscular injection</td>
</tr>
<tr>
<td>Site of administration</td>
<td>Antero-lateral aspect of thigh</td>
<td>on Day 0 at the deltoid area of arm</td>
</tr>
<tr>
<td>Duration of protocol therapy</td>
<td>3 months</td>
<td>Single vaccine dose followed by 24 hours observation.</td>
</tr>
</tbody>
</table>

Regular verbal discussions within the CDM and cross-functional teams (in particular-Project Management team, Medical Writing team, Clinical Operations Unit, Biostatistics Unit, Quality Assurance (QA), Pharmacovigilance (PVG) etc.) were carried out to gather individual team members perspectives and their understanding of the adopted CDM procedures to bring about company standards in the context of vaccine trials. There were no differences of
opinion within the CDM group. The various CDM parameters considered are listed in the Table 3.4.

Table 3.4: Feedback Questionnaire: Internal debate conducted in Panacea Biotec Ltd. [15]

<table>
<thead>
<tr>
<th>Details About the CDM parameters:</th>
</tr>
</thead>
<tbody>
<tr>
<td>With the adoption of ‘proposed CDM methodology for vaccine trials’ vs. ‘Excel® spreadsheets (invalidated)’ is it possible or not possible to address the following:</td>
</tr>
<tr>
<td>Compliance to Regulatory</td>
</tr>
<tr>
<td>Quality for very small quantity of Data</td>
</tr>
<tr>
<td>Quality for large quantity of Data</td>
</tr>
<tr>
<td>Remote Data Capture</td>
</tr>
<tr>
<td>Randomization</td>
</tr>
<tr>
<td>Blinding</td>
</tr>
<tr>
<td>Adapted (adaptive) Clinical Trials</td>
</tr>
<tr>
<td>Support to multi-arms study</td>
</tr>
<tr>
<td>Subject Confidentiality maintained</td>
</tr>
<tr>
<td>Double Data Entry</td>
</tr>
<tr>
<td>Data access by authorized person</td>
</tr>
<tr>
<td>Data security system maintained</td>
</tr>
<tr>
<td>Recorded audit trail</td>
</tr>
<tr>
<td>Hardware &amp; software - validation</td>
</tr>
<tr>
<td>Deletion of data once it has been entered</td>
</tr>
<tr>
<td>Adequate backup</td>
</tr>
<tr>
<td>Credibility of the data based on the study design</td>
</tr>
<tr>
<td>Data contained in the final report match with the original observations</td>
</tr>
<tr>
<td>Easy identification of the individual subjects, without compromising confidentiality for allocation of unique subject identification code.</td>
</tr>
<tr>
<td>Effective Data discrepancies handling</td>
</tr>
</tbody>
</table>
3.4 CDM model adopted: Myfive\textsuperscript{TM} vaccine trial

Figure 3.1 depicts the schematic of the methodology adopted for processing (data collection, transformation, analysis and report submission) of clinical trial data for the Myfive\textsuperscript{TM} vaccine trial, performed using paper CRF. The traditional method of use of paper CRFs are still used today in almost 30\% of active global trials [38].

The various CDM team participants, CDM a group activity to effectively deliver its responsibilities, were Project In-charge CDM (Project Manager), Database Administrator, Database Designer, DM (Data Manager), Data Entry Operator, Data Coordinator, SAS Programmer/Statistician, Quality Control Personnel, Quality Assurance Personnel, Medical Dictionary Coder and Trainer [17], [82]. These roles overlapped based on the project specific need and were clearly defined in the task ownership metrics of the data management plan (refer corresponding section 3.5.2).

![Diagram of CDM Model adopted for Myfive\textsuperscript{TM} vaccine trial](Image)
The entire CDM process was divided into three phases for the convenience of handling the project: Study Start-up Phase (Setup Phase), Study Conduct Phase (Execution Phase) and Study Closeout Phase (Report Phase) [83], [84], refer Figure 3.2. The major milestones of each of the phases are described below [16].

**Study Start-up**
- Paper CRF was used for the Myfive™ vaccine trial. Inputs were given for finalization of CRF from the viewpoint of database designer, data mangers, data entry operators and data coordinators [85]. This was sent back for review to the clinical operations team.
- On receiving final approved Protocol and CRF, following activities were done:
  - creation of DMP
  - creation of Annotated CRF
- Subsequent to this database designing was initiated. This included creation of study databases setup, views creation, validation programming as per final DVP (Data Validation Plan).
- Study was moved to production after Test Data Entry/ UAT (User Acceptance Test) on dummy subjects and updation of all QC findings
- Parallel to all the activates all desirable documentation was done and was filed in MDMF (Master Data Management File), to name few:
  - SEC (Self-Evident Correction) document
  - relevant QCs reports for each of the above listed task
  - DEG (Data Entry Guidelines)
- Handling of Lab (Laboratory ) Data (Non –CRF Data)
  - Data Coding Guidelines

**Study Conduct**
- Double data entry (1st Pass & 2nd Pass) was initiated once the (annotated) CRF was received.
- Tracking of the CRF was done through ‘CRF Data Tracking Log’.
- Data Entry was done based on DEG document
- Reconciliation of data entry mismatch i.e. between 1st & 2nd pass entry, was done to avoid transcription or entry errors.
- Batch validation job was prescheduled and was triggered by the system to run all the validation programs.
- Process of Discrepancy Management included review of all queries generated as a result of batch run.
- Final queries were sent to the site to seek resolution in the form known as DCFs.
- All DCFs were sent along with the ‘DCF Tracking Log’.
- On receipt of resolution of the query from the site with dated signatures and correct reason for the error, the same was updated into the database.
- All queries were resolved and the resolution status of the same was checked. It was ensured that they must have the system status of ‘closed’ or ‘resolved’ at the end.
- MDMF was updated with relevant documents and the relevant QC reports.

**Study Closeout**

- If needed fine tuning of the data extract views were done as per the requirements of the Bio-Statistician and/or SAS review.
- Data Coding was done by using MedDRA for AE (Adverse Event) and SAE (Serious Adverse Event) pages.
- All the SAEs were reconciled with PVG (Pharmacovigilance) database.
- Preparation for database lock was done based upon the activities mentioned in the database lock check list.
- Database lock (Study Access Revoke and relevant archiving task) was done after getting all the relevant permissions from all the departments.
- DHR (Data Handling Report) was prepared and given to the end user (biostatistician).
- Final (clean) data was extracted for analysis & submission.
- All the relevant documentation was done and filed in MDMF.
- All relevant system generated reports were created and filed in MDMF.
- Relevant QC of data and database (before/during/after lock) was done and the report was filed in MDMF.
Figure 3.2: Swim Lane Diagram for CDM Work Flow [86]
3.5 Details of steps adopted at different CDM phases

3.5.1 Finalization of CRF: CRF and Protocol Alignment (CAPA) Document

Refer Figure 3.3 for diagrammatic representation of CRF development process

![CRF Development Process](image)

Figure 3.3: CRF Development Process [87]

To collect the values (data) in a consistent way, information as required by the Myfive™ vaccine study protocol was collected in the CRF. Inputs were provided to finalize the CRF to align with the trial protocol, keeping in view the requirements needed for database designing [17], [85]. A well formatted CRF is expected to minimize errors and thereby ensuring regulatory compliance. The data collected in CRF was used to conclude the study outcome (evaluation of safety and efficacy end points); serving as the basis of CSR and publication.
3.5.1.1 CRF Development Process

- Global standard pages, where feasible, (for example, demographic information, safety endpoints etc.) were used in the CRF. However, study specific pages (for example, efficacy endpoints, Laboratory parameters), where applicable, were developed.

- Draft CRF was received from clinical research operations. It was reviewed with special focus to see its alignment with respect to the requirements in the protocol and also to check the feasibility of database designing.

- All comments were documented in CRF and Protocol Alignment Document (CAPA).

- The CAPA was rechecked and the information which was missed, duplicate or incorrect was corrected.

- Final document was sent to the clinical research operations team for necessary updation in the draft CRF.

- The needed approvals from other teams and regulatory authority (DCGI and Ethics Committee) were obtained. The approved CRF was used for initiating study database design. Outline Structure of CRF as an example is shown below:

Figure 3.4 shows, for example, that for Myfive™ study, blood sample for antibody titres were collected at visit 1 and at Follow-up visit at Day 84+7
<table>
<thead>
<tr>
<th>Procedure / Assessment</th>
<th>Screening/visit 1 / 1st dose of vaccine (Day 0)</th>
<th>2nd dose of vaccine (Day 28+7)</th>
<th>3rd dose of vaccine (Day 56+7)</th>
<th>Follow up (Day 84+7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informed consent</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demography</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical History &amp; Physical examination</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concomitant medication and vaccination details</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vital signs</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Blood sampling for antibody titres</td>
<td>✔</td>
<td></td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Vaccination</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>ADR monitoring for 30min</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>ADR monitoring on day 1</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>ADR monitoring on day 2</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>ADR monitoring on day 3</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Serious Adverse Events since last dose</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study completion</td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>

**Figure 3.4:** Outline Structure of CRF depicting schedule of events for Myfive™
Figure 3.5 depicts, for example, that for NUCOVAC® study, blood sample were collected at all 3 visits.

![Figure 3.5: Outline Structure of CRF depicting schedule of events for NUCOVAC®](image)

3.5.1.2 Specific Points considered for CRF design: inputs on CAPA

- Data collection as per the study protocol; nothing extra or less was collected.
- Data collection to facilitate the end users; easy for end users to fill the CRF at the site.
- Same, consistent and logical grouping of sequence of questions/fields
- Simple rather leading language of questions/fields, avoidance of the use of difficult technical words
- Avoidance of data redundancy and data duplication
- Data collection as per the requirements of database designer to overcome software limitations.
- Avoidance of use of free text as it makes the analysis difficult and creates inconsistency in data. Also, the chances of missing out important information increases with the use of free text.
- Throughout consistency in the order of code list.
- Data collection as per applicable regulatory requirements; certain information in AE and SAE datasets is mandatory for submission.
- Handling of the missing data, partial dates etc. as per laid out rules and scientifically acceptable framework.
- Consistent use of standard units of measurements.
- Consistent use of global standard pages.
- CRF completion guidelines (CCG) for the site, Figure 3.6.

**Figure 3.6:** General Instructions for transcribing data incorporated in the CRF
Refer Figure 3.7 for flow chart for finalization of CRF respectively.

![Flow Chart for finalization of CRF](image)

**Figure 3.7:** Flow Chart for finalization of CRF [15]

### 3.5.1.3 Standard Operating Procedures: CRF Design, CDM perspective

In-house SOPs were developed considering the following [85]:

- Steps for the CRF development process
- Procedure to amend the CRF and version control
- Role and responsibility of team members
- Use of standard format for CAPA
- Use of standard pages of CRF
3.5.2 Data Management Plan

Figure 3.8 depicts schematic of the DMP development process.
DMP, a cross-functional process involving members from several teams, was created to plan all the activities from study set up phase to conduct and close out phases of the project, pertaining to management of the trial data [88], [89], [90]. This is the project specific document which elaborated the key steps which were followed during the study, along with the information about the major milestone as well as the timelines and task ownership metrics. The document was prepared before the start of the CDM process and procedures. DMP, a live document, was updated throughout the course of the study till the database lock; two versions in Myfive™ and four versions in NUCOVAC®. These updates were due to changes in protocol, process or procedures, timelines, roles etc. The document served as a reference for the project specific knowledge sharing for all the stakeholders for not only finding out the task assigned, but also specific details like the guidelines which shall be adhered, data management SOPs, details about how individual tasks to be performed etc. The data management plan was prepared with an objective to consider the several facets of data management, data preservation, metadata generation, reporting and analysis before the vaccine project begins; to ensure that the data is not only well-managed with documented evidence but retrievable and reusable [88], [89] [90].

There are no specific guidelines or regulations which define framework for the content of DMP. Literature reveals that it may be a bulleted list of items with few pages or thick document with an index [90], or made as per internal policy of the organization. A well-defined format was created to maintain consistency in the content so that the overall information which it conveys remains uniform and thereby compliance.
3.5.2.1 Elements of DMP

The various components of DMP adopted for Myfive™ and NUCOVAC® vaccine studies are defined in the Table 3.5.

<table>
<thead>
<tr>
<th>Table 3.5: Various components of DMP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Elements of DMP</strong></td>
</tr>
<tr>
<td>Authorization details</td>
</tr>
<tr>
<td>Distribution list</td>
</tr>
<tr>
<td>Introduction</td>
</tr>
<tr>
<td>Purpose</td>
</tr>
<tr>
<td>Scope</td>
</tr>
<tr>
<td>Study data quality management plan</td>
</tr>
<tr>
<td>Data base design</td>
</tr>
<tr>
<td>Dictionary and coding management</td>
</tr>
<tr>
<td>Protocol summary</td>
</tr>
<tr>
<td>Visit flow diagram</td>
</tr>
<tr>
<td>Data flow diagram</td>
</tr>
</tbody>
</table>

**Authorization details**

A signature sheet, the dated signatures on this page indicate ‘review and approval’ of the final DMP document.

**Distribution list**

Details about the circulation of the document as authorization page may not have signatures of all the team members.

**Introduction Section**

Outlined the process for collecting, reporting and communicating CDM activities along with the study-specific metrics, how each step of the data management process will be carried out and what documents would be created or collected during study conduct.
Purpose
Defined the objectives of DMP with respect to its aim for laying down steps for data management including briefs about database design, with emphasis on identifying who will perform what task, and how the quality management plan will be implemented and other pertinent related details. DMP, a continuous process, communicated when the document is expected to be ready or updated.

Scope
Outlined choice of CDMS (Clinical Data Management System), CDMS documentation, and areas not within the scope. For example, title page of the CRF not captured in the database for individual subject as the information was already available in the CAPA.

Study data quality management plan
It defined study data quality management plan. Listed the details of QC, information about the authorized team member to perform the task, and designated personal in case of absence.

Database design
This section defined key steps, unlike a software user manual that give step by step information, about database designing, for example:

- the fields picked for designing were created or reused from global library or aligned to other valid standards
- use of extra fields i.e. the fields which were not in the CRF but still used to facilitate data cleaning or address software limitations: Question ‘EXIT_CURRENT_PAGE’ used for conditional branching, facilitated navigation during data entry and data discrepancy review during query management.

Dictionary and coding management
Defined the dictionary (MedDRA) and applicable rules used for coding of verbatim terms mentioned in the CRF, for example, SAE, AE information etc.

Protocol summary
Overview included topics not limited to the following: Study Title, Protocol No., Version with date, Phase of Development, Indication, Study Objectives, Study Design, Study Centre, Study Arms, Study population, Inclusion Criteria, Exclusion Criteria, Planned No. of Subjects, Investigational Product, Dosage and administration, Duration of protocol therapy,
Expected duration of study, Methodology, Study Endpoints Primary Endpoints, Secondary Endpoints.

Visit flow diagram
Detailed information about the subject’s visits described in the form of flow chart (Table 3.6), to help understand the protocol requirements from CDM perspective. For example, it was referred to reply to the auditors’ queries for the facts related to subject’s visits.

Table 3.6: Visit flow diagram based on study protocol

<table>
<thead>
<tr>
<th>Visit Flow Diagram for MYFIVE™ study</th>
<th>Visit Flow Diagram for NUCOVAC®</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Visit 1</strong>: Enrollment (First vaccination (Day 0))</td>
<td><strong>Visit 1</strong>: Screening (Day 0-4)</td>
</tr>
<tr>
<td>• Informed Consent</td>
<td>• Informed Consent</td>
</tr>
<tr>
<td>• Screening of subjects</td>
<td>• Screening of subjects</td>
</tr>
<tr>
<td>• Induction and Exclusion criteria</td>
<td>• Induction and Exclusion criteria</td>
</tr>
<tr>
<td>• Blood sampling for Immunogenicity assessment</td>
<td>• Blood sampling for Immunogenicity assessment</td>
</tr>
<tr>
<td>• Vaccine administration (1st dose)</td>
<td>• Vaccine administration (1st dose)</td>
</tr>
<tr>
<td><strong>Visit 2</strong>: Second Vaccination (Day 28)</td>
<td><strong>Visit 3</strong>: Enrollment and Vaccination (Day 0)</td>
</tr>
<tr>
<td>• Reconfirm induction/exclusion criteria</td>
<td>• Enroll and randomize subjects</td>
</tr>
<tr>
<td>• Vaccine administration (2nd dose)</td>
<td>• Blood sampling for baseline and pneumococcal antibody levels</td>
</tr>
<tr>
<td><strong>Visit 3</strong>: Third Vaccination (Day 56 + 1)</td>
<td>• UFT and urine for drug use</td>
</tr>
<tr>
<td>• Reconfirm induction/exclusion criteria</td>
<td>• Vaccine administration</td>
</tr>
<tr>
<td>• Vaccine administration (3rd dose)</td>
<td>• 24-hr observation at study center</td>
</tr>
<tr>
<td><strong>Dropout/Withdrawal</strong></td>
<td><strong>Visit 4</strong>: Blood sampling for post vaccination safety assessment, laboratory evaluation and vaccine immunogenicity.</td>
</tr>
<tr>
<td><strong>Visit 4</strong>: Post Vaccination Follow up (Day 14)</td>
<td>• Adverse event monitoring for 4 weeks post vaccination.</td>
</tr>
<tr>
<td>• Blood sampling for Immunogenicity assessment</td>
<td>• SIE monitoring throughout study duration</td>
</tr>
<tr>
<td>• Adverse event monitoring</td>
<td><strong>Visit 5</strong>: Follow up (Day 28 + 3)</td>
</tr>
<tr>
<td>• SAF monitoring</td>
<td>• Blood sampling for post vaccination safety assessment, laboratory evaluation and vaccine immunogenicity.</td>
</tr>
<tr>
<td><strong>Complete study completion</strong></td>
<td>• Adverse event monitoring, UFT, SAF monitoring, study completion termination</td>
</tr>
</tbody>
</table>
**Data flow diagram**

CRF information flow i.e. how the data flow in the CRF occurred and methodology adopted for tracking. Table 3.7 shows the milestones of CRF data flow, starting from the investigators site; transcribing of data into CRF till the time of database locked. The CRF data was tracked to assure that data is not lost.

**Table 3.7: Data flow diagram for Myfive™ vaccine**

<table>
<thead>
<tr>
<th>Data Flow Diagram</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRF completed on Investigator's site</td>
</tr>
<tr>
<td>CRA checks data &amp; verifies from source document</td>
</tr>
<tr>
<td>CRF retrieval &amp; submission to CDM by clinical monitor</td>
</tr>
<tr>
<td>Data entry (in QC)</td>
</tr>
<tr>
<td>CRF routed to Archive</td>
</tr>
</tbody>
</table>
Data Clarification Form (DCF) flow

Flow of queries generated for discrepant data points so as to get the resolution from the investigators site or medical monitor (Table 3.8).

Table 3.8: DCF flow diagram for Myfive\textsuperscript{TM} vaccine

<table>
<thead>
<tr>
<th>DCF Flow Diagram</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approval of legitimate queries from Medical Monitor / Designee on Draft DCF (as applicable)</td>
</tr>
<tr>
<td>DCF are raised</td>
</tr>
<tr>
<td>DCF sent to medical monitor/project manager/designee for Resolution</td>
</tr>
<tr>
<td>Resolutions are obtained on the DCF, along with dated signatures.</td>
</tr>
<tr>
<td>Resolved &amp; Unresolved Queries submitted to CDM</td>
</tr>
<tr>
<td>Resolved DCF for Database updation</td>
</tr>
<tr>
<td>Document in MDMF in DCF Tracking Log</td>
</tr>
<tr>
<td>Inputs from Medical Monitor</td>
</tr>
<tr>
<td>Inputs from Investigator</td>
</tr>
<tr>
<td>Unresolved to be re-raised/reviewed (To be solved by Medical Monitor/CRA/Investigator)</td>
</tr>
</tbody>
</table>
List of acronyms & definitions
Listed acronyms & definitions used in the DMP for the project.

Database security
Referred to the issues related with the database security not limited to the following-prevention of unauthorized access, inhibition of physical harm to database servers, addressing performance restrictions of the software, resolving issues related with the design imperfections, prevention of data corruption and/or loss etc.

Though most of the above mentioned tasks/responsibilities are outside the preview of CDM team and falls in the bracket of IT department, but CDM team coordinated and reorganized the tasks in order to ensure that these problems do not occur, or in case the problem occurred to perform UAT or PQ (Performance Qualification) to ensure that it has been resolved by IT to CDM and QA teams satisfaction.

Data handling (Data Entry and Processing): Overview
It addressed the data processing methodology adopted for the study in brief:

- Information for Study Design Document: The document was created to aid in database designing. This gave the information about the number of patents and patient-positions were created in OC, including screen failures. Also, it gave the information about the CRF submission to CDM unit for data processing.

- Project Kick-Off Meeting was conducted at the start to discuss the timelines, data entry guidelines etc.

- Data entry errors were minimized by performing double data entry to ensure correctness of copying the data

- Pre-entry review of the CRF was accomplished before data entry to deal with problem data like data entry of illegible, notations or comments in the CRF margins.

- All edits and changes to data were captured with audit trail. The CDMS system had the inbuilt feature to handle the same.

- Validation Procedures: To minimize the scope of data entry errors, the data was checked for inconsistencies’ by triggering programmed procedures using PL/SQL language (Procedural Language/Structured Query Language).
- Identifying and Managing Discrepancies:
  - All discrepancy related activities were controlled as per DCF flow for the study.
  - Discrepancies identified were related to blank values, ambiguous data, lab values (outliers), protocol violation etc.
  - DCF for each query was generated and was sent to the medical monitor/designee for review and resolution
  - DCFs were first created in the draft & were later updated to final version. However, where needed, DCF text was restructured based on the inputs of the medical monitor for enhanced clarity for the convenience of the study site staff.
  - The database was updated as per the valid resolutions received on the DCF. If the resolution(s) received was found to be invalid, the DCF was re-raised/resent and the cycle of DCF flow was repeated.

- Collecting Safety Data (AE/ SAE): All AE/SAE data was collected in the CRF and managed like any other data. However, MedDRA coding was done for verbatim term(s).

- Creating Report & Transferring Data: Due to seamless integration of O.C. database to SAS, there was no need for external data transfer/export. It is expected of CDM team to provide the clean data to the Biostatistician for analysis report generation.

- Closing: Closing: ‘Soft lock’ and ‘Hard Lock’ for the study.
  - Soft lock of the database was done as per the date mentioned in the ‘Information for study design in OC’ document.
  - Soft lock was done when- data entry was complete, MedDRA Coding done, SAE reconciliation over and all the queries were be resolved (as per checklist)
  - Care was taken that during the study soft lock the database had the status ‘lock’ & not ‘freeze’. Any inconsistency, if identified was rectified.

- This was followed by hard-lock at the planned interval, which was approximately between 6 to 10 months from the date of soft lock. Also, before the hard lock, it was ensured that the authorization from all the respective teams has been achieved.

- Sample Size and Sampling Method for QC: described the number of CRFs (subject) which had undergone QC process.
- 100% QC of the critical parameters was done. The critical data included- efficacy, safety, subject participation and study completion information.

**Communications**

This section detailed about the communications during the period of study conduct. Subsections were created for each task.

**List of investigators & investigational sites**

Listed the information about the study site (Pune for Myfive™ and Pondicherry for NUCOVAC®).

**List of SOPs used for the project**

Listed SOPs developed, updated and used during the course of the study. SOPs are not described in this thesis due to organizational confidentiality constraints, only the sample list is shared below:

- Screen Design, Data Entry Guidelines, Data Entry and Data Verification of Clinical Trial Data
- Master Data Management File
- Preparation of Data Handling Report
- Data Extraction of a Clinical Study from Oracle Clinical to SAS
- Data Protection, Security and Confidentiality
- Creation of Data Management Plan
- Creating Edit Specification and Data Validation Guidelines for a Clinical Study
- Designing a Database for a Clinical Trial
- Preparation an “Annotated CRF” for CDMS Database (OC Database)
- Query Generation and Resolution
- Change Control of Software and Computer Systems
- Receiving of CRF for Data Entry by CDM
- Reconciliation of Serious Adverse Events of Drug Products Undergoing Clinical Trials
- Quality Control of Database of Clinical Trial Data in Oracle Clinical
- Filling and Maintenance of Logs in a Clinical Trial
- Locking of Database of a Clinical Trial
- Coding of Clinical data
- Workflow in Clinical Data Management
- Allocation of Rights as per Roles defined for the Trials conducted in CDMS
- Quality Control and User Acceptance Test

**Team Structure, Scope of Work & Authorization (Task ownership & performance matrix with timelines)**

Detailed information about the teams involved in the study, and ‘Task Ownership & Performance’ metrics as a separate subsection.

**Regulatory guidelines**

This section detailed about the guidelines applicable for the study, for example

- Indian GCP Guidelines
- Ethical Guidelines for Biomedical Research on Human Subjects (ICMR)
- Schedule Y (Drug & Cosmetic Act, Govt. of India)

**Self-evident correction document**

Self-evident or obvious corrections (SECs) in the data that were carried out in the CRF data, identified throughout the study, judged to be incorrect as per the inputs of medical monitor. These corrections were not applied to efficacy and safety variables, which can change the meaning of the data and thereby misleading study outcome.

**3.5.2.2 DMP Development Process**

The Figure 3.9 depicts a Flow Chart for finalization of DMP.

- The draft version of the document was prepared as per the predefined template
- Once the draft document was ready, it was then circulated to the respective team members for procuring their suggestions/comments.
- As applicable, if a particular section was required from other team member(s), then the email request was sent for providing the section to the author.
- On receipt of the comments/sections, it was reviewed and if found suitable, the same was compiled in the study DMP.
- Subsequent to this, review of the document to check for consistency and correctness
- The signing off of final DMP was done. The activities listed under DMP elements were implemented and executed in the project as per the DMP plan.
- Any change during the study course called for repetition of the above listed DMP activates till the finalization of the next version of documents.
Figure 3.9: Flow Chart for finalization of DMP [15]
3.5.2.3 Specific points considered for creating DMP

- DMP was prepared in compliance with applicable regulatory guidelines.
- It clearly specified the data management processes involved in the project.
- Specified time lines for the first draft of all the documents to be ready.
- Mentioned key team members involved in the study and their opinion incorporated in the document, as applicable.
- Task ownership was defined along with the matrix i.e. timelines. For example- who shall do the task like data collection, data handling, data quality control etc. and when it will get completed documented.
- Documentation of project kick-off meetings called so that all the team members understand their task (role) thoroughly, along with the applicable data entry guideline.
- Up-to-date DMP keeping with version control. Current version signed and circulated to the concern team members.

3.5.2.4 Standard Operating Procedures: DMP

In-house SOPs were developed considering the following:
- Consistent format was used with sufficient number of sections to encompass all the information related with the planning of CDM steps.
- Established procedures to amend and version control of the document.
- Study specific task-ownership matrix
- Details about the deliverables
3.5.3 Database Setup, Data Processing and Documentation

Figure 3.10 depicts diagrammatic representation of Database Setup, Data Processing and Documentation.

![Database Setup, Data Processing and MDMF Documentation Diagram](image)

**Figure 3.10:** Database Setup and Data Processing

On completion of CRF designing and subsequent study initiation, CRF was sent to the site by clinical research operations team for data collection as per the study protocol. Filled CRF with patient data was sent back by the study site personal to the CDM unit for data processing.

The following steps were followed for database set-up:

- CRF Annotation
- Designing of Database
- Validation Programming
- Data view creation
**Database setup and Processing: Work Flow**

- CRF was annotated to initiate database designing.
- Database was designed based on the annotated CRF, on Oracle Clinical version 4.5.3
- Database was tested using dummy patients
- QC findings were updated
- Validation programming was done as per the pre-approved edit specification document. This document had the list of all the edit checks to validate data as per study protocol.
- QC of validation programming was done using dummy data.
- Updation of QC findings was done
- Final QC of database was performed
- Database was released in production for data entry
- Each CRF was reviewed for accuracy and completeness before initiation of data entry
- Double data entry was performed based on Data Entry Guidelines prepared after discussions with study medical monitor.
- Data verification / Data reconciliation was performed after data entry was complete; the differences in both passes of data were reconciled with the help of hard copy of the CRF.
- Total number of CRF entered was verified from the ‘CRF Submission Log’ to ensure that all the CRFs were entered in the database.
- All key documents were generated during the process and filed in MDMF

**3.5.3.1 Database Set-up**

Database, a clinical software application, was built to facilitate the CDM tasks [56], [89]. Database set-up was done using Oracle Clinical version 4.3.2 for capturing subject data, software that is suited to facilitate easy regulatory compliance by fulfilling the requirements, as mentioned in Indian GCP. Database was designed in such a way so that it could capture all the information exactly in the same manner as described in the CRF. Database setup was done to facilitate data cleaning, reviewing, analysis and reporting.
3.5.3.1.1 Annotation of CRF

An annotated-CRF was created that links the data points in the CRF with the fields in database. Annotated-CRF, identification of the variables in the database, served as the starting point for database design and creation of datasets for analysis or submission. Annotated CRF was reviewed by CDM member and biostatistician for its correctness. Standard annotations, where feasible, were consistently used to facilitate quick rollout of sister/extension studies, meta-analysis and for achieving process standardization. An example of an annotated CRF is given in Figure 3.11. In questions with discrete value options (like the variable ‘Tick’ having values Yes and No as responses), all possible options were coded, appropriately [17].

![Figure 3.11: Example of dummy Annotated CRF](image.png)

Specific Points considered for creating Annotated CRF

- Annotations reflected mapping of variable in the CRF with fields in study database
- Annotations were assigned as per in-house standards and also in reference to CDASH.
- Annotations were done using ‘text’ and were searchable [91].
- PDF file was used, as applicable.
- CRF annotations were not hand written.
- QC of the annotated CRF was done by independent CDM personnel before initiation of the database designing.
- Standard color codes were followed for CRF annotations. For example:
- Approved annotated CRF were assigned a version number and version date, before signing-off.

3.5.3.1.2 Designing of Database

Database design incorporated the procedure of creating a comprehensive data model of the study database. Coherent data model was created in OC version 4.5.3., which had required logical and physical design adoptions for storage, backup and archival of information related with the parameters given in the study CRF (Figure 3.12). Industry prevalent CDMS tools are CLINTRIAL, RAVE, ORACLE CLINICAL (OC), MACRO, and eClinical Suite [17]. The study specific database was designed based on Data Definition document, known as Collation List, having well-defined information about the metadata. Study details like objectives, clinically planned events/intervals, protocol visits, list of investigators, study sites and patients positions were defined into the database; character layouts for the CRF was designed for data entry [17], [92].

Following is the brief information about the study database design, Figure 3.12.
Designing of study database was done in the following sequence- DVG (Discrete Value Group), Questions, QG (Question Groups), DCM (Data Collection Module), DCI (Data Collection Instrument) and DCI Books, following hierarchal structure as in OC database.
Figure 3.13: Questions, Question Group and Discrete Value Group Created in Glib
The following Order of Organization of Oracle Clinical Database was followed:

- **Data Collection Instrument (DCI) Books:** The DCI book, a group of DCI, offered the order of data entry in OC.

- **Data Collection Instrument (DCI):** One DCI corresponded to a single page of CRF. The DCI has a single DCM or a collection of DCM.

- **DCM (Data Collection Module):** One DCM depicted a single CRF page or the screen for data entry. The DCM has a single QG or a collection of QG, in which the interrelated data can be transcribed for a clinical visit or CPE (Clinically Planned Event).

- **QG (Question Group):** The QG, a collection of questions, addressed the common or related information (Figure 3.13).

- **Questions:** Questions were created based on the Annotated CRF in the CDMS. Use of existing Questions from Glib (Global Library) was preferred. (Glib is the library of various database objects, created by CDM team as per company’s internal policy). Each variable or filed in the CRF which stores the information or data is known as Question in the OC database (Figure 3.13).

- **DVG (Discrete Value Group):** The character response for the question with fixed options was designed with the help of code list known as DVG (discrete value group) (Figure 3.13).

Based on the inputs (for example study duration, time unit for events etc.) mentioned in ‘Information for study designed’ document received from clinical operation team, study setup was done in OC, (Figure 3.14).

**MYFIVE™ study setup in Easy Study Design**

**NUCOVAC® study setup in Easy Study Design**

*Figure 3.14: Easy Study Design node of OC*
Once Question Groups was designed and study was setup in ‘Easy Study Design’ module of database, the designing of DCM (Data Collection Module) was initiated. DCM has one or more related Question Groups and each DCM was replica of single CRF page. DCM was prepared in the Definition node of the CRF (Figure 3.15).

DCM was prepared in a manner so that it was equivalent to data entry screen (character layout, Figure 3.15) for the study. DCM was mapped to DCI (Figure 3.16). Global standards pages were created using standard DCMs, as applicable, to achieve consistency.

**Figure 3.15:** Path for creation of DCM, Character Layout and Study DCM
Figure 3.16: An example DCI window for the study

DCI was then mapped to DCI Book. DCI Book is equivalent to a blank CRF (Figure 3.17).

Figure 3.17: An example: DCI Book for the study
Specific points considered for Designing Database

- On completion of study database designing, test data entry was performed using dummy patient data. Following parameters were checked for consistency [15].
  - Data Entry Screen Design is replica of the CRF
  - Cursor movement
  - Length of questions
  - Text of questions
  - Marking of Mandatory fields
  - Use of Correct Default prompt and SAS Labels
  - Correct use of DVG and its subsets
  - Applying upper and lower bounding (ranges) to the questions
  - Hard Coding of the fields, where required
  - Use of qualifying questions
  - Use of repeat questions groups for the tables
  - Use of indicator questions
  - Navigations through conditional branching
  - Use of subsets in DCM
  - Character layout generation

- Logically related Questions were grouped together, to form Question Group.

- Only study specific unique DVG, Questions, Question Group, and DCM were prepared in Glib. If these objects were already present in the Glib, they were reused to avoid duplication and redundancy.

- For Questions with fixed response options, as mentioned in the CRF, standard DVG (s) were used consistently throughout the study. Use of DVG facilitated harmonized responses and easy analysis.

Consideration of the above described points resulted in a well-designed database and served to avoid snags during validation programming, production data entry and discrepancy management. It facilitated copying these objects (validation, and views) from the global library into the DCM [15].
3.5.3.1.3 Validation Programming/Edits Checks

Cleaning of the raw (unprocessed), ambiguous and erroneous data was carried out by programming the checks into the database. This was done to validate the data as per protocol requirements, so that clear, consistent and quality study data was submitted to the biostatistician for analysis. Validation checks were programmed based on the pre-approved edit specification document. Where applicable, edit checks were copied from global standard pages to help achieve consistency of output, for example, for similar queries uniform text language was used in the DCFs. New programs in PL/SQL were developed for study specific unique pages.

Validation programming (Figure 3.18), was initiated after updation of the QC findings of test data entry. This was accomplished based on the approved edit check document i.e. DVP (Data Validation Plan) based on the inputs from medical monitor. Validation was done to ensure that the clinical trial data collected in the CRF complies strictly with the requirements of the protocol. Thus whenever a data point failed to meet the same then the query was triggered [92].
Path for creation of validation programming

Selection of Study For Validation Programming

Validation Procedure of MYFIVE™ vaccine

Validation Procedure of NUCOVAC®

Figure 3.18: Validation Programming for Discrepancy Management
Specific points considered for creation of DVP

- Possibility of protocol deviation was ruled out, for example, for violation of inclusion/exclusion criteria.
- Redundant/duplicate data were avoided, if collected must be checked, for example, date of signature by the investigator on every page of the visit.
- Designed to avoid data inconsistencies, for example date of medical history recorded should be prior to the date of Screening.
- Endpoints properly evaluated, for example, use of correct units in titer values.
- Standards check used for standard pages.
- UAT was performed to avoid programming errors.
- Inputs were taken from Clinical Operations, PVG, QA, and Biostatistician.

3.5.3.1.4 Data Extract View

Data views were generated in SAS format to extract the processed data to facilitate analysis by biostatistician. Most of the test views were created at the time of database designing but were later modified based on the comments of end user (biostatistician). Most of the view templates were copied to create subsets of DCMs to have consistency and thus to achieve standardization of CDM processes.

Figure 3.19 shows Data View created in OC.

![Figure 3.19: View Template for Data Extraction](image)
Specific points considered for creating Views for Data Extraction

- DCM and its subset have common Views
- Global Views copied from global pages
- Mapping of the variables in the views examined so that the values of columns should not get interchangeable or duplicated
- Views have correct SAS labels for analysis
- If the name of the variable is updated in the view then the same reflected in the annotated CRF
- View names are mentioned in the annotated CRF
3.5.4 Data Processing

Double data entry was performed after database was designed, tested and moved into production. It included first-pass and second-pass entry of the same data by two different entry operators. Data reconciliation was done by updation of differences in the data entry by verifying form the hard copy of the CRF.

- Data Entry
- Data Verification/Reconciliation

Brief Description about the Elements of Data Processing

3.5.4.1 Double Data Entry

Data entry was done based on the pre-approved document of the study; ‘Data Entry Guidelines’. The data was transcribed into the database by entering the same data twice by two different entry operators followed by reconciliation of the differences in the entities. This practice was not possible when traditional Excel® spread-sheets (invalidated) were used. This was trailed by reconciliation(s) and rectification(s) of points with variances in the data values. Since audit trail records the past history of every single task/changes for every transaction, any change in the data values was marked with the information about the user, time stamp, reason for change etc.

3.5.4.2 Data Verification/Reconciliation

Differences between first and second pass data were identified by manual comparison with the hard copy of the CRF. Differences, if any, were updated.
3.5.5 Quality Control Process

The operational techniques and activities within the system of QA were undertaken to verify requirements for quality of the trial related activities have been fulfilled [8], [93]. The data quality was measured in terms of comprehensiveness, legitimacy, uniformity, timeliness, relevance and accuracy to check the data suitability for analysis and clinical study report. QC was performed after completion of major study milestones (DMP, Database Setup, Data Processing etc.). The QC task was carried by an individual skilled in the QC subject area and did not perform the task being evaluated. The QC findings were addressed appropriately; all rectifications were documented with the reason for change. Following are the details of the QC reports.

3.5.5.1 Brief Explanation about QC Reports

- **QC Report of CRF annotation**: CRF annotations were checked for consistency.
- **QC Report of Database Designing and Entry Screen Layout**: QC was done to check that the database screens are exact replica of CRF pages. This facilitated fast data entry and helped prevent data entry errors.
- **QC report of 1st and 2ndPass Test Data Entry**: Dummy data was entered into the database to check for errors in designing.
- **QC Report of Validation Procedures**: Validation checks programming was tested by running them against the dummy data. This was done to identify programming errors for example duplicate discrepancies.
- **QC Report of View Definitions, View templates and Data Extraction**: Data Views were updated to suffice the inputs given by biostatistician.
- **QC Report of Subject Enrollment**: Based on the information given in the document ‘Information for Study Design in OC’ the creation of patient-position and their corresponding mapping with the Investigator’s site were checked.
- **QC Report of MedDRA Coding**: MedDRA coding was done to have consistency in coding of verbatim terms.
- **QC Report of Discrepancy Status in OC**: The discrepancy status was ‘Resolved’ prior to the database lock. Figure 3.20 shows the screen short of review statues assigned to discrepancies revealed during the data cleaning process:
Figure 3.20: Discrepancy review status

- **100% QC of Critical Data Points:** 100% QC of critical data points, procured form the study medical monitors, was performed.

- **100% QC of Random sample size using √n+1:** Randomly few CRFs were picked and 100% checking was done against the database for avoiding transcription/logical errors. The sample size was either greater than or equal to √n+1.

- **QC Report of DCFs:** All the DCFs were checked for validity of the resolution received from the site and had the dated signatures from the authorized site staff.

### 3.5.5.2 Specific points considered for performing QC

- ‘Format Check’ i.e.
  - Database permitted only numeric entry for the fields where numeric response was expected
  - Only date was permitted to enter in the fields where date value was expected, else query was raised.

- Eligibility Criteria not fulfilled
  - Warning flagged if inclusion and exclusion criteria of the study were not met.

- Ranges were checked with proper validation procedures or applying upper- and lower bonds, as applicable

- Conditional entry checks to reject unsuitable data, like:

- Comments filed were entered for medical history but the value for past medical history has been marked as “no”.

- Mandatory fields were marked properly for example, date of informed consent.

- Correct selection of the following: second pass reconciliation (visibility of the data to second pass data entry operator).
3.5.6 Documentation

MDMF (Master Data Management File) was created as soon as the protocol and CRF was received. Table 3.9 gives the list of documents which were created in MDMF.

Table 3.9: MDMF Documents

<table>
<thead>
<tr>
<th>TABLE OF CONTENTS FOR MDMF</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0 Data Management Plan</td>
<td>10.0 Data Clarifications Forms (DCFs)</td>
</tr>
<tr>
<td>2.0 Protocol Amendments</td>
<td>11.0 Non-CRF Data</td>
</tr>
<tr>
<td>3.0 Case Report Forms</td>
<td>12.0 Data Coding Report</td>
</tr>
<tr>
<td>4.0 CRF &amp; Protocol Alignment (CAPA)</td>
<td>13.0 Differences between 1st &amp; 2nd Pass Entry</td>
</tr>
<tr>
<td>5.0 Randomization</td>
<td>14.0 Data Tracking</td>
</tr>
<tr>
<td>6.0 Computer System Details</td>
<td>15.0 SAE Reconciliation</td>
</tr>
<tr>
<td>7.0 CDMS Software Documentation</td>
<td>16.0 Data Handling Report</td>
</tr>
<tr>
<td>8.0 Laboratory &amp; other data</td>
<td>17.0 Official Correspondences</td>
</tr>
<tr>
<td>9.0 Quality control by CDM</td>
<td>18.0 Quality Assurance</td>
</tr>
<tr>
<td>19.0 Miscellaneous</td>
<td></td>
</tr>
</tbody>
</table>

The data was documented ensuring accountability, facilitating coordination, evidence and back-up; if required to reconstruct the trial [94], [95]. It facilitated to reduce operational ambiguity, as training material, served as knowledge enhancing tool, and for the purpose of predictive analysis to improve the existing processes [96]. A MDMF was created to file all the approved documents; records that described the methods, conduct and results of the study, and corrective measures undertaken [8]. This facilitated audit(s) and inspection(s).

Each vaccine project had separate MDMF with documents filed in reverse chronological order. The older versions were stamped on front page ‘Superseded Copies’ in red. Although Indian GCP does not compels and provides the list of essential documents for CDM,
including vaccine trials, to be maintained before, during and after study, but, MDMF is important for audit and a documentary evidence that the task was performed, following the industry prevalent best practices.

### 3.5.6.1 Brief description of the content of MDMF

- Data Management Plan: Study specific DMP.
- Amendments: Final approved protocol and all subsequent amendments received from clinical trial team
- Case Report Forms: Final (blank): SF (Screening Form) and CRF for the trials.
- CRF Protocol Alignment Document: CAPA for the study
- Randomization: Randomization list procured from the biostatistician.
- Computer System Details: System dependability documentation; hardware and software used to create, maintain, modify, retrieve, archive, or transmit trial data [11].

The major components were:

- Name of the software/package system used with its version
- Description of the hardware platform
- Location of the database on the system
- CDMS Software Documentation
  - Annotated CRF
  - Collation list (study specific Glib specifications, in excel sheet)
  - Code lists (DVG mapping to the field i.e. field to code list association)
  - CRF Layout (Screen shorts of approved screen layout)
  - Database Specifications: information about the metadata details generated from OC, including the details about the few major reports:
    - Data Collection Module Questions for Study
    - Data Collection Module for Study
    - DCI Detail Listing
  - Edit Specifications (Details about the validation programming for discrepancy management: Procedure Details for the study).
  - Data Entry Guidelines (DEG): Live document that was updated throughout the study with all the versions of DEG.
- Data Import/Export details: How the data was transferred to the biostatistician.
- Database Closure: The procedure adopted for database lock to maintain its integrity.
- List of critical and non-critical parameters for the study as procured from medical monitor.
- Laboratory data (normal reference ranges & updates, if any).
- Quality control by CDM: QC reports and the path for soft copy of the system generated logs after testing validation programs.
- Data Clarifications Forms (DFCs): The list of system generated DCF Tracking Log and the unresolved DCFs.
- Non-CRF Data: Data which were not captured in the CRF
- Data Coding Report: The information about the study specific data coding guidelines with MedDRA. For example- the version of dictionary used for the study, the datasets which were coded, approval details of the codes, DCFs sent to the site based on coding etc.
- Report on differences between First and Second Pass Entry: Data entry differences which were observed between first and second pass.
- Data Tracking: All the CRFs and DCFs were tracked through logs so that they were not missed or lost.
- SAE Reconciliation: Details about the reconciliation done for SAE between CDM and PVG unit for the studies.
- Data Handling Report (DHR): Prepared after the lock and had the information if there were any deviations (minor) to Indian GCP, study DMP or SOP. DHR was later submitted to biostatistician.
- Official Correspondences: File Note and Other communications (Fax etc.)
- Miscellaneous section: other important information for the study like- staff transition etc.
3.5.6.2 Standard Operating Procedures: Database Setup and Processing

In-house SOPs were developed mentioning the following:

- Procedure for database design, CRF annotation, data entry, data reconciliation and verification.
- Procedure for validation programming, view for data extract, MDMF creation
- The responsibilities of the person designated for the database design, data entry, data reconciliation and verification.
- Procedure for QC activities and subsequent updation
- Procedure for version control, for example software
3.5.7 Data cleaning, MedDRA Coding, SAE Reconciliation

3.5.7.1 Data cleaning: Data Validation and Query Management

Figure 3.21 shows diagrammatic representation of Data Cleaning process.

**Data Validation**
Validation checks were triggered on completion of data entry. Validation of data was done incrementally by scheduling batch job or by triggering a single procedure, as applicable. Data was thoroughly reviewed for discrepant data points; flagged by database when the validation checks were run against the raw data or manual review of the data. The process was performed with an objective to scan for correctness and accuracy of data as per the requirements mentioned in the study protocol.

**Query Management**
Data cleaning was initiated after data validated programmatically and reviewed through appropriate QC procedures. It included incomplete, missing, inaccurate, irrelevant or out of range aspects of the data and then substituting, altering, or deleting this dirty data. Any discrepancy identified, DCFs were sent to the site for seeking resolution. DCF flow was maintained through DCF tracking logs to keep the track of turn-around-time (TAT) by the study site.
3.5.7.1.1  Data cleaning Process

- Batch Validation was run against raw data
- Discrepancies were reviewed; comparison against the hard copy of the filled CRF
- DCFs were generated and sent for resolution to the site
- Resolutions received on the DCFs were updated in the database.
- Batch validation re-run for secondary discrepancies, if any and the above process was repeated till all the discrepancies were resolved.

3.5.7.1.2  Specific points considered for Data Cleaning Process

- Up-to-date programming of edit checks and Active status not missed out.
- Discrepancies sent for resolution should be sent along with the ‘DCF Tracking Log’.
- Duplicate discrepancies were not sent for seeking resolution to the site.
- In case the resolution received was not satisfactory, DCFs were re-raised and resent.
- DCF text was kept simple

Figure 3.22 shows flow chart for data cleaning process
3.5.7.1.3 Standard Operating Procedures: Data cleaning Process

In-house SOPs were developed mentioning the following:

- Creation of Discrepancies (DCF)
- Updation of resolution to the discrepancy in database
- Preparing DCF Tracking log
- Database Design and Data Entry
- QC activities and subsequent updation
3.5.7.2 Medical Dictionary for Regulatory Activities (MedDRA) Data Coding

Diagrammatic representation of MedDRA coding process is shown in Figure 3.23.

![MedDRA Coding Process Diagram](image)

**Figure 3.23**: Steps for MedDRA Coding Process

Medical coding was performed to categorize the medical terms appropriately so that they are consistent/homogeneous, and therefore easily analyzed and reviewed [95]. It was done using MedDRA dictionary for verbatim terms mentioned in AE and SAE pages [96].

3.5.7.2.1 MedDRA Coding Process

- Verbatim terms (text as written on the CRF) were transformed to a standard terminology
- AE and SAE data was coded.
Verbatim terms were coded to corresponding LLT (Lower Level Term), PT (Preferred Term) and the primary SOC (System Organ Class).

Coding was done manually.

All coded terms were subjected to QC for checking the correctness

Documentation in the standard format.

Coded terms were sent for medical review.

Signing off, once coding was finalized.

3.5.7.2.2 Specific points considered: MedDRA Coding Process

- Coding was done as per the procedure’s described by the dictionary manual
- In case there is ambiguous information or text in the verbatim term, DCFs were sent to the site for clarification
- Coding was done at pre-specified regular intervals
- Coding was complete before database lock
- Coding was done as per the current version of the dictionary

3.5.7.2.3 Standard Operating Procedures: MedDRA Coding Process

In-house SOPs were developed mentioning the following:

- Data coding process
- Review of coding
- Procedure for updation of resolution of the discrepancy with respect to ambiguous verbatim term.
- Template for preparation of document for coding
Figure 3.24 shows flow chart for verbatim terms coding process using medical dictionary MedDRA.

Figure 3.24: MedDRA Coding Process Flow [15]
3.5.7.3 Serious Adverse Events (SAE) Reconciliation

Figure 3.25 give the diagrammatic representation of SAE reconciliation process.

![SAE RECONCILIATION PROCESS](image)

**Figure 3.25:** Steps for SAE Reconciliation Process

SAEs were reconciled in PVG and CDM safety databases so that the information in both the system is same and are exact match. All discrepant data points were identified, if any. If needed, DCFs were created to seek resolution from the site. This process was carried out manually as CDMS and PVG database was not integrated seamlessly. This process was documented and continued till database-lock, at predefined intervals.

### 3.5.7.3.1 SAE Reconciliation Process

- Manual process was adopted for SAE reconciliation
- Request was sent to PVG with the detail information about the dataset in CDMS.
- Datasets at both places (CDMS & PVG) were reconciled for differences in the:
  - Total number of SAE
  - Information about individual SAE
- In case of any discrepancy action was taken by sending the DCFs to the site and subsequent updation, till information in both the databases was found to be identical.
- In case there was no SAE were found, like in Myfive™ and NUCOVAC® studies, but all email communications were still documented.

3.5.7.3.2 **Specific Points considered: SAE Reconciliation Process**

- SAE reconciliation was completed prior to the database lock
- Reconciliation was done at pre-decided regular interval, visit wise.
- MedDRA Coding was checked in both PVG and CDM databases for consistency.
- For consistent coding of verbatim terms, if required, they were split

**Figure 3.25 shows flow chart of SAE Reconciliation Process**

![SAE Reconciliation Process Flow](image_url)

3.5.7.3.3 **Standard Operating Procedures: SAE Reconciliation Process**

In-house SOPs were developed mentioning the following:

- Steps for SAE Reconciliation in both CDM and PVG units
- Procedure for updation of resolution to the discrepancy
- QC activities and subsequent updation steps
- Template for preparation of document for coding SAE Reconciliation
3.5.8 Database Lock

*Database lock*

Database lock, locking or freezing of the study database, was done to maintain data integrity by preventing unintended or unauthorized changes to the database. This was done after data entry was complete, resolving all the discrepancies, completing data coding, reconciling the SAE information with that of PVG and CDM databases, and quality checks etc. as per database lock check list (Figure 3.26).

This was done by adopting the following procedures:

- **Soft Lock:** The process for database ‘Lock’ in Oracle Clinical, locking prevented alterations to formerly collected data. Once data was locked, it was not possible to unlock the same.

- **Hard Lock:** The process for database ‘Freeze’, freezing averted modifications or additions to data at the study, investigator, and patient levels. The study-level freeze excluded the study from any validation processing. Of note, it is possible to unfreeze CRFs by users with the required privileges, and the new data can be entered. This was not the case in this study. The rights for ‘privileged update’ were not granted to CDM.

3.5.8.1 Database Lock Process:

- It was checked that all the activities were completed as per database lock check list; CRFs entered, queries resolved, data coding done, SAE reconciled with PVG, and QC process completed. (Figure 3.26).

- All project team members were intimated of database lock including but not limited to- Clinical Operations team, QA, IT etc.

- Email was sent to Head – CR for seeking approval for the soft lock.

- Once approval was received the email was forwarded to DBA (data base administrator), IT Department.

- The process was documented in the standard format.

- Request for revoking the user rights was also sent to DBA

- Study hard lock was performed after a period of 4 to 9 months of soft lock.

- To seek hard lock, approval was requested from all the team members.
- The scanned copy of ‘Authorization Note’ was forwarded to IT Department for study database freeze.

- Database hard lock process was later documented.

- A provision to unfreeze (equates unlock) was introduced in the study only if:
  - There is any major updation in the efficacy data
  - There is significantly important updation in the safety data
  - Any other significant information, which may directly affect the data quality or standards or study outcome
  - Database unfreeze must be carefully reviewed of the needed updation and approval by all stake holders.
  - ‘Database Change Request Form’ shall be forwarded to DBA, with signatures and the reason for change clearly mentioned.
  - DBA shall unlock the study for needed updation.
  - Thereafter the process of lock, mentioned above, shall be repeated to maintain database integrity.

- No updation was required in the database after lock in this study.

3.5.8.2 Points considered in Database Lock Process

- Before proceeding for the lock, all the pre-requisite activities were completed as per the ‘database lock checklist’.
- Study access revoked to prevent unauthorized access.
- All team members were informed and the necessary approvals obtained before lock.
- Documentation was up-to-date and correct.

3.5.8.3 Standard Operating Procedures: Database Lock Process

In-house SOPs were developed mentioning the following:
- Points to be checked in database lock check list.
- Procedure for seeking approval for all team members
- Database lock- Soft and Hard Lock procedure
- Database unlocking (technically equates to un-freezing in OC) steps
- Documentation procedure
**Figure 3.26: Database Lock Check List**

<table>
<thead>
<tr>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. All expected CRFs have been received by Data Management Department.</td>
</tr>
<tr>
<td>2. All study data has been entered &amp; discrepancies have been resolved.</td>
</tr>
<tr>
<td>3. All reference ranges related to Lab data applied in the database have been checked.</td>
</tr>
<tr>
<td>4. Data validation has been completed and approved (refer to the Data Validation Guidelines document).</td>
</tr>
<tr>
<td>5. Coding has been completed (refer to the Coding Guidelines document).</td>
</tr>
<tr>
<td>6. All non-coded text and unsolicited comments have been reviewed, and any necessary actions have been taken.</td>
</tr>
<tr>
<td>7. Data extract in Excel, have been reviewed randomly.</td>
</tr>
<tr>
<td>8. QC of critical data items has been completed (refer to DMP, for sample size for QC, for study specific requirements).</td>
</tr>
<tr>
<td>9. For non-critical data points, QC of representative Sample of data (Sample Size = \sqrt{N+1}, N=Number of patients enrolled in study or more) has been performed.</td>
</tr>
<tr>
<td>10. An acceptable error rate has been reached, if calculated in QC Reports.</td>
</tr>
<tr>
<td>11. Data Entry report has been generated from QC for the purpose of QC (Data entry=&gt;Data Entry Report).</td>
</tr>
<tr>
<td>12. Audit Trail Verification has been done. (OC =&gt; Conduct =&gt; Conduct Reports=&gt;Data Validation=&gt;Audit History).</td>
</tr>
<tr>
<td>13. Data Handling Report is ready for approval.</td>
</tr>
<tr>
<td>14. Approval from Head-CR has been received and all respective team members have been informed.</td>
</tr>
<tr>
<td>15. SAE reconciliation done with PVG department.</td>
</tr>
</tbody>
</table>
3.6 CDM Processes: NUCOVAC® vaccine trial

The implemented CDM model for Myfive\textsuperscript{TM} vaccine, QA validated and regulatory compliant (audited) procedures, was replicated in NUCOVAC\textsuperscript{®} vaccine trial (Figure 3.27) to achieve CDM procedural standardization (in-house common data standard, Panacea Biotec Ltd.). The CDM procedure adopted for NUCOVAC\textsuperscript{®} vaccine trial was audited by QA Department to ensure that there are 1) no deviations in SOPs, 2) no major/critical audit findings in the adopted steps, and 3) no need to ‘Unlock’ database, thus thereby procedural standardization [80].

Figure 3.27: Summary: CDM Model Standardization for vaccine trial, Panacea Biotec Ltd