Efficiencies Realized in Building and Utilizing ADaM from SDTM

Shirish Nalavade,eClinical Solutions, a Division of Eliassen Group, Mansfield, MA
Parag Shiralkar,eClinical Solutions, a Division of Eliassen Group, New London, CT

Abstract

Although Study Data Tabulation Model (SDTM) serves the purpose of tabulation and submission, and Analysis Data Model (ADaM) serves the purpose of analysis and review of data, understanding interrelation of these models and utilizing efficiencies by their simultaneous implementation can be extremely beneficial to sponsor. Many initiatives have been taken by industry to implement ADaM when there is already in-built structure of SDTM based data models. SDTM/ADaM pilot project, which was conducted by collaborative efforts from industry, CDISC, and FDA, provided lot of guidelines and lessons learnt. In this paper, authors would like to share their experiences of simultaneous implementation of SDTM and ADaM for tabulation, analysis, and reporting of three clinical trials associated with same compound.

Introduction

Implementation of Study Data Tabulation Model (SDTM) provides lot of benefits through standardization of meta-data. We have been privileged to work on multiple clinical trials of the same drug and these trials were progressing simultaneously. Considering the benefits of Analysis Data Model, we decided to implement the ADaM for the analysis and reporting of the data. Although the SDTM specifications for each trial were finalized, the ADaM implementation methodology was tailored to the progress of trials. Considering this, the biostatistics and programming group undertook implementation and development of these data models for one trial at a time and tried to utilize efficiencies related to development of these models for subsequent trials. For trial specific reporting, authors found that having the data in SDTM model facilitated lot of traceability of data transformations required to develop ADaM. Input data in SDTM also facilitated lot of standardization of analysis dataset structures, and simplified the programming algorithms greatly. This also led to efficient utilization of programming resources. Benefits of having data in SDTM and analysis methodology developed and implemented by ADaM had lot of benefits in terms of re-usability and cross trial efficiencies when structure of these two models were utilized for subsequent trials and for pooled data analysis. Such efficiencies in implementation and reduction of programming and review time is measured for understanding the return on investment of implementation of ADaM when sponsor has SDTM dataset structure.

Methodology (Creation of SDTM and ADaM Datasets)

A brief description of our implementation methodology is provided below that will further help to understand the benefits of utilizing these standards.
Our implementation method considered a linear method to create analyses datasets, which is illustrated in Figure1. For more information on other methodologies, please refer Susan J. Kenny's paper mentioned in references.

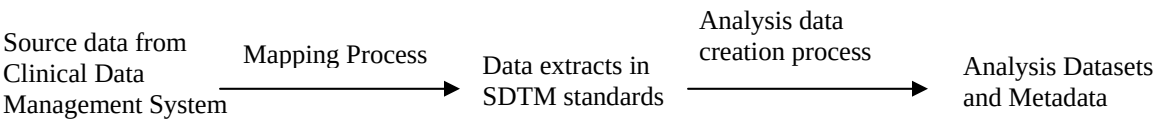

Figure1. Implementation method

Raw data from clinical database management system (CDMS) was processed and mapped as per SDTM standards by developing mapping macros in SAS® for each domain using SDTM IG v 3.1.2. Although the main purpose of these macros was to transform and modify the metadata as per SDTM standards, few variables were derived during this process, for ex: vitals sign baseline flags (VSBLFL), unique subject ID (USUBJID) which was concatenation of studyid–siteid and subjid. These variables were later carried forward as it is to analysis datasets.
Analysis datasets were developed from SDTM domain dataset and were used to create statistical summaries of efficacy and safety data. Data from different domains dataset was combined to create optimal number of datasets and new variables were derived during creation of analysis datasets. Derived analysis variables may be for statistical calculation of an important outcome measure, such as change from baseline or may represent the last observation for a subject while under therapy. These analysis decisions were detailed in study protocol and or SAP. The analysis datasets contained all variables required for creation of tables and figure as mentioned in SAP.

Efficiencies/Benefits realized

As linear approach was utilized for creation of SDTM and ADaM datasets and both these datasets were created according to standards, we were able to develop standard programs to create datasets and reports and reuse these programs across multiple trials resulting in efficient use of our resources

Programming Efficiencies: Programming efficiencies were achieved by standardization and re-usability of mapping, analysis and reporting processes. Data from CDMS was converted according to SDTM standards, which meant the data structure was consistent across trials to create analysis datasets. Because of this we were able to standardize most of our datasets creation process and to most part standard safety reporting and validation programs and reuse these programs across all studies.

Benefits were realized during all stages of the studies because of implementation of standards. During mapping process, data mapping team created Logical Data Map (LDM's) spreadsheets that provided mapping information and logical flow of data from CDMS to SDTM datasets for each domain. LDM's gave programmers with detailed information on transformation of metadata for each variable as per SDTM IG v3.1.2. Utility macros were developed to perform some repetitive tasks. For Ex. converting dates to ISO8601 format, and assigning metadata attributes to variables. These macros were then reused to convert data from all domains across all studies which assured that data is consistently converted as per SDTM standards across trials and in turn greatly reducing dataset programming time.

The structure of data coming from CDMS greatly facilitated our standards implementation process. The data was in vertical structure (i.e. one record per subject per visit and test) and the

data from CDMS had already derived some variables required for analysis. This was helpful because the programming team did not have to further process the data to derive some additional variables.
For Ex: In VS data, BMI was derived from HEIGHT and WEIGHT and data was arranged in vertical format as shown in figure2 below.

Vital Signs Test Name	Vital Signs Test Short Name
BMI	BMI
Diastolic Blood Pressure	DIABP
Heart Rate	HR
Height	HEIGHT
Respiratory Rate	RESPRAT
Systolic Blood Pressure	SYSBP
Temperature	TEMP
Weight	WEIGHT

Figure2. Vital Signs data in vertical format

After the input datasets were in SDTM standard, the next step was ADaM conversion to create ADaM datasets. ADaM dataset specifications were created based on the information from SAP. All transformations required to create analysis datasets were maintained in an excel sheet which was created by biostatistician as per individual trial information in SAP. As the studies were for the same compound most of the variable transformation were similar except for some of the efficacy endpoints where a little tweaking was required from trial to trial. Excel sheet contained metadata information such as variable name, label as per ADaM IG 2.0, and also the transformation required for each variable. In some cases, actual SAS® program was included in transformation column to facilitate programmers to derive the variable.

Because of ADaM standards, a lot of conversion processes common to all studies were standardized by macros and these macros were utilized across all studies. Macros were developed to read input data, to derive numeric date format from SDTM ISO8601 character format, to derive duration between two dates, computation of flags such as treatment emergent AE, prior AE and populate metadata information such as variable label. Utilizing these macros across studies reduced a lot of analysis dataset programming time and greatly expedited our process of creating the TLG's and subsequently allowed for early submission of reports to the sponsor.

Resource Utilization: A major area where efficiencies were greatly recognized was in resource utilization. Because of implementing standards and processes as mentioned above, a lot of cross trial reusability of programs was achieved which resulted in increased efficiency and reduced programming time and resources. As dedicated teams worked on each part of clinical trial, the standards implemented by each team for one study were carried forward to other studies. Therefore, not only an appreciable amount of time was reduced across trials but also resulted in following benefits

- Utilized less number of personnel that translated into substantial cost benefit to the sponsor
- As each programmer was responsible for specific task in trials, programmers developed a thorough understanding of the raw data and process which resulted in lowering the

learning curve for subsequent studies and time to complete programming of these studies.

- This understanding of study data further resulted in timely response to address any data queries or ad-hoc requests from sponsor
- Quick turn around of reports, well in advance of the specified timelines
- All these factors ultimately helped us in gaining trust and promoted a pleasant working relationship with the sponsor.

Review and Traceability: Since ADaM datasets were derived from SDTM datasets, the mapping sheet created during ADaM conversion process provided detailed information of the data flow and a level of metadata traceability between SDTM and ADaM variables. This mapping sheet had information about the variables that were carried forward from SDTM datasets, computation algorithm to derive variables required for analysis along with information about SDTM dataset name and variable names to be used for derivation. Using this document, we were able to easily trace back to any data queries from the sponsor and provide with a timely response to their queries. As a result, this document assisted to significantly reduce data and table review time of the sponsor and eventually the turnaround time of final reports. This was particularly critical during interim analysis when the sponsor was able to make an early decision on continuation of the study.

Scalability: Having data in ADaM standard across studies provide us with the capability to pool data from multiple studies. Standardization of data facilitates for trouble free data pooling as the metadata and the content of each variable are consistent across the studies. If the sponsor is planning to pool the data, we anticipate that merging/combining data for these studies should be straight forward and the sponsor could expect the same benefits as describe above.

Conclusion

Adapting SDTM and ADaM standards and implementation methodology depicted above was very successful and beneficial to all stakeholders involved in this project. Not only it benefited our programming group by standardizing and maintaining reusable programs to create datasets and TLG's across studies, it also benefited our sponsor in terms of time and financial aspects of the project. Implementing standards saved a significant amount time and resources by leveraging the process developed for one study and applying it to other studies.

Benefits of having data in SDTM were realized across all ADaM specific processes such as standardization of data mapping process, analysis dataset creation, creation and validation of standard safety reports. These benefits were also realized in terms of re-usability and cross trial efficiencies when structures of these two models were utilized for subsequent trials and for pooled data analysis. Moreover, implementing these standards provided our programming team an opportunity to gain a thorough understanding of SDTM/ADaM models and utilize these standards to efficiently and successfully close out clinical studies.

References:

1. Strategies for Implementing SDTM and ADaM Standards by Susan J. Kenny, Maximum Likelihood Solutions, Inc, Octagon Research Solutions, Inc., Chapel Hill, NC, Michael A. Litzsinger, SCHWARZ BIOSCIENCES, Inc., Research Triangle Park, NC

Your comments and questions are valued and encouraged. Please feel free to contact the author at:

Shirish Nalavade
eClinical Solutions, A Division of Eliassen Group
603 West St
Mansfield, MA 02048
Work Phone: 508-594-6337
E-mail: snalavade@eclinicalsol.com or at shirish.nalavade@gmail.com

Parag Shiralkar
eClinical Solutions, a Division of Eliassen Group
e-mail: pshiralkar@eclinicalsol.com or at pcshiralkar@gmail.com