

ADaM Compliance VS Easy-Reviewed in Design and Construct Efficacy Analysis Datasets

Huadan Li, MSD R&D (China) Co., Ltd., Beijing, China

ABSTRACT

CDISC ADaM provides a fundamental principle that enables statistical analysis of the data, while at the same time allowing the data reviewers to have a clear understanding of the data's lineage from collection to analysis to results. When develop the analysis datasets, compliant with the ADaM defined structure is first consideration. However, for the efficacy analysis data, if don't think about how to make the reviewer to understand the data's lineage from collection to results, that is, the analysis dataset is not easy-reviewed and must make more explanations in ADRG. This paper will demonstrate how to design and construct CDISC-ADaM compliance and easy-reviewed ADaM dataset.

Key word: CDISC ADaM Compliance, Easy-reviewed, Efficacy analysis dataset.

INTRODUCTION

The first CDISC Analysis Data Model (ADaM) Implementation Guide (ADaM IG v1.0) was published in 2009. It describes fundamental principles that apply to all analysis datasets, with the driving principle being that the design of ADaM datasets and associated metadata facilitate explicit communication of the content of, input to, and purpose of submitted ADaM datasets^[1]. By following the ADaM IG v1.0, the documents, The ADaM Basic Data Structure for Time-to-Event Analyses v1.0 in 2012 and ADaM Structure for Occurrence Data (OCCDS) v1.0 in 2016 are released to enhance the ADaM implementation in clinical trial statistical analysis of data. The Analysis Data Model supports efficient generation, replication, and review of analysis results^[1].

In March 2015, the CDISC ADaM Validation Checks v1.3 was written for ADaM IG v1.0 conformance checking. It provides more than 300 check rules and requests the data designer to create the compliant ADaM datasets.

It is a big challenge for the user creating compliant ADaM datasets, especially, the efficacy ADaM dataset, while allowing the data reviewers to have a clear understanding of the data's lineage from collection to analysis to results. In order to satisfy some specific analysis, the efficacy ADaM datasets have flexible data structure and content. Usually, the flexible data structure and contents are hard to meet the ADaM conformant and easy-review purpose simultaneously.

FUNDAMENTAL PRINCIPLES FOR ADaM DATASETS

Besides the ADaM conformant, a well-designed ADaM datasets should adhere to the fundamental principles described in ADaM IG section 2.1. The three principles below provide some basic rules how to create the easy-reviewed ADaM datasets.

- ADaM datasets and associated metadata must clearly and unambiguously communicate the content and source of the datasets supporting the statistical analyses performed in a clinical study^[1].
- ADaM datasets and associated metadata must provide traceability to show the source or derivation of a value or a variable (i.e., the data's lineage or relationship between a value and its predecessor(s)). The metadata must identify when and how analysis data have been derived or imputed^[1].
- ADaM datasets should have a structure and content that allow statistical analyses to be performed with minimal programming. Such datasets are described as "analysis-ready." ADaM datasets contain the data needed for the review and re-creation of specific statistical analyses. It is not necessary to collate data into "analysis-ready" datasets solely to support data listings or other non-analytical displays^[1].

According to the above rules, this paper will demonstrate how to design and construct ADaM compliance and easy-reviewed efficacy ADaM dataset from the four aspects.

- 1) Adopt CDISC ADaM Validation Checks to make first round ADaM compliance when design the ADaM specs
- 2) Organize the ADaM datasets associated metadata to make clearly data contents and data's lineage from SDTM to analysis results
- 3) Think more about the data structure and content, e.g., keep corresponding subject-level variables and create additional BDS variables and rows in efficacy ADaM datasets, to facilitate creation of statistical analysis and review of data
- 4) Arrange the variables order within efficacy ADaM datasets to facilitate the review

CDISC ADaM VALIDATION CHECKS

In March 2015, the CDISC ADaM Validation Checks v1.3 was written for ADaM IG v1.0 conformance checking. It provides more than 300 check rules and requests the data designer to create the compliant ADaM datasets. There are three types of ADaM structure checking rules: 1) ADSL:BDS; 2) BDS only and 3) BDS/ALL:SDTM. The data designer needs to follow the three types of rules to make first round eye-ball ADaM compliance checking when design the ADaM specs.

Table 1 Summary of ADaM validation check rules in the BDS related three types of rules

ADaM Structure Group	ADaM Variable Group	Frequency
ADSL:BDS	Flag Variables	4
ADSL:BDS	Timing Variables	10
ADSL:BDS	Treatment Variables	2
BDS	Analysis Parameter Variables	46
BDS	Analysis Visit Windowing Variables	5
BDS	Flag Variables	11
BDS	Population Indicator(s)	4
BDS	Time to Event Variables	9
BDS	Timing Variables	13
BDS	Toxicity and Range Variables	8
BDS	Treatment Variables	14
BDS/ALL:SDTM	Data Point Traceability Variables	5

The above table 1 gives the summary of ADaM validation check rules in the BDS related three types of rules. In total, there are 131 rules, the top three of frequency is Analysis Parameter Variables (46), Timing Variables (23) and Treatment Variables (16), especially, the user should pay attention for these rules and make sure the ADaM specs and programming to follow. Here, listing of individual rules is usually violated among the three top rules.

Table 2 Listing of individual rule are usually violated among the three top rules

ADaM Structure Group	ADaM Variable Group	Text from ADaM IG
ADSL:BDS	Timing Variables	APERIOD value must have corresponding TRTxxP/TRxxSDT/TRxxEDT variables
ADSL:BDS	Treatment Variables	TRTP/A must match at least one value in TRT01P/A-TRTxxP/A
BDS	Analysis Parameter Variables	If BASETYPE is populated for at least one record within a parameter then it must be populated for all records within that parameter

ADaM Structure Group	ADaM Variable Group	Text from ADaM IG
BDS	Analysis Parameter Variables	If both AVAL and AVALC are populated then there must be a one-to-one mapping
BDS	Analysis Parameter Variables	If both BASE and BASEC are populated then there must be a one-to-one mapping
BDS	Analysis Parameter Variables	PARAMCD values should follow SAS V5 variable naming conventions
BDS	Analysis Parameter Variables	PARAMN must be an integer
BDS	Analysis Parameter Variables	PARAMTYP has the same value for all records within a parameter
BDS	Treatment Variables	BDS must have TRTP variable
BDS	Treatment Variables	TRTAGy is required when TRTPGy is present and TRTA is present.

In general, in order to create an ADaM-compliant dataset, the user should get familiar with the CDISC ADaM validation checks and adopt the Checks to make first round ADaM compliance when design the ADaM specs.

ADaM DATASETS ASSOCIATED METADATA

Four types of ADaM metadata facilitates this communication by providing specification of details and links between the general description of the analysis, the analysis results, the data used in the analysis, and the SDTM domains

- 1) Analysis dataset metadata
- 2) Analysis variable metadata
- 3) Analysis parameter value-level metadata
- 4) Analysis results metadata

A well-designed ADaM datasets can clearly provide the traceability from source data to analysis results, also show the traceability back from analysis results to source data. The ADaM metadata, especially, analysis parameter value-level metadata and analysis results metadata, are useful tool for providing the mutual traceability from source data to analysis results, and vice-versa.

The analysis results metadata, the first part of ADaM data definition documents (define.xml), provide traceability from a result used in a statistical display to the data in the analysis datasets. They can be provided to assist the reviewer by identifying the critical analyses, providing links between results, documentation, and datasets, and documenting the analyses performed^[2].

The clear and unambiguous analysis results metadata are based on the ADaM datasets structure and contents with minimal programming. For example, the Figure 1 shows a One-Proc-Away statistical analysis in analysis results metadata from ADaM dataset. It just uses subset condition to get the statistical results and indicates the analysis dataset satisfies "Analysis-Ready" principle.

Therefore, in order to minimize the work in analysis results metadata, think more about the data structure and content, e.g., keep corresponding subject-level variables and create additional BDS variables and rows in efficacy ADaM datasets, to facilitate creation of statistical analysis and review of data

Figure 1 Analysis results metadata

Display	Table 14-3.01 Primary Endpoint Analysis: ADAS-Cog - Summary at Week 24 - LOCF (Efficacy Population)
Analysis Result	Dose response analysis for ADAS-Cog changes from baseline
Analysis Parameter(s)	PARAMCD = "ACTOT" (Adas-Cog(11) Subscore)
Analysis Variable(s)	CHG (Change from Baseline)
Analysis Reason	SPECIFIED IN SAP
Analysis Purpose	PRIMARY OUTCOME MEASURE
Data References (incl. Selection Criteria)	ADOSADAS [PARAMCD = "ACTOT" and AVISIT = "Week 24" and EFFF1 = "Y" and ANL01FL = "Y"]
Documentation	Linear model analysis of CHG for dose response; using randomized dose (0 for placebo; 54 for low dose; 81 for high dose) and site group in model. Used PROC GLM in SAS to produce p-value (from Type III SS for treatment dose). SAP Section 10.1.1
Programming Statements	[SAS version 9.2] <pre>proc glm data = ADQSADAS; where EFFF1='Y' and ANL01FL='Y' and AVISIT='Week 24' and PARAMCD="ACTOT"; class SITEGR1; model CHG = TRTPN SITEGR1; run;</pre>

SUBJECT-LEVEL VARIABLES AND CREATE ADDITIONAL BDS ROWS AND COLUMNS

A well-organized ADaM-BDS efficacy dataset is not only for creation of statistical analysis, but also for statistician or FDA reviewer's easy-review. The user creates an analysis-ready ADaM-BDS efficacy dataset, just keeping the core subject-level variables and BDS required variables, to satisfy various statistical analysis. However, sometimes, the statisticians and other reviewers, e.g., the FDA statistical reviewers, are hard to understand the data point and take much time to find out the logics and data point correlations across multiple ADaM datasets.

Besides by following the common convention of ADaM IG, how to select the subject-level variables and create additional BDS variables and rows to facilitate creation of statistical analysis and review is a big challenge for the data designer. Here is an example from Hepatitis C Virus (HCV) study to demonstrate the viewpoint how to create ADaM-BDS efficacy dataset to facilitate the review and creation of statistical analysis.

In HCV study, the primary efficacy analysis should be a comparison of the proportion of patients who achieve SVR₁₂ (sustained virologic response 12 weeks after stopping treatment) across trial treatment arms. The challenges are how to define the analysis visit windowing (e.g., 12 weeks after stopping treatment) and handling of missing data. In FDA HCV DAA Drugs for Treatment draft guidance, it has some paragraphs to describe the handling of missing data.

"For the primary analysis, sponsors can consider a patient as having achieved SVR₁₂ if the patient's week 12 follow-up HCV RNA measurement is missing and the patient achieved SVR₂₄. Sponsors should consider a patient not to have achieved SVR₁₂ if he or she discontinues from a trial before having an HCV RNA measurement at 12 weeks of follow-up or if the patient has missing HCV RNA values at the end of the scheduled 12- and 24-week follow-up periods."

"Analyses excluding patients with missing data or other post-treatment outcomes can be biased because patients who do not complete the trial may differ substantially in both measured and unmeasured ways from patients who remain in the trial."

"Appropriate sensitivity analyses should be performed to demonstrate that the primary analysis is robust to discontinuation and missing data. Sensitivity analyses can be performed using various methods for imputing missing post-treatment virologic results at 12 weeks of follow-up."

“We recommend that sponsors collect detailed data on confirmation of reasons for discontinuation (e.g., opportunity to enter another trial offering a promising new treatment, death or events leading to death, disease progression, adverse events, loss to follow-up, withdrawal of consent, noncompliance, pregnancy, protocol violations, not discontinued or not known to be discontinued but data were missing at the final visit). The underlying reasons for discontinuation should be interpreted. For example, the statistical analysis should include the number of patients who withdrew consent or were lost to follow-up, or who discontinued because of adverse events.”

From the FDA guidance description, we get some information for missing data handling and statistical analyses. It also requests the data designer to create the efficacy analysis dataset satisfying the requirements.

- 1) Identify the missing measurement in each time point, especially, follow-up 12 weeks;
- 2) The missing data type, intermittent missing or monotone missing and the discontinuation reasons for monotone missing;
- 3) Can perform appropriate sensitivity analyses to demonstrate that the primary analysis is robust to discontinuation and missing data

In the following HCV study efficacy example (Table 3.1-3.4), the HCV RNA assessment is scheduled to be performed at screening, Day 1, Day 7, Week 3, 4, 6, 8, 10,12, Follow-up week 4,12,24. The subject has missing assessment at week 3 and discontinued from study at Follow-up week 4. Therefore, there are three missing measurements at Week 3, Follow-up week 12 and 24; Week 3 data missing is intermittent missing and Follow-up week 12 and 24 are monotone missing due to the subject discontinued from study.

In order to identify the missing visit assessment, The ADaM dataset creates the “PHANTOM” rows (Row 6,13,14; See Table 3.2 and 3.3) for the missing visit assessment while indicating the missing data type, intermittent missing or monotone missing by using CRITy (See table 3.3 variables CRIT1, CRIT2 and CRIT3). Even if the CRITy variable is used for identifying a pre-specified criterion within a parameter, here, it is good solution for identifying the missing data type and it does not violate ADaM conformance checks. It’s very clear for the reviewer knowing the data points correlations.

At the same time, keep the study discontinuation status information (subject-level variables EOSSTT, EOSDT, EOSDY and DCSREAS; See Table 3.1) in ADaM dataset to facilitate the review, and support the data integrity.

Table 3.1 Subject-level variables show the discontinuation status

1	USUBJID	EOSSTT	EOSDT	EOSDY	DCSREAS	EOSTT	EOTDT	EOTDY	DCTREAS
2	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
3	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
4	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
5	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
6	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
7	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
8	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
9	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
10	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
11	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
12	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
13	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
14	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
15	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
16	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
17	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
18	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
19	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
20	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
21	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
22	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
23	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
24	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
25	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
26	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
27	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
28	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
29	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
30	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
31	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
32	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
33	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
34	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
35	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
36	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
37	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	
38	1111_0001	DISCONTINUED	7-Sep-2014	119	WITHDRAWAL BY SUBJECT	COMPLETED	3-Aug-2014	84	

Table 3.2 Analysis visit windowing

1	USUBJID	MFASRFL	PPROTFRFL	VISITNUM	VISIT	AVISIT	AVISITN	ADY	ADT	FUADY	AWRANGE	AWTARGET	AWTDIFF
2	1111_00001			10100	Screening	Screen	-1	-26	16-Apr-2014	-110	ADY < 0 day	-1	25
3	1111_00001			40001	Day 1	Day1	0	1	12-May-2014	-83	ADY = 1 day	1	0
4	1111_00001			40007	Day 7	TW1	1	8	19-May-2014	-76	2 days <= ADY <= 11 days	7	1
5	1111_00001			41020	Week 2	TW2	2	15	26-May-2014	-69	12 days <= ADY <= 18 days	14	1
6	1111_00001					TW3	3				19 days <= ADY <= 25 days	21	
7	1111_00001			41040	Week 4	TW4	4	28	8-Jun-2014	-56	26 days <= ADY <= 35 days (5WK)	28	0
8	1111_00001			41060	Week 6	TW6	6	42	22-Jun-2014	-42	36 days <= ADY <= 49 days (7WK)	42	0
9	1111_00001			41080	Week 8	TW8	8	56	6-Jul-2014	-28	50 days <= ADY <= 63 days (9WK)	56	0
10	1111_00001			41100	Week 10	TW10	10	70	20-Jul-2014	-14	64 days <= ADY <= 77 days (11WK)	70	0
11	1111_00001			41120	Week 12	TW12	12	84	3-Aug-2014	0	78 days <= ADY <= 98 days (14WK)	84	0
12	1111_00001			54004	Follow-up Week 4	FW4	104	119	7-Sep-2014	35	21 days (3WK) <= FUADY <= 69 days	28	0
13	1111_00001					FW12	112				70 days (10WK) <= FUADY <= 146 days	84	
14	1111_00001					FW24	124				147 days (21WK) <= FUADY	168	
15	1111_00001			10100	Screening	Screen	-1	-26	16-Apr-2014	-110	ADY < 0 day	-1	25
16	1111_00001			40001	Day 1	Day1	0	1	12-May-2014	-83	ADY = 1 day	1	0
17	1111_00001			40007	Day 7	TW1	1	8	19-May-2014	-76	2 days <= ADY <= 11 days	7	1
18	1111_00001			41020	Week 2	TW2	2	15	26-May-2014	-69	12 days <= ADY <= 18 days	14	1
19	1111_00001			41040	Week 4	TW4	4	28	8-Jun-2014	-56	26 days <= ADY <= 35 days (5WK)	28	0
20	1111_00001			41060	Week 6	TW6	6	42	22-Jun-2014	-42	36 days <= ADY <= 49 days (7WK)	42	0
21	1111_00001			41080	Week 8	TW8	8	56	6-Jul-2014	-28	50 days <= ADY <= 63 days (9WK)	56	0
22	1111_00001			41100	Week 10	TW10	10	70	20-Jul-2014	-14	64 days <= ADY <= 77 days (11WK)	70	0
23	1111_00001			41120	Week 12	TW12	12	84	3-Aug-2014	0	78 days <= ADY <= 98 days (14WK)	84	0
24	1111_00001			54004	Follow-up Week 4	FW4	104	119	7-Sep-2014	35	21 days (3WK) <= FUADY <= 69 days	28	7
25	1111_00001	Y	Y	41020	Week 2	TW2	2	15	26-May-2014	-69	12 days <= ADY <= 18 days	14	1
26	1111_00001	Y	Y	41020	Week 2	TW2	2	15	26-May-2014	-69	12 days <= ADY <= 18 days	14	1
27	1111_00001	Y	Y	41040	Week 4	TW4	4	28	8-Jun-2014	-56	26 days <= ADY <= 35 days (5WK)	28	0
28	1111_00001	Y	Y	41040	Week 4	TW4	4	28	8-Jun-2014	-56	26 days <= ADY <= 35 days (5WK)	28	0
29	1111_00001	Y	Y	41120	Week 12	TW12	12	84	3-Aug-2014	0	78 days <= ADY <= 98 days (14WK)	84	0
30	1111_00001	Y	Y	41120	Week 12	TW12	12	84	3-Aug-2014	0	78 days <= ADY <= 98 days (14WK)	84	0
31	1111_00001	Y	Y	54004	Follow-up Week 4	FW4	104	119	7-Sep-2014	35	21 days (3WK) <= FUADY <= 69 days	28	7
32	1111_00001	Y	Y	54004	Follow-up Week 4	FW4	104	119	7-Sep-2014	35	21 days (3WK) <= FUADY <= 69 days	28	7
33	1111_00001	N	N			FW12	112				70 days (10WK) <= FUADY <= 146 days	84	
34	1111_00001	N	N			FW12	112				70 days (10WK) <= FUADY <= 146 days	84	
35	1111_00001	N	N			FW12	112				70 days (10WK) <= FUADY <= 146 days	84	
36	1111_00001	N	N			FW24	124				147 days (21WK) <= FUADY	168	
37	1111_00001	N	N			FW24	124				147 days (21WK) <= FUADY	168	
38	1111_00001	N	N			FW24	124				147 days (21WK) <= FUADY	168	

Table 3.3 PARAMxx variables and CRITY variables

1	USUBJID	PARAM	PARAMCD	PARAMN	PARAMTY	CRIT1	CRIT1FL	CRIT2	CRIT2FL	CRIT3	CRIT3FL	DTYPE
2	1111_00001	Hepatitis C Viral RNA (IU/mL)	HCRNA	1								
3	1111_00001	Hepatitis C Viral RNA (IU/mL)	HCRNA	1								
4	1111_00001	Hepatitis C Viral RNA (IU/mL)	HCRNA	1								
5	1111_00001	Hepatitis C Viral RNA (IU/mL)	HCRNA	1								
6	1111_00001	Hepatitis C Viral RNA (IU/mL)	HCRNA	1		Intermittent missing	Y					PHANTOM
7	1111_00001	Hepatitis C Viral RNA (IU/mL)	HCRNA	1								
8	1111_00001	Hepatitis C Viral RNA (IU/mL)	HCRNA	1								
9	1111_00001	Hepatitis C Viral RNA (IU/mL)	HCRNA	1								
10	1111_00001	Hepatitis C Viral RNA (IU/mL)	HCRNA	1								
11	1111_00001	Hepatitis C Viral RNA (IU/mL)	HCRNA	1								
12	1111_00001	Hepatitis C Viral RNA (IU/mL)	HCRNA	1								
13	1111_00001	Hepatitis C Viral RNA (IU/mL)	HCRNA	1						Monotone missing	Y	PHANTOM
14	1111_00001	Hepatitis C Viral RNA (IU/mL)	HCRNA	1						Monotone missing	Y	PHANTOM
15	1111_00001	Hepatitis C Viral RNA (log10 IU/mL)	HCRNAL	2	DERIVED							
16	1111_00001	Hepatitis C Viral RNA (log10 IU/mL)	HCRNAL	2	DERIVED							
17	1111_00001	Hepatitis C Viral RNA (log10 IU/mL)	HCRNAL	2	DERIVED							
18	1111_00001	Hepatitis C Viral RNA (log10 IU/mL)	HCRNAL	2	DERIVED							
19	1111_00001	Hepatitis C Viral RNA (log10 IU/mL)	HCRNAL	2	DERIVED							
20	1111_00001	Hepatitis C Viral RNA (log10 IU/mL)	HCRNAL	2	DERIVED							
21	1111_00001	Hepatitis C Viral RNA (log10 IU/mL)	HCRNAL	2	DERIVED							
22	1111_00001	Hepatitis C Viral RNA (log10 IU/mL)	HCRNAL	2	DERIVED							
23	1111_00001	Hepatitis C Viral RNA (log10 IU/mL)	HCRNAL	2	DERIVED							
24	1111_00001	Hepatitis C Viral RNA (log10 IU/mL)	HCRNAL	2	DERIVED							
25	1111_00001	HCV RNA Undetectable (TND) at Treatment Week 2	WK2TND	3	DERIVED							
26	1111_00001	HCV RNA < 15 IU/mL (either TD(u) or TND) at Treatment Week 2	WK2TDU	4	DERIVED							
27	1111_00001	HCV RNA Undetectable (TND) at Treatment Week 4	WK4TND	5	DERIVED							
28	1111_00001	HCV RNA < 15 IU/mL (either TD(u) or TND) at Treatment Week 4	WK4TDU	6	DERIVED							
29	1111_00001	HCV RNA Undetectable (TND) at Treatment Week 12	WK12TND	7	DERIVED							
30	1111_00001	HCV RNA < 15 IU/mL (either TD(u) or TND) at Treatment Week 12	WK12TDU	8	DERIVED							
31	1111_00001	HCV RNA Undetectable (TND) at 4 Weeks After End of Treatment	SVR4TND	9	DERIVED							
32	1111_00001	Sustained Viral Response 4 Weeks After End of Treatment (SVR4)	SVR4	10	DERIVED							
33	1111_00001	Sustained Viral Response 12 Weeks After End of Treatment (SVR12)	SVR12	11	DERIVED					Monotone missing	Y	PHANTOM
34	1111_00001	Sustained Viral Response 12 Weeks After End of Treatment (SVR12)	SVR12	11	DERIVED					Monotone missing	Y	TRD=F
35	1111_00001	Sustained Viral Response 12 Weeks After End of Treatment (SVR12)	SVR12	11	DERIVED					Monotone missing	Y	M=F
36	1111_00001	Sustained Viral Response 24 Weeks After End of Treatment (SVR24)	SVR24	12	DERIVED					Monotone missing	Y	PHANTOM
37	1111_00001	Sustained Viral Response 24 Weeks After End of Treatment (SVR24)	SVR24	12	DERIVED					Monotone missing	Y	TRD=F
38	1111_00001	Sustained Viral Response 24 Weeks After End of Treatment (SVR24)	SVR24	12	DERIVED					Monotone missing	Y	M=F

Table 3.4 Analysis results variables

1	USUBJID	LBORRES	LBORRESU	AVAL	AVALC	BASE	ABLFL	CHG	ANL01FL	SAVISIT	SRCSEQ	SRCDOM	SRCVAR
2	1111_00001	5000000	IU/mL	5000000	TDq						1	LB	LBORRES
3	1111_00001	3400000	IU/mL	3400000	TDq	3400000	Y	0	Y		2	LB	LBORRES
4	1111_00001	300	IU/mL	300	TDq	3400000		-3399700	Y		3	LB	LBORRES
5	1111_00001	HCV RNA not detected	IU/mL	1	TND	3400000		-3399999	Y		4	LB	LBORRES
6	1111_00001				MISSING				Y				
7	1111_00001	HCV RNA not detected	IU/mL	1	TND	3400000		-3399999	Y		5	LB	LBORRES
8	1111_00001	HCV RNA not detected	IU/mL	1	TND	3400000		-3399999	Y		6	LB	LBORRES
9	1111_00001	HCV RNA not detected	IU/mL	1	TND	3400000		-3399999	Y		7	LB	LBORRES
10	1111_00001	HCV RNA not detected	IU/mL	1	TND	3400000		-3399999	Y		8	LB	LBORRES
11	1111_00001	HCV RNA not detected	IU/mL	1	TND	3400000		-3399999	Y		9	LB	LBORRES
12	1111_00001	HCV RNA not detected	IU/mL	1	TND	3400000		-3399999	Y		10	LB	LBORRES
13	1111_00001				MISSING				Y				
14	1111_00001				MISSING				Y				
15	1111_00001	5000000	IU/mL	6.69897									
16	1111_00001	3400000	IU/mL	6.53148		6.53148	Y	0	Y				
17	1111_00001	300	IU/mL	2.47712		6.53148		-4.05436	Y				
18	1111_00001	HCV RNA not detected	IU/mL	0		6.53148		-6.53148	Y				
19	1111_00001	HCV RNA not detected	IU/mL	0		6.53148		-6.53148	Y				
20	1111_00001	HCV RNA not detected	IU/mL	0		6.53148		-6.53148	Y				
21	1111_00001	HCV RNA not detected	IU/mL	0		6.53148		-6.53148	Y				
22	1111_00001	HCV RNA not detected	IU/mL	0		6.53148		-6.53148	Y				
23	1111_00001	HCV RNA not detected	IU/mL	0		6.53148		-6.53148	Y				
24	1111_00001	HCV RNA not detected	IU/mL	0		6.53148		-6.53148	Y				
25	1111_00001	HCV RNA not detected	IU/mL	1	SUCCESS				Y				
26	1111_00001	HCV RNA not detected	IU/mL	1	SUCCESS				Y				
27	1111_00001	HCV RNA not detected	IU/mL	1	SUCCESS				Y				
28	1111_00001	HCV RNA not detected	IU/mL	1	SUCCESS				Y				
29	1111_00001	HCV RNA not detected	IU/mL	1	SUCCESS				Y				
30	1111_00001	HCV RNA not detected	IU/mL	1	SUCCESS				Y				
31	1111_00001	HCV RNA not detected	IU/mL	1	SUCCESS				Y				
32	1111_00001	HCV RNA not detected	IU/mL	1	SUCCESS				Y				
33	1111_00001				MISSING				Y				
34	1111_00001				MISSING				Y				
35	1111_00001			2	FAILURE				Y				
36	1111_00001				MISSING				Y				
37	1111_00001				MISSING				Y				
38	1111_00001			2	FAILURE				Y				

In this study, analysis timepoints are defined by relative day (ADY) or relative follow-up day (FUADY), then the variables (AWRANGE, AWTARGET and AWTDIFF) in Table 3.2 may be used along with ADY or FUADY to clarify how the record representing each analysis timepoint was chosen from among the possible candidates. The three variables AWRANGE, AWTARGET and AWTDIFF won't be used for statistical analysis, but it can be used for checking analysis window correction.

For some complicated statistical analysis, add rows within a parameter with DTYPE="Nominal -Value" is good practice. In this study, DTYPE="TRD=F" or "M=F" are derived for the sensitivity analysis. Usually, add the rows within a parameter are more convenient for analysis and reduce variable creation.

In some cases, the SDTM variables are kept in ADaM as well for the traceability and review. In the HCV study example, the schedule visit variables VISIT and VISITNUM, measurement source value LBORRES and LBORRESU are carried in ADaM. These SDTM variables in ADaM have identical formats.

In general, a column or row is supportive if it is not required in order to perform an analysis but is included in order to facilitate traceability or review.

ORDERING OF VARIABLES

Most of the data designer usually ignores the ordering of variables in ADaM datasets. However, the well variable order usually facilitates the review of the ADaM dataset. The variables in the ADaM dataset follow a logical ordering, not simply alphabetic. The proposal ordering of variables are in below.

- 1) Variables within same variable group should be arranged together, e.g., record-level treatment variables, timing variables, analysis parameters variables

- 2) The SDTM variables should be arranged appropriately in some variable group
- 3) Some supportive variables should be closed to main variables
- 4) The subject-level variables carried from adsl should be the same ordering as adsl.

For example, in Table 3.3, the variables CRIT1, CRIT2 and CRIT3 indicated the missing data type, intermittent missing or monotone missing, are closed to DTYPE. They really facilitate the reviewers to know how handling the missing data.

CONCLUSION

A well-organized ADaM-BDS efficacy dataset is not only for creation of statistical analysis, but also for statistician or FDA reviewer's easy-review. Here is the proposal how to create the ADaM compliance and easy-reviewed analysis datasets.

- 1) The user should get familiar with the CDISC ADaM validation checks and adopt the Checks to make first round eyeball ADaM compliance checking
- 2) Organize the ADaM datasets associated metadatas to make clearly data contents and data's lineage
- 3) Think more about the data structure and content, like supportive row and column
- 4) ordering of variables

REFERENCES

[1] CDISC ADaM Implementation Guide, Version 1.1

[2] CDISC Analysis Data Model, Version 2.1

[3] CDISC ADaM Validation Checks v1.3

[4] Michelle Barrick, John Troxell (2014) A Guide to the ADaM Basic Data Structure for Dataset Designers, PharmaSUG 2014

[5] FDA guidance: Chronic Hepatitis C Virus Infection: Developing Direct-Acting Antiviral Drugs for Treatment, May 2016

CONTACT INFORMATION

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

Name: Huadan Li
Enterprise: MSD R&D (China) Co., Ltd.
Address: 1-13F, Building 21 Rongda Road, Wangjing R&D Base, Zhongguancun Electronic Zone West Zone, Chaoyang District,
City, State ZIP: Beijing 100012
Work Phone: +86 10 58609432
E-mail: hua.dan.li@merck.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.