PhUSE 2011

Paper PP15

Extension of the Six Steps Approach: How to go from an SDTM findings domain to an ADaM-compliant analysis dataset

Qian Wang, MSD, Brussels, Belgium
Carl Herremans, MSD, Brussels, Belgium

ABSTRACT

The pharmaceutical industry is experiencing a learning curve of building CDISC Analysis Data Model (ADaM)-compliant datasets from SDTM source data. In 2010 the authors presented six basic steps to transform an SDTM finding dataset into an ADaM-compliant Basic Data Structure (BDS) analysis dataset. This paper will demonstrate an extension of this systematic approach with examples to show how easily steps can be added to and/or removed from the existing framework to support analyses.

INTRODUCTION

Since the release of the Analysis Data Model (ADaM) and its implementation guide, more and more people are becoming familiar with the standard and are beginning to appreciate the benefits it brings. In recent years, a lot of interests are focused on how the standard can be applied to support statistical analyses. In early 2011, the CDISC team released the ADaM Basic Data Structure for Time-to-Event Analyses Document for public review. The guidance on data structure to support analysis of incidence data is also being developed.

At PHUSE 2010, the authors presented 6 steps to convert an SDTM finding data structure into an ADaM-compliant Basic Data Structure (BDS) [Wang & Herremans 2010]. This paper extends our proposed systematic approach and illustrates how steps can be easily inserted or removed to support analysis of interest. Detailed examples are provided for analysis of:

- Predefined limits of change
- Time to event

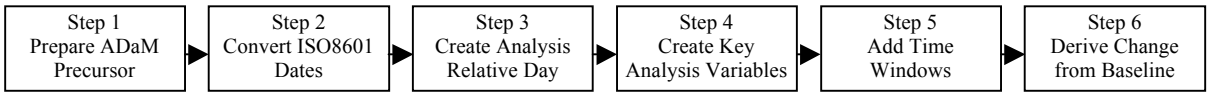
ADaM BDS DATA STRUCTURE

The ADaM Implementation Guide describes two ADaM standard data structures: the subject level analysis dataset (ADSL) and the Basic Data Structure (BDS).

A BDS dataset contains one or more records per subject, per analysis parameter and per analysis timepoint (depending on the analysis). It describes the data being analyzed and also includes variables to support the analysis (e.g. covariates) as well as information to facilitate traceability. Although a BDS dataset also has a long and skinny structure similar to a SDTM finding domain, it may be derived from all classes of SDTM domains, other ADaM datasets or combination of those.

SIX STEPS TO CONSTRUCT AN ADaM BDS DATASET

In this section, the six generic steps [Wang and Herremans 2010] are summarized:



STEP 1: PREPARING FOR ADaM PRECURSOR BY ADDING SUPPQUAL TO SDTM

The corresponding supplemental variables stored in the SUPPQUAL domain are added to the parent SDTM domain.

STEP 2: CONVERTING ISO8601 DATES

SDTM contains date and time values in the ISO8601 format (YYYY-MM-DDThh:mm:ss) as character variables. This step converts the ISO8601 formatted dates/datetimes to a numerical date format as required by the ADaM specifications.

STEP 3: CREATING ANALYSIS RELATIVE DAY

This steps creates the timing variable ADY expected in the BDS dataset. ADY typically describes the number of days from Analysis Date to a reference date.

STEP 4: CREATING KEY ANALYSIS PARAMETER VARIABLES PARAMCD/PARAM AND AVAL

This step defines the analysis parameters and their values as illustrated in the example below:

SDTM LB

LBTESTCD	LBTEST	LBSTRESN	LBSTRESU
NA	Sodium	139	mmol/L
K	Potassium	3.5	mmol/L

ADaM ADLB

PARAMCD	PARAM	AVAL
NA	Sodium (mmol/L)	139
K	Potassium (mmol/L)	3.5

STEP 5: ADDING TIME WINDOWS

The variables which describe the analysis time window are added in this step:

ADY	PARAMCD	AVAL	AVISIT	AVISITN	AWTARGET	AWRANGE	ABLFL	ANL02FL
-18	NA	141	Baseline	0	1	-70 to 7		
1	NA	140	Baseline	0	1	-70 to 7	Y	Y
14	NA	145	Treatment 1	1	45	8 to 90		
46	NA	149	Treatment 1	1	45	8 to 90		Y

STEP 6: DERIVING CHANGE FROM BASELINE

In step 6, baseline value (BASE) and change from baseline (CHG) are derived:

ADY	PARAMCD	AVAL	AVISIT	AVISITN	ABLFL	BASE	CHG
-18	NA	141	Baseline	0		140	1
1	NA	140	Baseline	0	Y	140	0
14	NA	145	Treatment 1	1		140	5
46	NA	149	Treatment 1	1		140	9

EXTENSION OF THE BASIC STEPWISE APPROACH

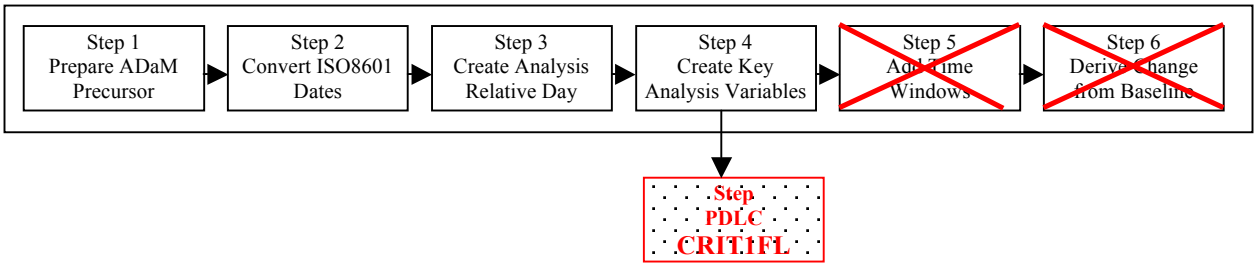
Like building blocks, the six steps can be re-assembled, replaced with project specific implementation, removed or added quite easily to support different analyses. The authors provide in this paper two such applications: analysis of predefined limits of change and analysis of time to event.

ANALYSIS OF PREDEFINED LIMITS OF CHANGE (PDLC)

Analysis of predefined limit of change categorizes information based on whether the endpoint exceeds certain specified limit of interest. The 'hit' criteria can range from simple rules like 'the post-baseline Serum Sodium result greater than 146 mmol/L' to more complex ones like 'sitting systolic blood pressure ≥ 180 mm Hg with a 20 mm Hg increase from baseline'. Although the criterion might vary depending on the data/population of interest and its clinical justification, it can be implemented as an independent step and inserted into the basic stepwise framework.

EXAMPLE 1: PREDEFINED LIMIT CRITERION IS MET WHEN THE POST-BASELINE SERUM SODIUM RESULT>146 MMOL/L

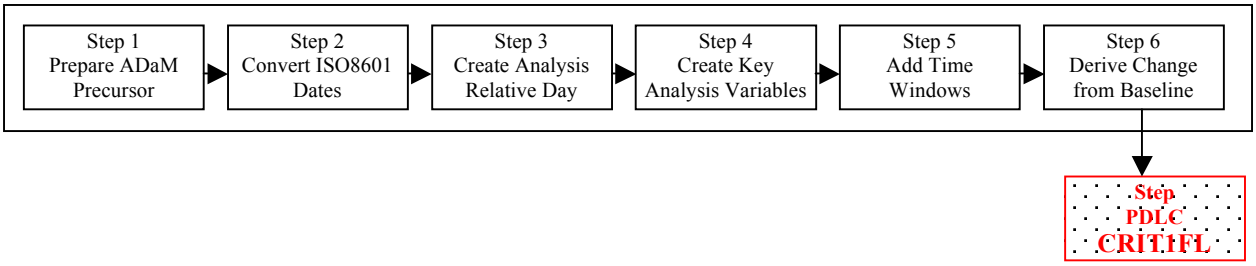
As this analysis does not require time window or the baseline/change from baseline information, Step 5 and 6 are no longer needed. The PDLC module can be added after step 4 and creates the pair of CRIT1 and CRIT1FL to flag the record which meets the criterion.



PARAMCD	ADY	PARAM	AVAL	CRIT1	CRIT1FL
NA	-18	Sodium (mmol/L)	141		
NA	1	Sodium (mmol/L)	140		
NA	14	Sodium (mmol/L)	145		
NA	46	Sodium (mmol/L)	149	Sodium > 146 mmol/L	Y

EXAMPLE 2: PREDEFINED LIMIT CRITERION IS MET WHEN THE POST-BASELINE SITTING SYSTOLIC BLOOD PRESSURE ≥ 180 MM HG WITH A 20 MM HG INCREASE FROM BASELINE

Although applied to VS instead of LB with a more complex criterion, this step can be added to the existing 6 steps in a similar way.



PARAMCD	ADY	AVISIT	ABLFL	AVAL	BASE	CHG	CRIT1	CRIT1FL
SYSBP	-18	Baseline		104	120	-6		
SYSBP	1	Baseline	Y	120	120	0		
SYSBP	14	Treatment 1		180	120	60	Result ≥ 180 mm Hg and change from baseline > 20 mm Hg	Y

TIME TO EVENT ANALYSIS

In ADaM Implementation Guide v1.0, CDISC introduced a list of analysis descriptor variables for the BDS datasets to support time to event analysis. The recent ADaM Basic Data Structure for Time-to-Event Analyses draft provides further guidance on time to event dataset ADTTE.

The table below illustrates such ADTTE BDS using an example of single event (DEATH) with a binary value for censoring variable (CNSR):

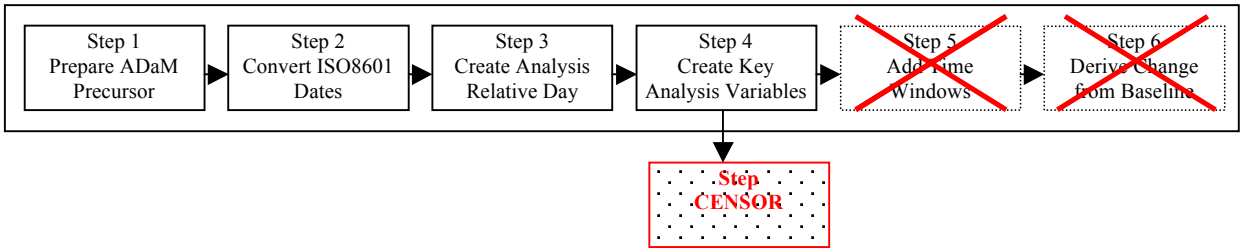
USUBJID	PARAMCD	PARAM	STARTDT	ADT	AVAL	CNSR	EVNTDESC
1001-0001	DEATH	Time to Death (days)	01JAN2007	15JAN2007	15	0	DEATH
1001-0002	DEATH	Time to Death (days)	01JAN2007	17JUN2007	168	1	COMPLETED THE STUDY
1001-0003	DEATH	Time to Death (days)	01JAN2007	30APR2007	120	1	LOST TO FOLLOW-UP
1001-0004	DEATH	Time to Death (days)	01JAN2007	17JUN2007	168	1	COMPLETED THE STUDY
1001-1005	DEATH	Time to Death (days)	01JAN2007	30JAN2007	30	0	DEATH
1001-1006	DEATH	Time to Death (days)	01JAN2007	04JAN2007	4	1	ADVERSE EVENT

Some important variables specific to the ADTTE BDS are listed in the table below:

Variable	Label	Description
ADT	Analysis Date	Date of the event/censoring
STARTDT	Time to Event Origin Date for Subject	The original date of risk for the time-to-event analysis. Typically this will be for example the date of the first drug intake.
AVAL	Analysis Value	ADT – STARTDT = relative date of the event
CNSR	Censor	Defines whether the event was censored
EVNTDESC	Event or Censoring Description	Describes the event of the interest, an event that warrants censoring, or an event represented by the censoring date

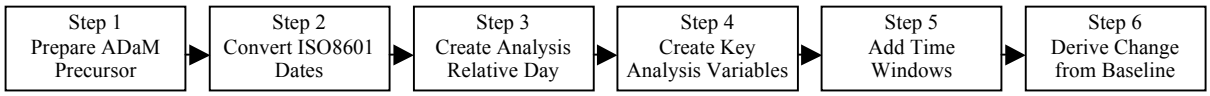
The existing step-wise approach can be adjusted in the following way by removing steps no longer needed (step 5 and step 6), replacing some steps with analysis specific implementation (step 2 and step 4) and by adding a new component for censoring:

Step 1: Merge of SDTM datasets (AE, DS…) with the SUPQUAL in order to build the ADaM precursors.
Step 2: Conversion of the ISO 8061 dates to valid SAS dates (e.g. for ADT and STARTDT).
Step 3: Creation of the relative day (ADY).
Step 4: Derivation of PARAMCD, PARAM and AVAL.
Step CENSOR: Creation of additional time to event variables CNSR and EVNTDESC.



CONCLUSION

In this paper, the authors reviewed the generic stepwise framework we presented in 2010 Phuse Conference to transform an SDTM finding dataset into an ADaM-compliant BDS analysis dataset:



The framework allows flexibility to delete/replace/add modules and its applications are illustrated using 3 examples to support analysis of:

- Predefined limits of change
- Time to event

References

Analysis Data Model (ADaM) Implementation Guide, Version 1.0. Final version, published by CDISC December 17, 2009. Available for download at http://www.cdisc.org.

Analysis Data Model (ADaM), Version 2.1. Final version, published by CDISC December 17, 2009. Available for download at http://www.cdisc.org.

The ADaM Basic Data Structure for Time-to-Event Analyses. Draft 2011

Qian Wang and Carl Herremans "How to go from an SDTM Finding Domain to an ADaM-Compliant Basic Data Structure Analysis Dataset: An Example" PHUSE 2010, available for download at http://www.lexjansen.com/phuse/2010/cd/cd05.pdf

ACKNOWLEDGMENTS

The authors would like to express appreciation to Frederic Coppin for his input during the preparation of this paper.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Please contact the authors at:

Qian Wang
Statistical Programming Analyst
MSD
Clos du Lynx 5
B-1200 Brussels [Belgium]
Tel: ++32 27766303
E-mail: qian_wang@merck.com
Web: http://www.merck.com

Carl Herremans
Senior Statistical Programming Analyst
MSD
Clos du Lynx 5
B-1200 Brussels [Belgium]
Tel: ++32 27766305
E-mail: carl_herremans@merck.com
Web: http://www.merck.com