International Journal of PharmTech Research
CODEN (USA): IJPRIF, ISSN: 0974-4304
Vol.8, No.1, pp 146-153, 2015

Insights in paper Case Report Form Design from Vaccine Trials in an Indian Pharmaceutical Company: Clinical Data Management Prospective

Bajpai N.[1]*, Chatterjee A.[2], Dang S.[3], Sharma S.K.[4]

[1]HCL Technologies, Noida, India
[2]Biological E Ltd., Hyderabad, India
[3]Department of Biotechnology, Jaypee Institute of Information Technology, Sector-62, Noida, India
[4]Sanjeev Kumar Sharma, Head – Department of Biotechnology, Jaypee Institute of Information Technology, Noida, India

Abstract: CRF serves as an important tool for the data collection and are tailored as per the study requirements in conformance with the protocol.[1] The design of the CRF is imperative not only from the viewpoint of clinical research operations and the study site staff, but the perspective of data management team should not be ignored. A well designed CRF with correct layouts, can aid in study conduct by enhancing the performance, as the data collected is expected to be credible and accurate with minimum errors. Format(s) for CRF designs, may be readily available but void exists in terms of applicable tricks and tactics needed for the task from the standpoint of database designer, data mangers, data entry operators and data coordinators.[2] It is noteworthy to acknowledge cross-functional domain knowledge in developing study CRF. This article narrates the possible methodologies which if adopted during 'study-setup phase' at the time of paper CRF designing, will result in optimal quality of data thus leading to overall reduction in the time to database lock, by augmenting productivity thereby benefiting the organization to shrink its budget for the study. The information incorporated here is derived from our experiences of data management of vaccine studies, conducted by a prominent Indian bio-pharmaceutical company under Indian regulatory environment.
Key-words: Case Report Form, (CRF), Clinical Data Management (CDM), Indian Good Clinical Practices (GCP), CDM Metrics.

Introduction

Indian Good Clinical Practice (GCP) defines the CRF as: 'A document designed in consonance with the Protocol, to record data and other information on each trial subject. The Case Record Form should be in such a form and format that allows accurate input, presentation, verification, audit and inspection of the recorded data. A CRF may be in printed or electronic format'.[1]

CRF may be considered as a lifeline of the trials, because selection of wrong layouts may have negative implications leading into repercussions which may result in corruption of data and noncompliance with regulatory requirements. Designing of CRF should be done to suffice the requirements of cross functional teams, i.e. by taking inputs from data procurers, data provider, data owners, data managers, data analysts and

data custodians,[3]at the planning stage of study setup phase. Standardization of CRF pages with correct choice of underline formats is vital for the successful study conduct.

It is the responsibility of the sponsor to design the CRF prior to the investigator's meeting and site initiation visit for the study. With the start of patient recruitment, the study CRF would be needed to capture protocol related information about the subjects. Timely completion of designing of CRF which accurately represents precise data points is very essential. Correct data fields incorporated into standard pages will facilitate accurate analysis by biostatisticians. If standardized formats are used then it will ease in meta-analysis and subsequent rollout of extension and/or sister studies.

Depending upon the standard operating procedures (SOPs) of the organization, the CRF may or may not be designed by CDM team. However, the important piece in the task is to incorporate the inputs and comments into the document based on overall the perspectives all the stake holders.

The layout of a CRF should be determined by data types and the structure of the dataset in the database while considering a cluster level of fields in one page as well as a plan of data review by scheduled visits.[4]

This article describes the key inputs given by CDM team during CRF design of vaccine clinical trials (Myfive™ and Pneumococcal vaccine NUCOVAC®).[5] However, due to confidentiality constraints, examples which are quoted in the paper are the replica created for the purpose of demonstration and are not the exact copy the CRF handled during the study conduct.

This report is an important milestone and a suggested step towards CDM process standardization.

Methodology

Work flow: Overview of CRF Design

Following gives the brief about the workflow adopted for development of the CRF of vaccine studies:

- As far as possible global standard pages were used in the CRF. These pages were designed in as logs based on visits. However, study specific pages, where applicable, was developed based on the available literature or else as required. This task was performed by clinical trial operation and medical writing teams.
- Draft CRF was provided to CDM Team by clinical research operations department.
- CRF was reviewed by CDM unit, with special focus to see its alignment with respect to the requirements of the protocol and also to check the feasibility of database designing, data cleaning (discrepancy management) and data entry.
- The comments from this review process were documented in CRF and Protocol alignment document (CAPA). The document was thoroughly reviewed by QC personnel before its finalization.
- Final CAPA was sent to the clinical research operations team for necessary updation in the draft CRF.
- CRF was finalized after incorporating the inputs by the all the stake holders.

Standard Operating Procedures for CRF Design

Standard operating procedure (SOP) is a document consisting of step by step information on how to execute a particular task.[6]

Following in-house SOPs were developed to perform the task related with the designing of the CRF:

- Steps for the design and development of CRF
- Task ownership metrics for the process
- The standard format and pages for the CRF
- Process of Quality control (QC), Quality assurance (QA), CRF Approval
- Procedures for amendment and version control of CRF
- CRF completion instruction/guidelines
- Process of training

Points considered while providing inputs on CAPA, by CDM Team for Vaccine studies

A well-designed CRF facilitates data collection and entry, and directly benefit other facets of data management and statistical analysis.[4]To achieve this CDM team has given following inputs for the CRF design in order to address requirements, not limited to the following.

- **Collection of data as per the study protocol**

CRF must be designed to collect only the information as defined in the protocol. It must not be designed to collect less or more than that required. Flow of information, to be collected in the CRF, must be as per the protocol.

If there is version change of the CRF, due to protocol amendments, then policy must be in place to handle the following business scenarios-

- Version change of CRF: before the start of study at the site
- Version change of CRF: for an ongoing study
- Use of different versions of the CRFs at different sites for the same study.

For the three scenario mentioned above, the data cleaning process may vary, this will have direct impact on database design, validation/edit checks programming and discrepancy management.

In the following example (Table 1a & 1b), it can be seen that the exclusion criteria number nine was removed from the study as protocol was amended, however the CRF designer has missed out deleting the information from the demographic page of the study. This has led to capture of extra information than that needed in the protocol.

Table 1a: Exclusion Criteria Number 9 was removed from the study protocol

Exclusioncriteria		
Please tick(√)as applicable		
S.No.	Details of Exclusion Criteria	Response
9.	~~Wastheinfant bornbefore37weeksof gestation?~~	Yes [__] No[__]

Table 1b: CRF designer has missed out deleting the corresponding field from demographic data, resulting into the collection of extra information at point number 3.

Demographic Data	
1-Hospital/OPD No. (ifapplicable)	[__I__I__I__I__I__I__I__I__]
2-Gender	Male[__] Female [__]
3-Gestational ageat timeofbirth(inWeeks), should be >=37 week	[__I__]

- **Collection of data to facilitate the end user*s***

CRF must be designed keeping in mind the needs of end users, who will fill the form at the site and who will perform data entry.

- **Sequence of questions/fields**

Data flow must be correctly arranged so that the sequence of questions asked is accurate and the placement of text is also in the right order. This could be seen as per the example (Table 2a & 2b) mentioned below in 'Reminder' page of the CRF, where information about the diary card has been asked from the subject at the site.

Table 2a: In the below table the order of questions asked in the CRF is wrongly arrange.

Reminders(Visit1)	
1-Parent/LAR instructed to come with the subject and Diary card on the next scheduled visit?	Yes [__] No[__]
2-Parent/LAR instructed to fill the diary card appropriately?	Yes [__] No[__]
3-Subject Diary Card issued to Parent/LAR?	Yes [__] No[__]

Table 2b: In the below table the order of questions asked in the CRF is correctly arrange. Point Number 3 should have been asked before asking question number 1 and 2.

Reminders(Visit1)	
Subject Diary Card issued to Parent/LAR?	Yes [__] No[__]
Parent/LAR instructed to fill the diary card appropriately?	Yes [__] No[__]
Parent/LAR instructed to come with the subject and Diary card on the next scheduled visit?	Yes [__] No[__]

- **Language of questions/fields**

The language used in the CRF must be simple and easy to understand by the end used. It should be without any spelling mistakes. Following table (3a) reflects the diary card reconciliation data, solicited local reaction post vaccination, in the CRF.

Here, the parameters are not very simple to understand, because usually the person, who fills the data in the CRF at the site, is not the investigator himself, it may be the site's clinical coordinator or any other authorized designee.

Table 3a: Use of tough language

S.No.	**Parameter**	**Category**	
1.	Erythema	[__] None	[__] Mild
		[__] Moderate	[__] Severe
2.	Induration	[__] None	[__] Mild
		[__] Moderate	[__] Severe

In the following table (3b), it is much easier to comprehend the information about the above parameters by the end user:

Table 3b: Use of simple language

S.No.	**Parameter**	**Category**	
1.	Redness/ Erythema	[__] None	[__] Mild
		[__] Moderate	[__] Severe
2.	Hardness/Induration	[__] None	[__] Mild
		[__] Moderate	[__] Severe

- **Grouping of questions/fields**

All fields which details with the common topic should be collected together. For example- all question related with demographic data must be place together, while the question related with vital signs & physical examination should be grouped together.

Table 4: Wrong grouping of questions can be seen in the following:

Questions	Response	Dataset
Gender	Male[__] Female [__]	Demographic Information
HeartRate	[__I__ I__] per minute	Vital Signs
DateofBirth (DD-MM-YY)	[__I__]-[__I__]-[__I__]	Demographic Information
General Appearance	Normal[__] Abnormal [__]	Physical Examination
BirthWeight of the infant (Kg)	[__I__]-[__I__]	Demographic Information
Skin	Normal[__] Abnormal [__]	Physical Examination

The last column gives the mapping for the responses. It can be seen that the responses are mapped to different datasets.

- **Other precautions**

Each page of the CRF should have limited number of questions so that the page must have sufficient space to capture the response and must not look over stuffed with too much information, which might be difficulty to read by data entry operator.

As far possible leading questions should be avoided, rather the questions should be clear and concise.

- **Collection of data so as to avoid redundancy**

Collection of duplicate information must be avoided as it adds to the redundancy. For example date of birth and age must not be collected together, as age can be derived easily from date of birth. Similarly, if BMI has to be collected than only height and weight of the subject should be captured in the CRF. This will maintain homogeneity and avoid calculation mistakes.

In the following examplethe 'Site of vaccine administration' was collected twice i.e. in 'Vaccine Administration Record' (Table-5a) & 'Diary Card'(Table-5b) pages. Information must not be collected twice, just to confirm its correctness, as chances are that the site personal can fill different replies for the same questions, leading to data queries

Table 5a: 'Site of vaccine administration' marked red, collected in Vaccine Administration Record of CRF

Vaccine Administration Record : Dose 7			
Was study vaccine administered to the subject enrolled?	No □	Yes □	
Kindly confirm following information for the Vaccine administered:			
Date (DD-MM-YY)	[__I__]-[__I__]-[__I__]	Time(HH:MM)	[__I__]:[__I__]
□ **ARM 1:**TEST	***Vaccine 1*** given on anterolateral aspect of thigh Right thigh □ Left thigh □ **Lot No:** _____		
□ **ARM 2:** COMPARATOR	***Vaccine 1*** given on anterolateral aspect of thigh Right thigh □ Left thigh □ **Lot No:** _____		

Table 5b: 'Site of vaccine administration' marked red, collected again in Post Vaccination Record of the CRF

<table>
<tr><th colspan="6">POST VACCINATION RECORD (DOSE 7)</th></tr>
<tr><th colspan="6">SOLICITED 'LOCAL REACTIONS'(up to 27 minutes after/ post vaccination)</th></tr>
<tr><th>S. No.</th><th>Factors</th><th colspan="2">Group</th><th>Relationship##</th><th>Outcome*</th></tr>
<tr><td colspan="2">Anterolateral thigh</td><td colspan="2">Left □ Right □</td><td>□ □Vaccine 1□ □ □Vaccine 2</td><td></td></tr>
<tr><td rowspan="2">1.</td><td>Pain/Tenderness#</td><td>□ Severe
□ Mild</td><td>□ Moderate
□ None</td><td>[__]</td><td>[__]</td></tr>
<tr><td colspan="5">#Severe: Very painful;Moderate:painful; Mild: Insignificant; None: Nothing</td></tr>
<tr><td>2.</td><td>Swelling@</td><td>□ None
□ Mild</td><td>□ Moderate
□ Severe</td><td>[__]</td><td>[__]</td></tr>
<tr><td colspan="6">@<i>Severe:</i><i>>25mm;</i> <i>Moderate:</i> <i>10-25mm;</i> <i>Mild</i><i>: <10mm;</i><i>None</i><i>: Nothing</i></td></tr>
<tr><td colspan="6">##Causal relationship with the vaccine: Z. Unrelated; U. Unlikely; P. Possible; C. Certain;</td></tr>
<tr><td colspan="6">* Outcome: 1. Resolved 2. Unresolved/ongoing 3.Death 4. Unknown.</td></tr>
</table>

Duplicate information can serve as an important tool, when it comes to CRF header. CRFs are unusually sent visit wise from the site for data processing. If a page is left out in the site and sent later, to the sponsor, then a cross-check of patient ID with initials for the data on every page against enrollment information has proven invaluable in identifying mislabeled pages.[7]

- **Use of free text**

Refer to table 5a, for the filed 'Lot No.' a single liner space is provided which will result in the errors for the person transcribing the information in the CRF at the site and for data entry operator who has to key-in the same into the database.

Instead, for lot number, proper boxed should have been provided with appropriate length. This should have been done in a manner so that single box may be used to fill a letter or a number.

- **Change in the order of code list**

Refer to table 5b, the order of codes list associated with point number 1 and 2 has changed. This will result in major problems for the person who is filling the CRF at the site and will drastically increase the rate of data entry errors.

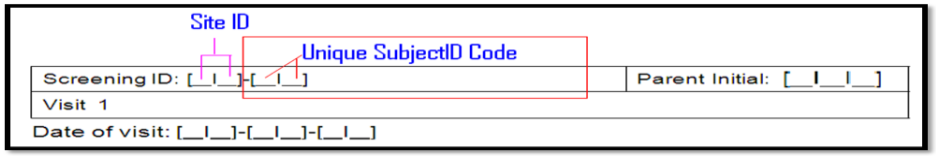
Where the responses are fixed/ expected, dropdown list should be associated with the filed for consistency.

- **Collection of data as per as per the necessities of database designer and applicable regulatory requirement**

It is very important for the group who has the responsibility to contribute for CRF designing, should suffice all the regulatory requirements. As per Indian GCP-the Sponsor should use an explicit Subject identification code that allows identification of all the data reported for each Subject.[1]Thus to capture this information sufficient space must be there in the CRF so that the uniqueness in subject identification code is maintained.

For instance if there are one thousand subjects to be enrolled in a study, then if space is provide as follows (Table 6) in the header, then for the subject with identification code 1000, the site personal will have problem in capturing this information.

Table 6: Screenshot of CRF, showing incorrectly designed header, with less space for writing down subject ID 1000:



Use of standard pages, as far as possible should be encouraged, especially for those which are common for almost all trials, to name few- demographic information, vital and physical examination, Adverse Events (AE) /Serious Adverse Event (SAE) information etc.

Except for very limited number of study specific pages, almost all other pages may be standardized based on the therapeutics group or kind of trial. For example standardized pages may be created for vaccine trials, oncology or diabetic trials etc.

These standard pages, if provided by the Indian regulatory authorities to the industry then it would be one of the most welcomed step and shall have more benefits than the ones mentioned below[11].

- Time will be saved for each stake holder involved in the review process of the CRF. The benefit of time saving will also get passed on to the database designers, validation programmers, followed by faster implementation at the site. This may contribute in decreasing the overall training time for all, as anyone who has worked in one segment can capitalize on his/her previous experience and finally having clean data with comparatively less efforts than before. It may mean decreased time for regulatory review and subsequent approvals.
- The chances of missing out vital information to be collected in the CRF, specific to one particular medical segment, will be decreased. People can have more detailed SOPs, probably with more elaborate checklists or procedures incorporated in it, there by the chances of repeating previous mistakes and committing new ones will be drastically reduced[11].
- Analysis of data, creation of data listings and summary tables would be much faster and comparatively easier. As each table of the CRF would be predefined and pre-written codes can be reuse.
- This will also ensure the use of standardized laboratory values with units. Batch data uploads would be much easier.

Discussions

An efficiently structured CRF with correct arrangement of data points and appropriate text will offer for flawless approach for the collection of primary efficacy and safety endpoints for the study.[8] It will provide the users a clear direction for the collection of required responses.[4]

However, giving too much importance to single perspective might proof problematic. For example, database entry operators would prefer to have tabulated CRF based on the visits, matching with flow of information given in the protocol. But the pages generated to capture adverse event (AE) information- are usually preferred to be designed as the 'log form'. This is because it is easy to compare the previous information related with ongoing AE during source data verification (SDV) or discrepancy management. For this particular example, it is must to evaluate the requirement of both i.e the person transcribing the information as well as the one who is monitoring (CRA). Thus equal weightage should be given to the viewpoints of multidisciplinary teams, including but not limited to the Medical Writers, Medical Monitors, Biostatistician, Monitors (CRA), Database Programmers/ Designers, Investigator/Site staff, Data Entry Operators, QA personnel and also from the people from regulatory department of the company.

Some of the other aspects of process of CRF design are:

- Adequate rules should be incorporated to capture dates and partial dates in the CRF completion guidelines.
- Use of consistent color and font size should be encouraged.
- Use Of different Software for formatting, designing of the layouts and their version control. Adequate Training should be imparted to all the team to help them perform the task in compliance with SOP, protocol, GCP and regulatory standards.

Scope of Improvement Identified in the process:

- For finalization of CRF, inputs were provided internally by all the interdisciplinary teams, including CDM, with documentary evidence. However, there should be an SOP in place to take the input from the sites personal, who are the end users of the CRF.
- Availability of standard pages of the CRF with minimum content defined, from regulatory authorities is needed.

Regardless of the time and effort spent conducting the trial, the correct data points (response to a CRF question/data is entered) must be collected; otherwise, a meaningful analysis may not be possible.[9]

Conclusion

An informative and structured CRF simplifies database design and data validation processes as well as manipulation of data during statistical analysis.[4]Extraneous data can adversely affect overall data quality by drawing the attention of site personnel away from the key variables.[10]

A systematically designed CRF may promise to check protocol adherence by the investigator. CRF should be designed with an aim to have analysis ready data from which precise and unbiased conclusions could be easily interpreted.

Disclosure

The opinions and interpretations stated in the article represent individual's viewpoint only, there being no conflict of interest.

References

1. Indian Good Clinical Practice : http://cdsco.nic.in/html/GCP.htm (Accesses on 27th Sep, 2013) http://unpan1.un.org/intradoc/groups/public/documents/apcity/unpan009867.pdf&http://www.cmchvellore.edu/static/research/Files/Good%20Clinical%20Practices%20For%20Clinical%20Research%20In%20India.pdf(Assessed on 23rd April, 2015).
2. Bajpai, N; Chatterjee, A; Dang, S; Sharma, S. K. "Clinical Data Management Patrons: Positions & skill requirements in the industry". The Clinical Research Plus, vol. 3, no 1, pp. 18-21, 2012.
3. Martin E., "Data Management Best Practices and Standards for Biodiversity Data. Applicable to Bird Monitoring Data. Compiled by. 1 and Grant Ballard". http://www.nabci-us.org/aboutnabci/bestdatamanagementpractices.pdf (Assessed: 13 April, 2015).
4. Moon K.K., "Techniques for Designing Case Report Forms in Clinical Trials Considerations for Efficient Data Management and Statistical Analysis".ScianNewsvol. 9, no., 2006. http://www.scian.com /media/ 1533/sciannews91b.pdf (Assessed on 23rd April, 2015).
5. Bajpai, N., Sharma, M., Chatterjee, A., Dang, S., Sharma, S. K. "Standardization of procedural implementation in Clinical Data Management", with reference to the trials: DTwP-HepB-Hib vaccine (MyfiveTM) vs. Pneumococcal vaccine (NUCOVAC®), Indian Journal of Scientific Research (IJSR), vol. 04, no. 02, pp. 179-191, 2013.
6. Satyendra, Standard Operating Procedures in Management, Oct 2014. http://ispatguru.com/standard-operating-procedures/ (Assessed on 23rd April, 2015)
7. Practical Guide to Clinical Data Management, by Susanne Prokscha, Second edition. by Taylor & Francis Group, 2007.
8. Belgium: Society for Clinical Data Management; 2009. Good Clinical Data Management Practices. http://www.swaggerinfosolutions.pro/sridevi/New%20Folder/ICRI/GCDMP_complete_OCT2009.pdf (Assessed on 14thJune , 2013)
9. Bellary S., Krishnankutty B.,and LathaM. S. "Basics of case report form designing in clinical research".PerspectClin Res. ,vol 5, no.(4), 2014.
10. Califf R.M., Karnash S.L., Woodlief L.H.,"Developing systems for cost-effective auditing of clinical trials". Controlled Clinical Trials. vol. 18,1997.
11. Good Clinical Data Management Practices, Published 2013, SCDM
12. Case Report Form Design Issues by Phyllis Rathman, Biotechnical Services, Inc. www.biotechnical services.
13. Clinical Data Interchange Standards Consortium (CDISC). (Assessed on 23rd April, 2015)
14. Nahm M., Shepherd J., Buzenberg A., et al. "Design and implementation of an institutional case report form library". Clin Trials. vol.8, 2011.

*****