PhUSE US Connect 2019

Paper DH04

Expectations and Collaboration – Recipe for Successful CDISC Deliverables!

Santosh Lekkala, Vita Data Sciences, Waltham, USA
Bhavin Busa, Vita Data Sciences, Waltham, USA

ABSTRACT

When working in a CRO/Vendor model, we may have projects with different Sponsors where the data collection standards and requirements may differ based on the study-specific and therapeutic area needs. Also, when work is outsourced to a CRO/Vendor model, the interpretation and the ways of handling data can differ between Sponsor and the vendor, causing re-work at a later stage which results in delayed deliverables, decreased productivity and increased cost. This paper focuses on the process we follow at our organization before and during the Specification development/programming of the SDTM datasets to achieve consistent, reliable and timely deliverables, thus benefiting both the Sponsor and the vendor. This paper explains in detail the pre-requisite steps that can be followed before writing the specifications, quality checks and data handling during the development process of the datasets and informed notes during the delivery of datasets to the Sponsor to ensure the expected outcome.

INTRODUCTION

The process of generating the SDTM datasets starts with understanding the requirements, annotating the CRF, developing the specifications, programming and validation of the datasets and checking for compliance against the current validation rules. The process of achieving consistency and reliability of datasets is a 2-way coordination effort which requires a clear set of expectations from the sponsor, meticulous attention to detail and timely response from the vendor to make sure that expectations are met. When a clear set of expectations for various components is not done beforehand, it results in re-work at a later stage where implementation becomes difficult and this increases the overall cost.

Though each vendor may have their own SDTM development process, we will discuss the steps that can be implemented during the development phase, thereby ensuring that we have the required information right before the start. This also helps in making oneself aware of the various components which can be implemented beforehand.

THE PRE-REQUISITE STEPS THAT CAN BE FOLLOWED BEFORE WRITING THE SPECIFICATION:

The pre-processing steps are those that are to be taken into consideration before proceeding with writing the specifications for the study. The following questions answered beforehand would be beneficial and this would involve a few confirmations from the sponsor. The lead programmer must make sure they have the following information.

Pre-requisite to SDTM Specifications:

Topic	Required Checks
Protocol (All versions)	All the protocol versions are to be requested from the sponsor before proceeding to write the specifications and the programmer needs to make sure if the latest version is being used. This would help in creating the TI Domain where changes if any, to the inclusion and exclusion criteria are required to be specified. This would also be beneficial as TI and IE datasets are to have the same TEST and TESTCD as per the guidelines.
SDTM Model and IG Version	The general practice is to use the latest version available, but in some cases, if the sponsor determines to use a previous version to make it consistent with the other studies that are already present, the Sponsor may request to use a previous version.

Topic	Required Checks
CT, MedDRA and WHO Drug Dictionary versions	Though the MedDRA and WHO Drug versions would be updated before submission to the FDA, It would be better to have the version being used beforehand
eCRF With Raw Annotations	Electronic CRF with RAW annotations would be helpful to navigate and search for required variable names and assist during specification writing.
Source Data Analysis	It would be better to have a list of all the raw datasets that are being used in the study along with the associated SDTM dataset to which they are being mapped to be listed and confirmed from the sponsor before writing the specifications. This would help avoid re-work which when done at a later stage, might affect the quality and timeline.
External Data Transfer Specifications	Review and let the sponsor know of any changes that might be required such as the TEST and TESTCD, addition of new variables, format of the transfer, frequency of the transfer, Original and Standard units for the data received, Conversion factor if required. This review helps to make sure all the required variables are present and checks for any data inconsistencies that could affect the specification writing/programming.
Frequency of Transfer	We need to determine the frequency of transfer to plan and execute the delivery process so that delays can be minimized. These include the frequency of transfer for - RAW data, external data, if any and Transfer of SDTM datasets to the sponsor. Appropriate care must be taken to have enough buffer time between receiving RAW transfer and transferring the datasets to the sponsor
P21/ Define Version	The version of P21 used to validate the define and datasets along with the version of the define required by the sponsor needs to be confirmed.
Study Tracker	A study tracker is to be prepared with all the tasks such as the datasets planned, production and QC programmer names, target and actual dates of completion and point of contact for any questions.
Randomization	Randomization if any, in the study needs to be taken care of. A dummy randomization file needs to be requested that should include the various fields that would be available when the final randomization file is received, so that it can be incorporated during programming and would not require an update at a later stage.
Study Milestones	Determining the study milestones is an important part which provides an idea on various deliverable components to both the Sponsor and the Vendor. It determines various important steps and deliverables during the study and helps in tracking and meeting timelines
CRF Annotation	It is recommended to annotate the CRF without any bookmarks for sponsor review before proceeding with specification writing.

QUALITY CHECKS AND DATA HANDLING DURING THE DEVELOPMENT PROCESS

The quality checks and data handling are done during the specification development and programming. Once the specifications are ready, it is recommended to have the sponsor review the specifications. A checklist is recommended to be included along with the specifications that lists the major points of focus which requires critical attention, so that errors are minimized at a later stage. The following points would not only help the reviewer get an overall picture but also makes the review process faster and easier. The following checklist will be provided to the sponsor:

Domain	Topic	Values Expected/Derivation used	Agree/Disagree
TA/DM	ARM and ARMCD	'XX01', 'XX02'	
TA/DM	Actual vs Planned Arm	'XX01', 'XX02'	
TA/TE/S	Elements	Screening, Treatment, Follow-up	
TA/TE	Epoch	SCREENING, TREATMENT, FOLLOW-UP	

Domain	Topic	Values Expected/Derivation used	Agree/Disagree
TV/SV	Planned Visits	Visit -1, Visit-2, etc.,	
DM	RFSTDTC vs RFXSTDTC	First day and time of dosing in EX	
DM	RFENDTC vs RFXENDTC	Last drug dose date in EX	
DM	RFPENDTC	Last available date for a subject in the entire study	
SV	Data sets to be considered for VISITS	Considerations for creation of SV dataset. If just only the master dataset if available is to be used to create the dataset or if all the datasets are to be used. Also check from the sponsor if any external data such as the LB or PC data dates are to be included in SV dataset.	
SV	Handling of “UNSCHEDULED VISITS”	Naming of UNSCHEDULED VISITS such as 14.01, 14.02	
TA/TE/SE	Determining the Epochs	Ex: SCREENING, TREATMENT, FOLLOW-UP	
TS	TSPARAM/TSVAL	Missing information or information unavailable through open sources for various TSPARAM’s are to be confirmed from sponsor.	
All Findings	BLFL derivation method	Last non-missing value collected before study drug administration.	
AE	Derivation of AETRTEM	Set to “Y” for all AE’s occurring after treatment start date (or) existing AE’s which have worsened.	
AE	Control terminology used for AEREL, AEACN, AEENRF, etc.	Control terminology used for mapping the collected values are to be confirmed by the sponsor.	
DS	Values for DSTERM	Screening, Randomization, Withdrew consent, etc.	
All Findings Domain	TEST and TESTCD as per CT	TEST and TESTCD’s being used for all findings domains are to be confirmed.	
All Domains as required	Determination of EPOCH derivation.	If SE is being used or any other derivation method is being used to determine EPOCH. If so, is the time part is being considered.	
All Domains as required	Mapping of External data	Example: SDTM.LB dataset. Merging process to be followed and how missing data is being represented (EX: LBALL for missing TESTS by each category)	
XX	Custom datasets	Custom datasets are to be confirmed by the sponsor for the approach and appropriate type into which they are being mapped to like events or interventions.	

During the programming and validation part the following steps need to be taken care of:

- Perform metadata difference for each of the RAW Data transfer received to make sure no new fields are added or highlight any changes to the fields already mapped.
- Pass the generated datasets through P21 validator and address all issues as required.
- A manual check of all the datasets is required before sending the delivery to the sponsor (such as running customized edit checks) for those that cannot be determined through P21 and requires a diligent Human Eye. (Ex: P21 doses not through an Error (or) Warning when RFSTDTC is present but ACTARM is provided as “Not Assigned”.)
- It is recommended to generate and pass the define along with the datasets when validating through P21 to make any changes as required to either define, datasets or both.

INFORMED NOTES DURING THE DELIVERY OF DATASETS TO THE SPONSOR:

The informed notes provided during the delivery of Datasets to the sponsor is an important part which ensures that the basic high-level information regarding the delivery has been provided to the sponsor and this may include information such as:

- The total number of subjects present in the transfer (if possible a breakdown by number of subjects per ARM, such as 10 Screen failures, 25 treated to drug, 25 to placebo, etc.)
- The total number of datasets in the transfer
- Missing datasets, if any.
- Any information regarding specific subjects that the vendor thinks would be helpful for review by the sponsor.
- Data Issues recorded in a log.
- The Data cut for each of the RAW data source such as RAW data cut date, External data cutoff date, etc.,
- Clarification Log for any information or questions that require sponsor response.
- Information regarding any blank datasets that have not been included in the transfer but are part of the delivery due to lack of corresponding information.
- Any errors, warnings from P21 that standout from the normal ones.

CONCLUSION

Achieving consistent and reliable datasets starts right from determining what the expectations are and is an ongoing process until the end of the study. The different components mentioned above aids in making sure that the interpretation and handling of data is done correctly to factor in all the pieces such as Specification development, programming, compliance and ease of understanding are met. This in turn, would benefit both the sponsor and vendor by making sure that all the deliverables are met on time without any delay or re-work. By making sure that the steps above are followed as part of the SDTM development process, it would be easier to achieve reliable datasets right from the start.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Author Name: Santosh Lekkala
Company: Vita Data Sciences
Address: 281 Winter Street, Suite 100
City / Postcode: Waltham, MA 02451
Phone: 713-851-7559
Email: slekkala@vitadatasciences.com, lsantosh280@gmail.com
Web: www.vitadatasciences.com

Author Name: Bhavin Busa
Company: Vita Data Sciences
Address: 281 Winter Street, Suite 100
City / Postcode: Waltham, MA 02451
Work Phone: 781-373-8455
Email: bbusa@vitadatasciences.com
Web: www.vitadatasciences.com