

Creation of Submission and Analysis Data Sets by following the SDTM and ADAM requirements

Yvane Boudraa, sanofi-aventis pharma, Antony, France

ABSTRACT

In Biostatistics & Programming (B&P) department, programmers and statisticians must create SAS datasets to perform the analysis and the reporting of clinical data. In case of submission, these datasets called Analysis Data Sets (ADS) will be provided to the health authorities. In addition, the health authorities require other datasets that contain all the collected information on subjects. These datasets are called Submission Data Standards (SDS) that will be mainly used for the medical review.

To facilitate the review, the Food and Drug Administration (FDA) has endorsed (July 21, 2004) a standard submission data format, called the Study Data Tabulation Model (SDTM) developed by the Clinical Data Interchange Standards Consortium (CDISC). In terms of analysis data standards, CDISC Analysis Data Model (ADaM) team has provided metadata models and examples for analysis datasets used to generate the statistical results for a regulatory submission.

The standardization of the SDS & ADS creation process involves a validated tool by following a computerized methodology.

INTRODUCTION

This paper presents the different assumptions in term of standards, databases and computerized methodology.

STANDARD DEFINITIONS

Standard definitions of ADS/SDS are internal to sanofi-aventis, based on SDTM v3.1.1 Implementation Guide, SDTM v1.1 and ADaM v2.0, and adapted according to sanofi-aventis data flow and to sanofi-aventis standard analysis (including reporting).

The Standard definitions of the datasets contain the structure of it

- SDS definition follow the SDTM documents where the following points are fixed:
 - o Datasets type and datasets names.
 - o Variables type (all in character), names, label.
 - o Variables required, expected or permissible notion.
 - o Variables values which must follow some format ISO, or other code list
- ADS definition follow the ADaM general consideration:
 - o Datasets names: ADxxxxxx.
 - o Maintain SDTM variable attributes if the identical variable also exists in an SDS dataset.
 - o Variables required.
- Standard Sanofi-aventis team defined the following points:
 - o Variables length
 - o ADS data and SDS data will be put together to build ADS/SDS dataset, according to the sanofi-aventis data flow.

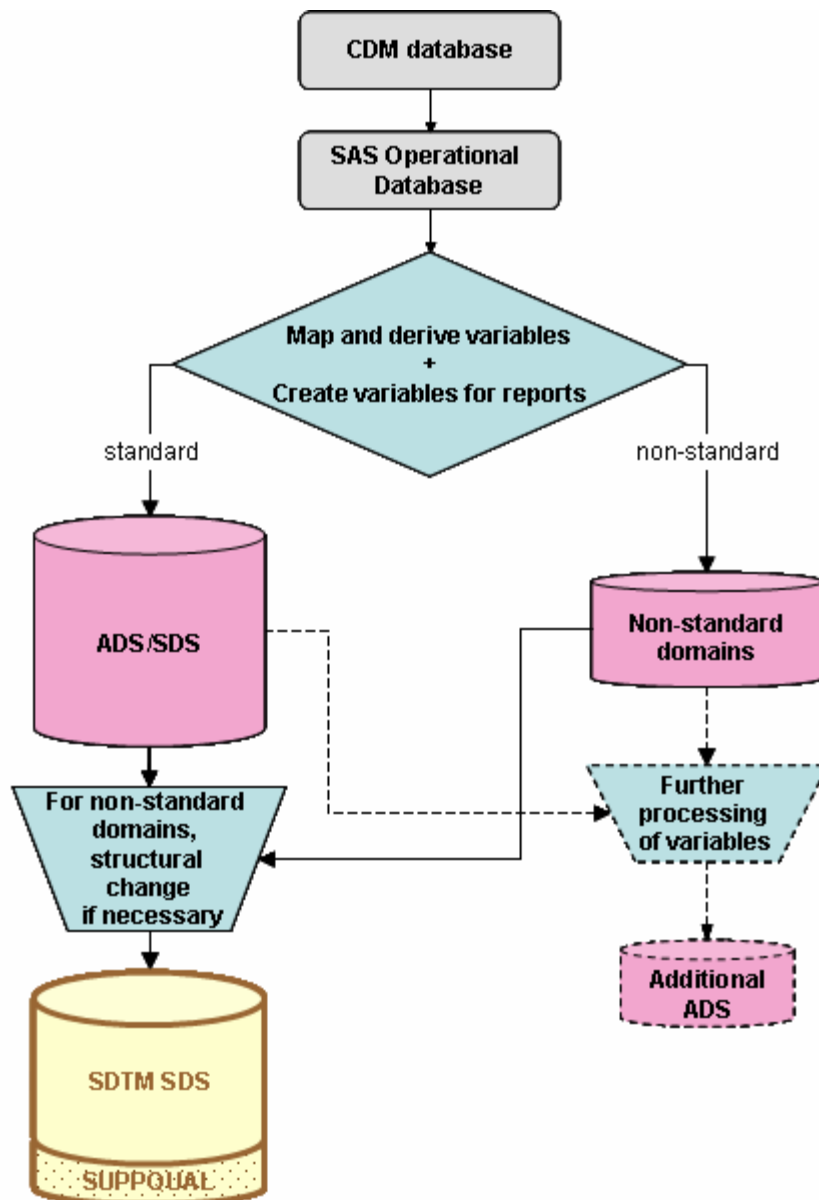
The Standard definitions of the datasets contain also the contents

- o By applying the standard rules of derivations for derived variables.
- o By keeping as much as possible the values of variables existing in the Clinical Data Management database.

STANDARD DATA FLOW

The creation of ADS/SDS follows a sanofi-aventis data flow process:

- Based on the CDM original datasets, SDS and ADS information are developed in the same datasets at the 1st step.
- SDS and Supplemental qualifiers datasets are directly taken from these mixed datasets in a second step.
- Only the standard side could be included in a standard application in term of derivation.



Note: The color pink stands for ADS and are used for reporting.
Both the ADS and the SDTM SDS/SUPPQUAL data are submitted.
The dotted line means when needed.

STANDARD RULES OF DERIVATIONS

Variables defined as standard in the ADS/SDS can come directly from the CRF, then the CDM database, or can be derived.

The derivations have been discussed within the standard team and the statistician and programmers to put in place some rules. These rules must be followed as much as possible and must take into account the study design and the study phase.

In addition, according to the reporting, the derivation must follow some display of values.

STANDARD DATABASES

Two Clinical Data Management databases are available: Clintrial 4 and Oracle Clinical that must be taken into account in the specification of the tool, mainly in the derivation macro development.

Values available in these CDM database would not be modified to avoid some interpretation, except for the missing codes and some coding where the interpretation is not possible e.g: sex=Male in the database becomes sex=M in the SDS/ADS.

COMPUTERIZED METHODOLOGY

The computerized methodology explains the way to conduct a project, including a complete documentation. This is separated in several phases: initiation, requirements, specification, building, testing, deployment and finally the production of the deliverables.

The methodology includes also the interaction between all the actors of the project:

- Biostatistics and Programming department being the sponsor of the project and representing the user community
- System department including in the Biostatistics and Programming department concerning the SAS development and the project lead
- Informatics department concerning the architecture
- Quality department to ensure the quality compliance in term of data, data privacy, informatics process

In addition, external members are consulted according to the project needs, as the Clinical Data Management concerning the original databases, the Standard team including in the Biostatistics and Programming department.

CONCLUSION

To put in application FDA /CDISC requirements involves to have a coordination between several actors coming from different departments (Biostatistics and Programming department: statistician and programmer, standard and system, CDM, IS, Quality).

The development of a standard tool is all the more complex since several sources must be followed and taken into account, and in the same time, a methodology must be followed.

To permit the success of this type of application, it is needed to have a development and a support team to follow the standard update and to satisfy the user needs as quicker as possible by following the computerized system methodology.

REFERENCES

SDTM v3.1.1 Implementation Guide, SDTM v1.1 and ADaM v2.0.

ACKNOWLEDGMENTS

N/A

RECOMMENDED READING

N/A

CONTACT INFORMATION

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

Yvane Boudraa
Sanofi-aventis
20 avenue Raymond Aron
Antony 92165
Work Phone: (33) 1 55 71 63 95
Email: yvane.boudraa@sanofi-aventis.com

Claude Guyot
Sanofi-aventis
1, avenue Pierre Brossolette
Chilly-Mazarin 91385 Cedex
Work Phone: (33) 1 69 79 72 72
Email: claud.guyot@sanofi-aventis.com

Brand and product names are trademarks of their respective companies.