

Overview of ADAM – Analysis Data Model

Anagha Bhatkhande

Varsha Mahajan

November 1, 2017

TCS confidential

| Copyright © 2017 Tata Consultancy Services Limited

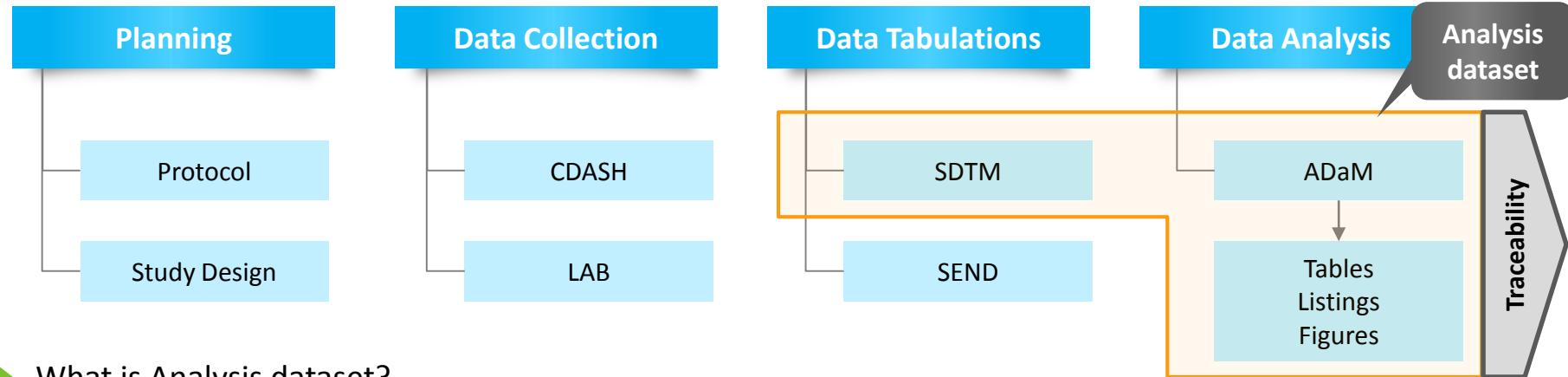
Confidentiality Statement



This document contains information which is commercially confidential to Tata Consultancy Services Ltd. ("TCS") and is provided for the sole purpose of permitting the recipient to evaluate the proposal submitted within the document. In consideration of the receipt of this document, the recipient agrees to maintain the information it contains in confidence and not to reproduce or otherwise disclose the contents of this document to any person outside the recipient's evaluation team or nominated advisors. The document may not be used, reproduced or disclosed to others, except as specifically permitted by TCS in writing. The recipient of this document, by its retention and use, agrees to take reasonable endeavours to protect the document and the information it contains from loss or theft.

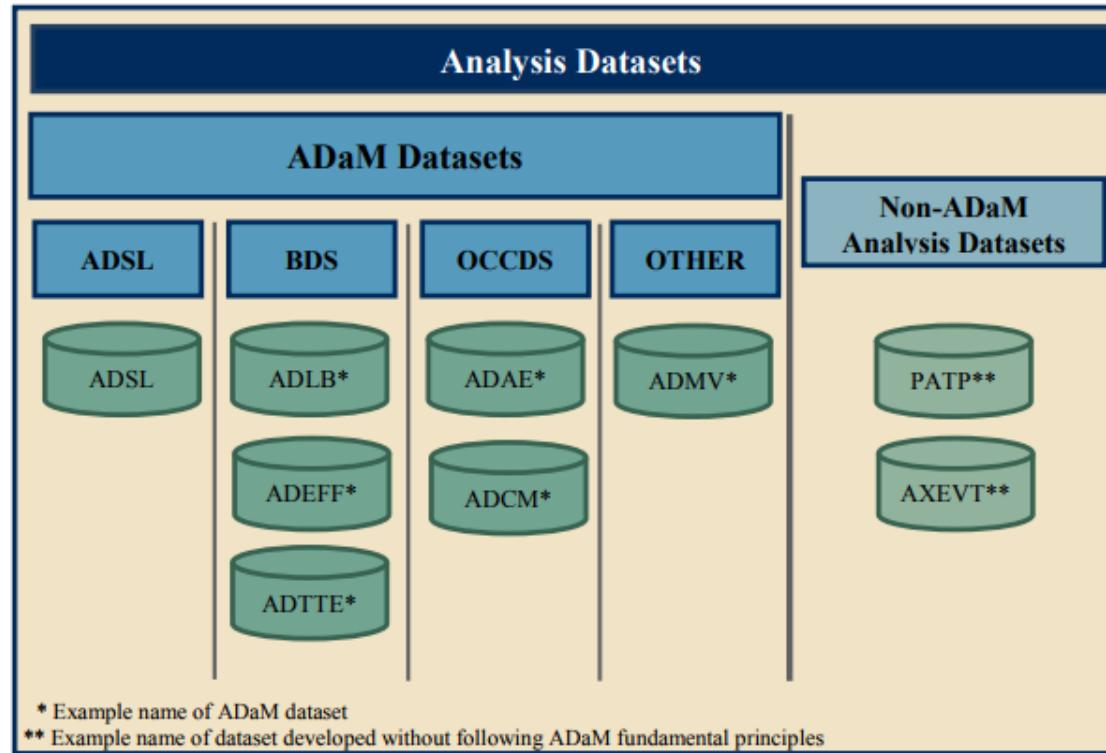
© 2017 - TATA Consultancy Services Limited (TCS)

Why do we need ADaM?



- ▶ What is Analysis dataset?
- ▶ What is ADaM dataset?
- ▶ Does ADaM have any basic structure?
- ▶ What are Non-ADaM analysis dataset?
- ▶ How to differentiate ADaM and Non-ADaM datasets via the naming conventions?

Analysis Datasets



Picture courtesy ADaM Implementation guideline CDISC.org



ADaM Data Structures

- ▶ Clear and Unambiguous connect between the Analysis datasets and associated Metadata supporting the statistical analyses performed in a clinical study – communicate the content and source of the datasets
- ▶ Provide Traceability – data's lineage or relationship between analysis value and its predecessor (derived, imputed)
- ▶ Readily usable with commonly available software tools
- ▶ Ease to perform statistical analysis with minimal programming – one proc away
- ▶ Ease data review and re-creation of specific statistical Analysis



Traceability

- Assumption: Source for ADaM datasets are SDTM
 - Reality : Source for ADaM can be another ADaM
- Why Traceability?
- A CDISC compliant submission includes both SDTM and ADaM datasets ; relationship between SDTM and ADaM must be clear.
 - ADaM tries to balance 'Traceability' and 'Analysis-Ready' with a predictable structure
 - Contains derived data along with directly imported SDTM datasets
 - In case of records/observation not used for analysis are dropped then appropriate analysis flags are added to ease analysis
 - Traceability is key for reviewers since it helps identify
 - Information that exists in the submitted SDTM data
 - Information that is derived or imputed within the ADaM dataset
 - The method used to create derived or imputed data
 - Information used for analyses versus not used for analyses but is included in the dataset to support traceability or future analysis

Traceability

- Two levels of Traceability:
 - Metadata traceability:
 - Required for ADaM compliance
 - Establishes the relationship of the analysis variable to its source dataset(s) and variables by describing the algorithm used to derive or populate it
 - Establishes the relationship between an analysis result & analysis dataset(s)



Metadata traceability

HCSEQ	AVISITN	AVISIT	VISITNUM	VISIT	HCRES	HCSTRESN	ANL02FL	AVAL	DTYPE	Dataset	ADHC							
1	900	Screening	20001	SCREENING	9763165	9763165		9763165		Program Name	Adhc.sas							
2	901	Day -1	21001	DAY -1	12396132	12396132		12396132		Description	HCV RNA Analysis Dataset							
3	902	Day 1 Pre-Dose	30001	DAY 1	5076583	5076583		5076583		Unique Identifier Variables	USUBID APHASEN AVISITN HCDCD HCRES							
3.5	950	Baseline					Y	9763165	MEDIAN	STRUCTURE	One record per HCV RNA assessment per time point per subject							
4	1001.06	Day 1 6H	30001	DAY 1	6390354	6390354	Y	6390354		General Class	Findings							
										Input Datasets	HC, DM, DS							
										Notes	Includes all enrolled subjects							
5	1001.12	Day 1 12H	30001	DAY 1	5410749	5410749	Y	5410749		Variable Name	Variable Label	Type	Length	Controlled Terms or Formats	Origin	Role	Comments	Core
6	1002	Day 2	30002	DAY 2	825410	825410	Y	825410		USUBID	Unique Subject Identifier	Char	40		HC.USUBID	Identifier	Equivalent to studyid '_' strip(siteid) '_' strip(subjid)	Req
7	1004	Day 4	30004	DAY 4	645024	645024	Y	645024		HCSEQ	SEQUENCE NUMBER	Num	8		HC.HCSEQ	Identifier	Equals to HC.hcseq For a calculated baseline record (avisitn=950), the value is derived from HC.hcseq(where hcfl='Y') +0.5. For placeholder records hcseq is 0.01 more than the sequence number corresponding to the previous HCV RNA assessment. This variable is mainly used to establish traceability	Perm
8	1008	Week 1	30008	WEEK 1	1191916	1191916	Y	1191916		APHASEN	Phase Number	Num	8	AOHASEN (APHASE): (1) 0 = Pre-Treatment Phase (2) 1 = On-Treatment Phase (3) 2 = Post-Treatment Phase	Derived	Timing	If HC.hcseq<DM.rfstdtc then aphasen=0; Else if DM.rfstdtc<=HC.hcseq<=DM.rfendtc+14 aphasen=1; Else if HC.hcseq>DM.rfendtc+14 then aphasen=2;	Perm
9	1015	Week 2	30015	WEEK 2	392325	392325	Y	392325		ANL02FL	Analysis Record Flag 02	Char	2	YESF: (1) Y	Derived	Analysis	Indicates the analysis record in a visit window in Overall treatment and Follow-up phase Populated only for records with ontrfl="Y" or abfl="Y" or aphasen=2 If there are multiple records in a visit window then one closest to target date is set to "Y". In case of two equidistant records from the target date then the latest record in time is set to "Y"	Cond
10	1022	Week 3	30022	WEEK 3	386255	386255	Y	386255		AVAL	Analysis Value	Num	8		Derived	Analysis	Equals to hcstresn for all other records except avisitn=950. For avisitn=950 use the median of pre-dose HCV RNA assessments	Req
11	1029	Week 4	30029	WEEK 4	96117	96117	Y	96117										
12	1057	Week 8	30057	WEEK 8	7096	7096	Y	7096										
13	1085	Week 12	30085	WEEK 12	412	412	Y	412										
14	1113	Week 16	30113	WEEK 16	38	38	Y	38										
16	1141	Week 20	30141	WEEK 20	<25	17.5	Y	17.5										
18	1141	Week 20	30141	WEEK 20	<25	17.5		17.5										
20	1169	Week 24	30169	WEEK 24	UNDETECTED	5	Y	5										
22	1253	Week 36	30253	WEEK 36	<25	17.5	Y	17.5										
23	1253	Week 36	80001	UNSCHEDULED	<25	17.5		17.5										
24	1253	Week 36	80001	UNSCHEDULED	<25	17.5		17.5										
25	1337	Week 48	80001	UNSCHEDULED	UNDETECTED	5		5										
26	1337	Week 48	80001	UNSCHEDULED	<25	17.5		17.5										
27	1337	Week 48	80001	UNSCHEDULED	UNDETECTED	5	Y	5										
28	1337	Week 48	30337	WEEK 48 (EOT)	UNDETECTED	5		5										
29	2029	Antiviral Follow-up Week 4	70004	SAFETY FOLLOW-UP	105		105	Y	105									

Metadata traceability

HCSEQ	ANL02FL	AVAL
1		9763165
2		12396132
3		5076583
3.5	Y	9763165
4	Y	6390354
5	Y	5410749
6	Y	825410
7	Y	645024
8	Y	1191916
9	Y	392325
10	Y	386255
11	Y	96117
12	Y	7096
13	Y	412
14	Y	38
16	Y	17.5
18		17.5
20	Y	5
22	Y	17.5
23		17.5
24		17.5
25		5
26		17.5
27	Y	5
28		5
29	Y	105

HCSEQ	SEQUENCE NUMBER	Num	8		HC.HCSEQ	Identifier	Origin	Derivation	Perm
							Equals to HC.hcseq For a calculated baseline record (avisitn=950), the value is derived from HC.hcseq(where hcfl=Y) +0.5. For placeholder records hcseq is 0.01 more than the sequence number corresponding to the previous HVC RNA assessment This variable is mainly used to establish traceability		

ANL02FL	Analysis Record Flag 02	Char	2	YESF: (1) Y	Derived	Analysis	Role	Type	Cond	
							Indicates the analysis record in a visit window in Overall treatment and Follow-up phase Populated only for records with ontrtfl="Y" or abfl="Y" or aphasen=2 If there are multiple records in a visit window then one closest to target date is set to "Y". In case of two equidistant records from the target date then the latest record in time is set to "Y"			

AVAL	Analysis Value	Num	8		Derived	Analysis	Req
							Equals to hcstresn for all other records except avisitn=950. For avisitn=950 use the median of pre-dose HCV RNA assessments

Traceability

- Data point traceability:
 - Implemented if practical and feasible
 - Points directly to the specific predecessor record from SDTM to ADaM i.e. SDTM variables that are copied in to ADaM and should not be modified
 - Helps to trace the path of a complex data manipulation
 - Primary candidates for data point traceability:
 - SDTM Domain variable value
 - SDTM Source variable name
 - Relevant SDTM domain –SEQ value
 - These 3 with the analysis variable help link back to the source SDTM records

Variable Name	Variable Label	Type	CDISC Notes
SRCDOM	Source Domain	Char	The 2-character identifier of the SDTM domain that relates to the derived analysis value
SRCVAR	Source Variable	Char	The name of the column (in the SDTM domain identified by SRCDOM) that relates to the derived analysis value
SRCSEQ	Source Sequence Number	Num	The sequence number SEQ of the row (in the SDTM domain identified by SRCDOM) that relates to the derived analysis value

Data Point Traceability

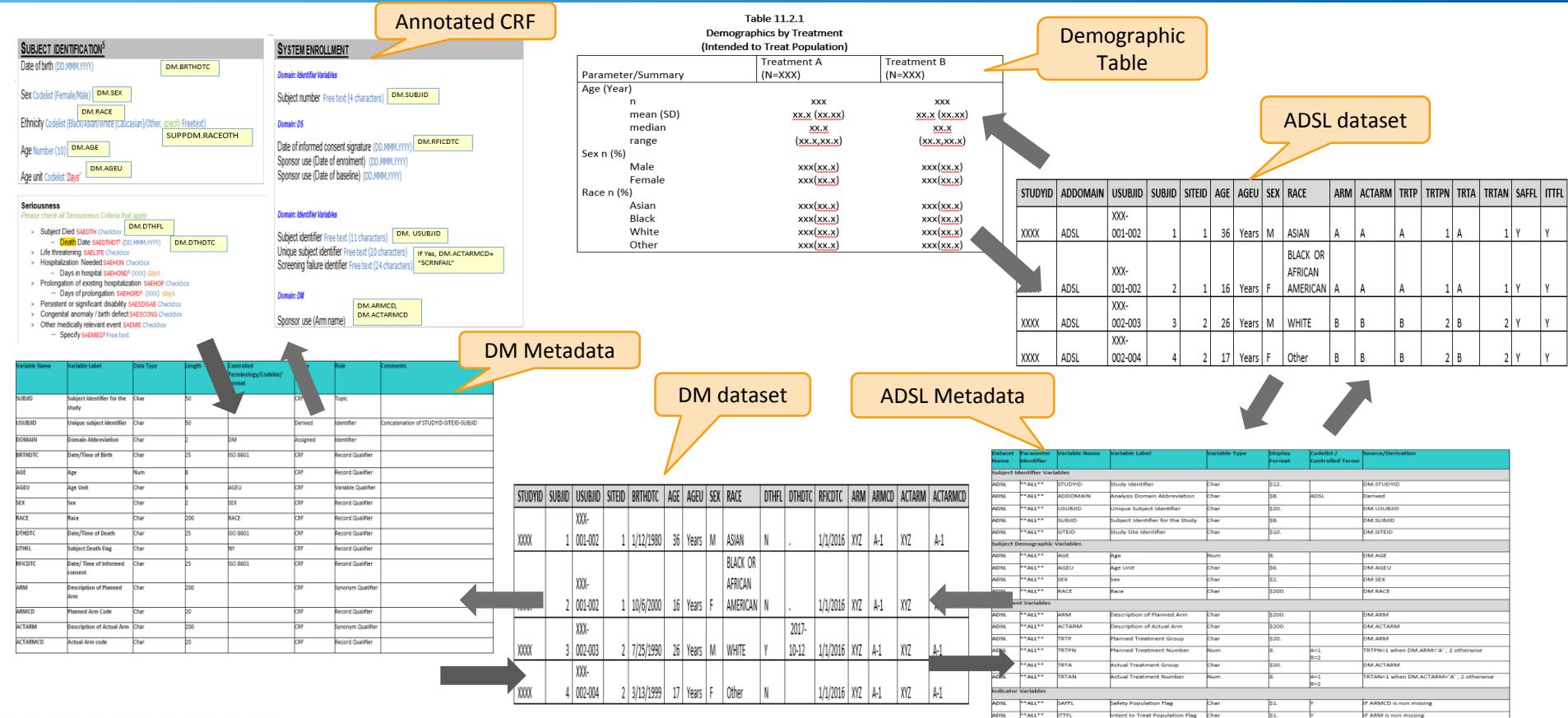
HCSEQ	AVISITN	AVISIT	ANL02FL	AVAL	RVRFL	RVRFN	SRCDOM	SRCVAR	SRCSEQ	UNDW24FL	UNDW24FN	DTYPE
1	900	Screening		3744926								
2	901	Day -1		4371834								
3	902	Day 1 Pre-Dose		2541566								
3.5	950	Baseline	Y	3744926								
4	1001.06	Day 1 6H	Y	6764238								
5	1001.12	Day 1 12H	Y	3567054								
6	1002	Day 2	Y	5081353								
7	1004	Day 4	Y	1225725								
8	1008	Week 1	Y	2653658								
9	1015	Week 2	Y	1819450								
10	1022	Week 3	Y	543214								
11	1029	Week 4	Y	104363	N		0	HC	HCSTRESN	11		
12	1057	Week 8	Y	1799								
13	1085	Week 12	Y	211	.	HC	HCSTRESN	13				
14	1113	Week 16	Y	53								
15	1141	Week 20	Y	17.5								
16	1169	Week 24		17.5								
17	1169	Week 24	Y	5	.	HC	HCSTRESN	17	Y			1

Derivation type

Variable Name	Variable Label	Type	Length	Controlled Terms or Formats	Comments	Metadata
RVRFL	Rapid Viral Response Flag	Char	2		Equals to "Y" if a subject has undetectable HCV RNA at Week 4 i.e. 3 SAS Global Forum 2012 Posters HCRES="UNDETECTED" at avisitn=1029 and anl02fl="Y". Else equals to "N".	
RVRFN	Rapid Viral Response Flag (N)	Num	8	YESNOFN (RVRFL); (1) 1 = Y (2) 0 = N	Equals to 1 if rvrfi="Y". Equals to 0 if rvrfi="N"	
UNDW24FL	Undetectable HCV RNA at Week 24	Char	2		Equals to "Y" if a subject has undetectable HCV RNA at Week 24 i.e. HCRES="UNDETECTED" at avisitn=1169 and anl02fl="Y". Else equals to "N".	
UNDW24FN	Undetectable HCV RNA at Week 24 (N)	Num	8	YESNOFN (UNDW24FL); (1) 1 = Y (2) 0 = N	Equals to 1 if undw24fl="Y" Equals to 0 if undw24fl="N"	
SRCDOM	Source Domain	Char	4		T="Y" T="N"	
SRCVAR	Source Variable	Char	8		Equal to "HCSTRESN" for avisitn=1029 and anl02fl="Y" Equal to "HCSTRESN" for avisitn=1169 and anl02fl="Y"	
SRCSEQ	Source Sequence Number	Num	8		= HCSEQ for avisitn=1029 and anl02fl="Y" = HCSEQ for avisitn=1169 and anl02fl="Y"	

Data point

Traceability Cycle



ADaM Data Structures

- ▶ Subject Level analysis Structure: The subject level analysis dataset: One record per subject e.g.: ADSL
- ▶ Occurrence Data Structure (ODS): Contains all the event domains of SDTM. e.g.: ADAE
- ▶ Basic Data Structure(BDS): Contains one or more records per subject, per analysis parameter, per analysis time point. Mostly contain the finding domains of SDTM e.g.: ADLB
- ▶ Others



Standard ADaM Variables

- ADaM variables are classified as:
 - Required (Req) : This variable must be included in the dataset
e.g. : STUDYID, USUBJID, SUBJID, SITEID, etc.
 - Conditional (Cond) : This variable must be included in the dataset in certain circumstances
e.g. : FASFL, SAFFL, ITTFL, etc.
 - Permissible (Perm) : This variable may be included in the dataset, but is not required
e.g. : SITEGRy, TRTxPN, etc.

All ADaM variables are populated as per their classification ; nulls are allowed



ADaM variable Conventions

The variable conventions are defined to ensure compliance with SAS version 5 transport file format and Oracle constraints

General Variable Conventions

- Variable names should be less than or equal to 8 characters in length
- Variable names to start with a letter and not underscore
- Variable names can be comprised only of letters (A-Z), underscore(_) and numerals (0-9)
- Variable labels should be less than or equal to 40 characters in length
- All ADaM character variable length should not be more than 200 characters in length
 - In SDTM if a character variable has length more than 200 characters then the remaining portion is added in SUPP domain but in ADaM it is in the same domain.



ADaM variable Conventions

General Variable Conventions

- Any SDTM variable directly copied in to ADaM dataset should maintain its attributes as in the SDTM (same name, same meaning, same values)
- Any ADaM standard variable defined must be used, even if the content is a direct copy of the SDTM variable e.g. AVAL variable in ADLB is just a copy of LBSTRESN the dataset must contain AVAL
- In case of paired variables, the secondary variable is presented in the dataset only in the presence of primary variable. Relevant suffix is used only for the secondary variable (e.g. C-character, N-numeric) e.g. TRTP and TRTPN



General Variable Conventions

- Character variables converted to Numeric have to be suffixed with 'N'. In case the variable name is already 8 characters then the last character is replaced by 'N'
- The character date variables converted to Numeric have the following naming convention *DT, *TM and *DTM
- Character flag variables end in 'FL'; numeric in 'FN'
- Variables whose names end in GRy, Gy, or CATy are grouping variables, where y refers to the grouping scheme or algorithm (not the category within the grouping); Numeric version of the variable will use the suffix GRyN (or GyN)
- Grouping variable created from a predecessor variable should begin with the name of the variable being grouped e.g. AGEN, AGECAT

Variables with xx, y and zz parameters

- xx Parameter : The xx parameter represents a two-digit period. Valid values are 01–99 e.g. : TRTxP etc.
- y Parameter : The y parameter represents a single-digit grouping or category, an analysis criterion, or an analysis range. Valid values are 1–9 e.g. : SITEGRy etc.
- zz Parameter : The zz parameter represents a two-digit index for the zzth record selection algorithm e.g. : ANLzzFL etc.

Timing variable Conventions

Variables with names ending in

- DT are numeric dates
- DTM are numeric datetimes
- TM are numeric times
- DY are relative day variables

Names of timing start variables end with

- A 'S' followed by the two characters indicating the type of timing (e.g., SDT, STM)

Names of timing end variables end with

- An 'E' followed by the two characters indicating the type of timing (e.g., EDT, ETM)



Flag and Group variables

- Subject level population flags in the analysis datasets end in FL and FN (Character and Numeric); values (Yes, No and 1, 0)
- Parameter level population flags end in PFL and PFN; values (Yes,1) null records not included
- Record level population flags end in RFL and RFN; values (Yes,1) null records not included
- Date time imputation flags end in F
 - *DTF represent Date imputation flags
 - *TMF represent Time imputation flags
 - Both *DTF and *TMF describe the level of imputation
 - *DTF can take values Blank, D, M, Y
 - *TMF can take values H, M, S



Additional Fragments

Fragments	Definition
BL	Baseline, position relative to type of variable. E.g. SBP01BL, SBP02BL Systolic blood pressure for two periods
CHG	Change, position relative to type of variable. E.g. SBP01CHG Systolic blood pressure change from baseline for period 1
FU	Follow-up, position relative to type of variable. E.g. FUENDT Follow-up end date
OT	On treatment, position relative to type of variable. E.g. ONTRTFL On treatment Flag
RU	Run-in, position relative to type of variable. E.g. RUSDT Run-in start date
SC	Screening, position relative to type of variable. E.g. WEIGHTSC or WTSC
U	Units, suffix
WA	Washout, position relative to type of variable. E.g. WA01SDT, WA01EDT, WA02SDT and WA02EDT Wash out start and end dates for two periods



ADaM Treatment Variables

- Treatment variables are required in all analysis datasets
- Actual Treatment variables are required when they do not match with the planned treatment

Variable Name	Label	Type	Source
TRTxnP	Planned Treatment for Period xx	Char	TRTxnP might be derived from the SDTM DM variable ARM. At least TRT01P is required
TRTxpN	Planned Treatment for Period xx (N)	Num	The numeric code variable for TRTxnP. One-to-one mapping to TRTxpN within a study
TRTxxA	Actual Treatment for Period xx	Char	Subject-level identifier that represents the actual treatment for the subject for period xx. Required when actual treatment does not match planned.
TRTxAN	Actual Treatment for Period xx (N)	Num	TRTxAN Actual Treatment for Period xx (N). One-to-one mapping to TRTxxA within a study



ADaM Common Variables

Identifiers
USUBJID,
STUDYID

**Demographic
Variables e.g.**
**AGE, RACE,
SEX etc.**

**Population
Flags e.g.**
**ITTFL, SAFFL,
FASFL etc.**

**Baseline Flag
variable i.e.**
ABLFL

**Treatment
Variables e.g.**
**TRTxnP,
TRTxxA**



Navigating through ADaM IG

Anagha Bhatkhande

Varsha Mahajan

November 1, 2017

TCS confidential

| Copyright © 2017 Tata Consultancy Services Limited

The screenshot shows a PDF document with a dark grey header bar. On the far left, there is a vertical sidebar containing a 'Bookmarks' section with a tree view of the document's structure. The 'Cover Page' is the currently selected item. Other items in the tree include 'Table of Contents', '1 Introduction', '2 Fundamentals of the ADaM Standard', '3 Standard ADaM Variables', '4 Implementation Issues, Standard Solutions, and Examples', and 'Appendices'. The main content area features the CDISC logo at the top, followed by the title 'Analysis Data Model Implementation Guide' in large, bold, serif font, and 'Version 1.1' below it. A subtitle 'Prepared by the' is followed by 'CDISC Analysis Data Model Team'. Below this is a thick black horizontal bar. At the bottom of the page, there is a callout box with a black border containing the heading 'Notes to Readers' and the text 'This Implementation Guide is Version 1.1 and corresponds to Version 2.1 of the CDISC Analysis Data Model.' To the right of the callout box, there is a small icon of a red PDF file and the text 'ADaM IG'.

CDISC

**Analysis Data Model
Implementation Guide**

Version 1.1

Prepared by the

CDISC Analysis Data Model Team

Notes to Readers

This Implementation Guide is Version 1.1 and corresponds to Version 2.1 of the CDISC Analysis Data Model.

ADaM IG



354.03 213.48
+0.13% +0.48%

End of Session

ADSL – Subject Level Analysis Dataset

Anagha Bhatkhande

Varsha Mahajan

November 1, 2017

TCS confidential

| Copyright © 2017 Tata Consultancy Services Limited

ADSL : Subject-Level Analysis dataset

- ▶ ADSL dataset and its related metadata are required in a CDISC-based submission of a clinical trial even if no other ADaM datasets are submitted
- ▶ The structure of ADSL is one record per subject; for any clinical trial design
- ▶ ADSL is used to provide key facts about the subject ; enabling analysis and interpretation
- ▶ ADSL does not contain time varying data and endpoints
- ▶ ADSL is a source for subject-level variables used in other datasets to enable analysis of the BDS dataset
- ▶ Only one ADSL per study contains variables integral to design and/or randomization
- ▶ Structure and content make this analysis dataset a cornerstone for any trial

ADSL : Subject-Level Analysis dataset

► ADSL contains variables such as

- Subject-level population flags
- Treatment variables – planned and actual
- Completion status
- Important dates
- Disease history
- Baseline Information
- Demographic variables
- Stratification variables
- Statistical model covariates
- Subgrouping variables etc.

ADSL : Subject-Level Analysis dataset

► Notes:

- To create ADSL dataset, it is necessary to review the following sections in SAP to understand how the analysis was planned: Population definitions, Schedule of assessments, Definition of baseline values, Analysis covariates and Definition of prior/concurrent medications.
- The Demographics Domain Model (DM) and Exposure (EX) domains from the SDTM model are typically required.
- Screen Failures : Whether analysis datasets include data for subjects not analyzed (e.g., screen failures) is a sponsor decision and should be communicated with the reviewers or users of the data. If these data are included, they should be incorporated in the appropriate analysis datasets such as ADSL (as opposed to separate datasets for non-analyzed subjects) using appropriate flag variables to clearly differentiate these records. The metadata must specify that these data are included and how to distinguish them.

ADSL : Subject-Level Analysis dataset

Must match SDTM variables and values in metadata

Table 3.2.1 ADSL Identifier Variables

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core	CDISC Notes
STUDYID	Study Identifier	Char		Req	DM.STUDYID
USUBJID	Unique Subject Identifier	Char		Req	DM.USUBJID
SUBJID	Subject Identifier for the Study	Char		Req	DM.SUBJID. SUBJID is required in ADSL, but permissible in other datasets.
SITEID	Study Site Identifier	Char		Req	DM.SITEID. SITEID is required in ADSL, but permissible in other datasets.
SITEGRy	Pooled Site Group y	Char		Perm	Character description of a grouping or pooling of clinical sites for analysis purposes. For example, SITEGR3 is the name of a variable containing site group (pooled site) names, where the grouping has been done according to the third site grouping algorithm, defined in variable metadata; SITEGR3 does not mean the third group of sites.
SITEGRyN	Pooled Site Group y (N)	Num		Perm	The numeric code for SITEGRy. One-to-one mapping to SITEGRy within a study.
REGIONy	Geographic Region y	Char		Perm	Character description of geographical region. For example, REGION1 might have values of 'Asia', 'Europe', 'North America', 'Rest of World'; REGION2 might have values of 'United States', 'Rest of World'.
REGIONyN	Geographic Region y (N)	Num		Perm	The numeric code for REGIONy. Orders REGIONy for analysis and reporting. One-to-one mapping to REGIONy within a study.

Table 3.2.2 ADSL Subject Demographics Variables

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core	CDISC Notes
AGE	Age	Num		Req	DM.AGE. If analysis needs require a derived age that does not match DM.AGE, then AAGE must be added.
AGEU	Age Units	Char	(AGEU)	Req	DM.AGEU
AGEGRy	Pooled Age Group y	Char		Perm	Character description of a grouping or pooling of the subject's age for analysis purposes. For example, AGEGR1 might have values of "<18", "18-65", and ">65"; AGEGR2 might have values of "Less than 35 y old" and "At least 35 y old".
AGEGRyN	Pooled Age Group y (N)	Num		Perm	The numeric code for AGEGRy. Orders the grouping or pooling of subject age for analysis and reporting. One-to-one mapping to AGEGRy within a study.
AAGE	Analysis Age	Num		Cond	Age used for analysis that is derived differently from DM.AGE. AAGE is required if age is calculated differently than in SDTM.
SEX	Sex	Char	(SEX)	Req	The sex of the subject is a required variable in ADSL; must be identical to DM.SEX.
RACE	Race	Char	(RACE)	Req	The race of the subject is a required variable in ADSL; must be identical to DM.RACE.
RACEGRy	Pooled Race Group y	Char		Perm	Character description of a grouping or pooling of the subject's race for analysis purposes.
RACEGRyN	Pooled Race Group y (N)	Num		Perm	The numeric code for RACEGRy. Orders the grouping or pooling of subject race for analysis and reporting. One-to-one mapping to RACEGRy within a study.

ADSL : Subject-Level Analysis dataset

Must match SDTM variables and values in metadata

Table 3.2.3 ADSL Population Indicator Variables

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core	CDISC Notes
FASFL	Full Analysis Set Population Flag	Char	Y, N	Cond	These flags identify whether or not the subject is included in the specified population. A minimum of one subject-level population flag variable is required in ADSL.
SAFFL	Safety Population Flag	Char	Y, N	Cond	Not all of the indicators listed here need to be included in ADSL. As stated in Section 3.1.4, Item 2, only those indicators corresponding to populations defined in the statistical analysis plan or populations used as a basis for analysis need be included in ADSL.
ITTFL	Intent-To-Treat Population Flag	Char	Y, N	Cond	This list of flags is not meant to be all-inclusive. Additional population flags may be added.
PPROTFL	Per-Protocol Population Flag	Char	Y, N	Cond	The values of subject-level population flags cannot be blank. If a flag is used, the corresponding numeric version (*FN, where 0=no and 1=yes) of the population flag can also be included. Please also refer to Section 3.1.4.
COMPLFL	Completers Population Flag	Char	Y, N	Cond	
RANDFL	Randomized Population Flag	Char	Y, N	Cond	
ENRLFL	Enrolled Population Flag	Char	Y, N	Cond	

Must have at least one population flag

All applicable subject-level population flags must be present

Table 3.2.4 ADSL Treatment Variables

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core	CDISC Notes
ARM	Description of Planned Arm	Char		Req	DM.ARM
ACTARM	Description of Actual Arm	Char		Perm	DM.ACTARM
TRTxnP	Planned Treatment for Period xx	Char		Req	Subject-level identifier that represents the planned treatment for period xx. In a one-period randomized trial, TRT01P would be the treatment to which the subject was randomized. TRTxnP might be derived from the SDTM DM variable ARM. At least TRT01P is required.
TRTxpN	Planned Treatment for Period xx (N)	Num		Perm	The numeric code variable for TRTxnP. One-to-one mapping to TRTxnP within a study.
TRTxxA	Actual Treatment for Period xx	Char		Cond	Subject-level identifier that represents the actual treatment for the subject for period xx. Required when actual treatment does not match planned and there is an analysis of the data as treated.
TRTxxAN	Actual Treatment for Period xx (N)	Num		Perm	The numeric code variable for TRTxxA. One-to-one mapping to TRTxxA within a study.

Must have at least one TRTxnP variable

ADSL : Subject-Level Analysis dataset

Added if required as per the study design and analysis

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core	CDISC Notes
TRTSEQP	Planned Sequence of Treatments	Char		Cond	Required when there is an analysis based on the sequence of treatments, for example in a crossover design. TRTSEQP is not necessarily equal to ARM, for example if ARM contains elements that are not relevant to analysis of treatments or ARM is not fully descriptive (e.g., "GROUP 1," "GROUP 2"). When analyzing based on the sequence of treatments, TRTSEQP is required even if identical to ARM.
TRTSEQPN	Planned Sequence of Treatments (N)	Num		Perm	Numeric version of TRTSEQP. One-to-one mapping to TRTSEQP within a study.
TRTSEQA	Actual Sequence of Treatments	Char		Cond	TRTSEQA is required if a situation occurred in the conduct of the trial where a subject received a sequence of treatments other than what was planned and there is an analysis based on the sequence of treatments.
TRTSEQAN	Actual Sequence of Treatments (N)	Num		Perm	Numeric version of TRTSEQA. One-to-one mapping to TRTSEQA within a study.
TRxxPGy	Planned Pooled Treatment y for Period xx	Char		Perm	Planned pooled treatment y for period xx. Useful when planned treatments (TRTxnP) in the specified period xx are pooled together for analysis according to pooling algorithm y. For example when in period 2 the first pooling algorithm dictates that all doses of Drug A (TR02PG1="All doses of Drug A") are pooled together for comparison to all doses of Drug B (TR02PG1="All doses of Drug B"). Each value of TRTxnP is pooled within at most one value of TRxxPGy.
TRxxPGyN	Planned Pooled Trt y for Period xx (N)	Num *		Perm	The numeric code for TRxxPGy. One-to-one mapping to TRxxPGy within a study.
TRxxAGy	Actual Pooled Treatment y for Period xx	Char		Cond	Actual pooled treatment y for period xx. Required when TRxxPGy is present and TRTxXA is present.
TRxxAGyN	Actual Pooled Trt y for Period xx (N)	Num *		Perm	The numeric code for TRxxAGy. One-to-one mapping to TRxxAGy within a study.
TSEQPGy	Planned Pooled Treatment Sequence y	Char		Perm	Planned pooled treatment sequence y. Useful when planned treatment sequences (TRTSEQP) are pooled together for analysis according to pooling algorithm y. For example, this might be used in an analysis of an extension study when the analysis is based on what the subject received in the parent study as well as in the extension study.
TSEQPGyN	Planned Pooled Treatment Sequence y (N)	Num		Perm	Numeric version of TSEQPGy. One-to-one mapping to TSEQPGy within a study.
TSEQAGy	Actual Pooled Treatment Sequence y	Char		Cond	Actual pooled treatment sequence y. Required when TSEQPGy is present and TRTSEQA is present.
TSEQAGyN	Actual Pooled Treatment Sequence y (N)	Num		Perm	Numeric version of TSEQAGy. One-to-one mapping to TSEQAGy within a study.

* TRxxPGyN and TRxxAGyN were mistakenly indicated as character variables in ADaMIG v1.0. The error is corrected above.

ADSL : Subject-Level Analysis dataset

Table 3.2.5 ADSL Dose Variables

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core	CDISC Notes
DOSExxP	Planned Treatment Dose for Period xx	Num		Perm	Subject-level identifier that represents the planned treatment dosage for period xx.
DOSExxA	Actual Treatment Dose for Period xx	Num		Perm	Subject-level identifier that represents the actual treatment dosage for period xx.
DOSExxU	Units for Dose for Period xx	Char		Perm	The units for DOSExxP and DOSExxA. It is permissible to use suffixes such as "P" and "A" if needed, with labels modified accordingly.

Table 3.2.6 ADSL Treatment Timing Variables

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core	CDISC Notes
TRTSDT	Date of First Exposure to Treatment	Num		Cond	Date of first exposure to treatment for a subject in a study. TRTSDT and/or TRTSDTM are required if there is an investigational product. Note that TRTSDT is not required to have the same value as the SDTM DM variable RFXSTDTC. While both of these dates reflect the concept of first exposure, the ADaM date may be derived to support the analysis which may not necessarily be the very first date in the SDTM EX domain.
TRTSTM	Time of First Exposure to Treatment	Num		Perm	Time of first exposure to treatment for a subject in a study.
TRTSDTM	Datetime of First Exposure to Treatment	Num		Cond	Datetime of first exposure to treatment for a subject in a study. TRTSDT and/or TRTSDTM are required if there is an investigational product.
TRTSDF	Date of First Exposure Imput. Flag	Char	(DATEFL)	Cond	The level of imputation of date of first exposure to treatment. If TRTSDT (or the date part of TRTSDTM) was imputed, TRTSDF must be populated and is required. See Section 3.1.3 .
TRTSTMF	Time of First Exposure Imput. Flag	Char	(TIMEFL)	Cond	The level of imputation of time of first exposure to treatment. If TRTSTM (or the time part of TRTSDTM) was imputed, TRTSTMF must be populated and is required. See Section 3.1.3 .
TRTEDT	Date of Last Exposure to Treatment	Num		Cond	Date of last exposure to treatment for a subject in a study. TRTEDT and/or TRTEDTM are required if there is an investigational product. Note that TRTEDT is not required to have the same value as the SDTM DM variable RFXENDTC. While both of these dates reflect the concept of last exposure, the ADaM date may be derived to support the analysis which may not necessarily be the very last date in the SDTM EX domain.
TRTEM	Time of Last Exposure to Treatment	Num		Perm	Time of last exposure to treatment for a subject in a study.

Selected trial dates required in case of an investigational product

ADSL : Subject-Level Analysis dataset

Added if required as per the study design and analysis

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core	CDISC Notes
TRTEDTM	Datetime of Last Exposure to Treatment	Num		Cond	Datetime of last exposure to treatment for a subject in a study. TRTEDT and/or TRTEDTM are required if there is an investigational product.
TRTEDTF	Date of Last Exposure Imput. Flag	Char	(DATEFL)	Cond	The level of imputation of date of last exposure to treatment. If TRTEDT (or the date part of TRTEDTM) was imputed, TRTEDTF must be populated and is required. See Section 3.1.3 .
TRTETMF	Time of Last Exposure Imput. Flag	Char	(TIMEFL)	Cond	The level of imputation of time of last exposure to treatment. If TRTETM (or the time part of TRTEDTM) was imputed, TRTETMF must be populated and is required. See Section 3.1.3 .
TRxxSDT	Date of First Exposure in Period xx	Num		Cond	Date of first exposure to treatment in period xx. TRxxSDT and/or TRxxSDTM are only required in trial designs where multiple treatments are given to the same subject, such as a crossover design, but are permissible for other trial designs. Also useful in designs where multiple periods exist for the same treatment (i.e., multiple cycles of the same study treatment).
TRxxSTM	Time of First Exposure in Period xx	Num		Cond	The starting time of exposure to treatment in period xx. TRxxSTM and/or TRxxSDTM are only required in trial designs where multiple treatments are given to the same subject, such as a crossover design (but are permissible for other trial designs), and time is important to the analysis.
TRxxSDTM	Datetime of First Exposure in Period xx	Num		Cond	Datetime of first exposure to treatment in period xx. TRxxSDT and/or TRxxSDTM are only required in trial designs where multiple treatments are given to the same subject, such as a crossover design, but are permissible for other trial designs.
TRxxSDTF	Date 1st Exposure Period xx Imput. Flag	Char	(DATEFL)	Cond	The level of imputation of date of first exposure to treatment in period xx. If TRxxSDT (or the date part of TRxxSDTM) was imputed, TRxxSDTF must be populated and is required. See Section 3.1.3 .
TRxxSTMF	Time 1st Exposure Period xx Imput. Flag	Char	(TIMEFL)	Cond	The level of imputation of time of first exposure to treatment in period xx. If TRxxSTM (or the time part of TRxxSDTM) was imputed, TRxxSTMF must be populated and is required. See Section 3.1.3 .
TRxxEDT	Date of Last Exposure in Period xx	Num		Cond	Date of last exposure to treatment in period xx. TRxxEDT and/or TRxxEDTM are only required in trial designs where multiple treatments are given to the same subject, such as a crossover design, but are permissible for other trial designs.
TRxxETM	Time of Last Exposure in Period xx	Num		Cond	The ending time of exposure to treatment in period xx. TRxxETM and/or TRxxEDTM are only required in trial designs where multiple treatments are given to the same subject, such as a crossover design, and ending time is important to the analysis, but are permissible for other trial designs.
TRxxEDTM	Datetime of Last Exposure in Period xx	Num		Cond	The datetime of last exposure to treatment in period xx. TRxxEDT and/or TRxxEDTM are only required in trial designs where multiple treatments are given to the same subject, such as a crossover design, but are permissible for other trial designs.
TRxxEDTF	Date Last Exposure Period xx Imput. Flag	Char	(DATEFL)	Cond	The level of imputation of date of last exposure to treatment in period xx. If TRxxEDT (or the date part of TRxxEDTM) was imputed, TRxxEDTF must be populated and is required. See Section 3.1.3 .
TRxxETMF	Time Last Exposure Period xx Imput. Flag	Char	(TIMEFL)	Cond	The level of imputation of time of last exposure to treatment in period xx. If TRxxETM (or the time part of TRxxEDTM) was imputed, TRxxETMF must be populated and is required. See Section 3.1.3 .

ADSL : Subject-Level Analysis dataset

Important Notes

- A set of timing variables can be included in ADSL only if the definitions are set fixed across the study
- Varying timepoint definitions for period start and stop even if required of safety or efficacy analysis should not be included
- Additional subject level subperiod, phase timing, trial experience variables can be added
- EPOCH variable becomes important to be retained in all ADaM datasets in case of submission

Refer ADaM Implementation Guide V1.1 for all the details





354.03 213.48
+0.13% +0.48%

End of Session

BDS – Basic Data Structure

Anagha Bhatkhande

Varsha Mahajan

November 2, 2017

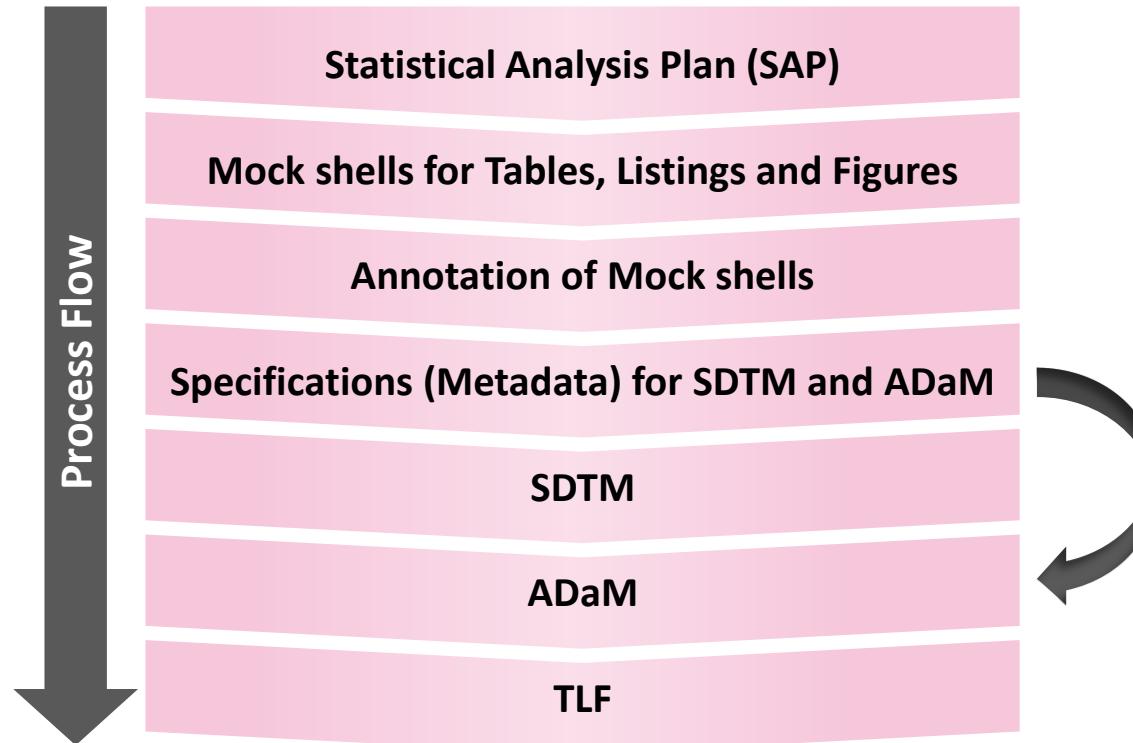
TCS confidential

| Copyright © 2017 Tata Consultancy Services Limited

BDS : Basic Data structure

- BDS dataset contains one or more records per subject, per analysis parameter, per analysis timepoint (only if present)
- Classification of variables in the BDS datasets
 - Subject Identifier variables
 - Treatment variables
 - Timing variables
 - Analysis Parameter variables
 - Analysis Descriptor variables
 - Indicator variables
 - Data point traceability variables
 - Analysis-Enabling variables

Relation between ADaM BDS and Mock up tables



Example Mock shell

Table 14.4.2
Change from Baseline for Biochemistry parameter
(Safety Population)

Parameter: Creatinine

Following variables will be needed to create this output:

1. Population variable – SAFFL
2. Laboratory Parameter – PARAMCD, PARAM
3. Treatment variable – TRTAN, TRTA
4. Baseline information – BASE
5. Change information – CHG
6. Postbaseline – AVAL and AVISIT

Treatment	Baseline						Postbaseline						Change					
	n	Mean	SD	Median	Min	Max	Mean	SD	Median	Min	Max	Mean	SD	Median	Min	Max		
Treat1 (N=XXX)	xxx	xx.x	xx.xx	xx.x	xx.x	xx.x	xx.x	xx.xx	xx.x	xx.x	xx.x	xx.x	xx.xx	xx.x	xx.x	xx.x		
Treat2 (N=XXX)	xxx	xx.x	xx.xx	xx.x	xx.x	xx.x	xx.x	xx.xx	xx.x	xx.x	xx.x	xx.x	xx.xx	xx.x	xx.x	xx.x		

n: patients with valid baseline and postbasline information

Postbaseline is the last non-missing information for the parameter post the intervention which is the Endpoint

Programming note: Repeat for all Biochemistry parameters in alphabetical order



Annotated Mock Shell

Table 14.4.2
Change from Baseline for Biochemistry parameter
(Safety Population)

ADSL.TRTAN	ADLB.BASE	Baseline					Postbaseline					Change				
		n	Mean	SD	Median	Min	Max	Mean	SD	Median	Min	Max	Mean	SD	Median	Min
Treatment ↗																
Treat1 (N=XXX)	xxx	xx.x	xx.xx	xx.x	xx.x	xx.x	xx.x	xx.x	xx.xx	xx.x	xx.x	xx.x	xx.x	xx.xx	xx.x	xx.x
Treat2 (N=XXX)	xxx	xx.x	xx.xx	xx.x	xx.x	xx.x	xx.x	xx.x	xx.xx	xx.x	xx.x	xx.x	xx.x	xx.xx	xx.x	xx.x

n: patients with valid baseline and postbasline information

Postbaseline is the last non-missing information for the parameter post the intervention which is the Endpoint

Programming note: Repeat for all Biochemistry parameters in alphabetical order

ADLB – BDS Structure

Dataset Name	Parameter Identifier	Variable Name	Variable Label	Variable Type	Display Format	Codelist / Controlled Terms	Source/Derivation
Subject Identifier Variables							
ADLB	**ALL**	STUDYID	Study Identifier	Char	\$12.		ADSL.STUDYID
ADLB	**ALL**	ADDOMAIN	Analysis Domain Abbreviation	Char	\$8.	ADLB	Derived
ADLB	**ALL**	USUBJID	Unique Subject Identifier	Char	\$20.		LB.USUBJID
ADLB	**ALL**	SUBJID	Subject Identifier for the Study	Char	\$8.		ADSL.SUBJID
ADLB	**ALL**	SITEID	Study Site Identifier	Char	\$10.		ADSL.SITEID
Treatment Variables							
ADLB	**ALL**	TRTA	Actual Treatment Group	Char	\$20.		ADSL.TRTA
ADLB	**ALL**	TRTAN	Actual Treatment Number	Num	8.		ADSL.TRTAN
Timing Variables							
ADLB	**ALL**	AVISIT	Analysis Timepoint Description	Char	\$50.	BASELINE VISIT 1	LB.VISIT
ADLB	**ALL**	AVISITN	Analysis Timepoint Number	Num	8.	0=BASELINE 1=VISIT 1	LB.VISITNUM
Analysis Parameter Variables							
ADLB	**ALL**	PARAM	Parameter Description	Char	\$100.		LB.LBTEST + unit
ADLB	**ALL**	PARAMCD	Parameter Code	Char	\$8.		LB.LBTESTCD
ADLB	**ALL**	AVAL	Analysis Value	Num	8.		LB.LBSTRESN
ADLB	**ALL**	BASE	Numeric Baseline Value	Num	8.		ADLB.AVAL at AVISIT="BASELINE"
ADLB	**ALL**	CHG	Change from Baseline	Num	8.		ADLB.AVAL - ADLB.BASE

New columns added

ADLB – BDS Structure

Dataset Name	Parameter Identifier	Variable Name	Variable Label	Variable Type	Display Format	Codelist / Controlled Terms	Source/Derivation
Descriptor Variables							
ADLB	**ALL**	DTYPE	Derivation Type	Char	\$10		<p>LOCF = last observation carried forward for post baseline endpoint calculation</p> <p>Create a new row per parameter with last non-missing post-baseline value, with the value of AVISITN=99 and AVISIT='ENDPOINT', retain the rest of the information from the selected record.</p> <p>AVG = average of pre-screening, screening, baseline values for baseline record</p> <p>Create a new row per parameter with average of baseline and screening and pre-screening values, with the value of AVISITN=1 and AVISIT=BASELINE, retain the rest of the information from the latest selected record.</p>
Indicator Variables							
ADLB	**ALL**	SAFFL	Safety Population Flag	Char	\$1.	Y	ADSL.SAFFL
ADLB	**ALL**	ABLFL	Baseline Flag	Char	\$1.	Y	'Y' at ADLB.AVISIT='BASELINE'
Data Point Traceability Variables							
ADLB	**ALL**	SRCDOM	Source Domain	Char	\$8.	LB	
ADLB	**ALL**	SRCVAR	Source Variable	Char	\$8.	LBSTRESN	
ADLB	**ALL**	SRCSSEQ	Source Sequence Number	Num	8.		LB.LBSEQ

New row created

ADLB – BDS Data

STUDY	ADDOMAIN	USUBJID	SUBJID	SITEID	TRTA	TRTAN	AVISIT	AVISITN	PARAM	PARAMCD	AVAL	BASE	CHG	DTYPE	SAFFL	ABLFL	SRCDOM	SRCVAR	SRCSEQ
XXXX	ADLB	XXX-001-001	1	1	A	1	PRE-SCREENING	-10	CREATININE (mg/L)	CREA	110	.	.	.	Y	.	LB	LBSTRESN	LBSEQ
XXXX	ADLB	XXX-001-001	1	1	A	1	SCREENING	1	CREATININE (mg/L)	CREA	109	.	.	.	Y	.	LB	LBSTRESN	LBSEQ
XXXX	ADLB	XXX-001-001	1	1	A	1	BASELINE	2	CREATININE (mg/L)	CREA	110	.	.	.	Y	Y	LB	LBSTRESN	LBSEQ
XXXX	ADLB	XXX-001-001	1	1	A	1	BASELINE	2	CREATININE (mg/L)	CREA	109.7	109.7	.	AVG	Y	Y	LB	LBSTRESN	LBSEQ
XXXX	ADLB	XXX-001-001	1	1	A	1	VISIT 3	3	CREATININE (mg/L)	CREA	114	109.7	4.3333	.	Y	.	LB	LBSTRESN	LBSEQ
XXXX	ADLB	XXX-001-001	1	1	A	1	VISIT 4	4	CREATININE (mg/L)	CREA	102	109.7	-7.667	.	Y	.	LB	LBSTRESN	LBSEQ
XXXX	ADLB	XXX-001-001	1	1	A	1	VISIT 5	5	CREATININE (mg/L)	CREA	121	109.7	11.333	.	Y	.	LB	LBSTRESN	LBSEQ
XXXX	ADLB	XXX-001-001	1	1	A	1	ENDPOINT	99	CREATININE (mg/L)	CREA	121	109.7	11.333	LOCF	Y	.	LB	LBSTRESN	LBSEQ



Summary of Key BDS parameters

Parameter (PARAM) and Parameter Code (PARAMCD)

- PARAM describes, and must uniquely and sufficiently identify, the contents of the relevant analysis value variable AVAL or AVALC. PARAMCD is the code for PARAM is map one to one

Differences between PARAM and SDTM xxTEST

- SDTM xxTEST is designed to work in conjunction with other variables called qualifiers in order to describe the collected result and PARAM is the only variable that describes AVAL or AVALC, qualifiers are not allowed.

Parameter Category (PARCATy)

- PARCATy is used to group parameters into categories

Parameter Type (PARAMTYP)

- Parameter Type (PARAMTYP) flags "derived" parameters ; often misleading when used with DTYPE and will be retired in the next version



Summary of Key BDS parameters

Derivation Type (DTYPE)

- DTYPE serves two functions: when populated on a given record,
 - (1) it indicates that the record is derived from other records within the same parameter and
 - (2) it identifies the algorithm used to derive the analysis value (AVAL or AVALC) on the record

Baseline Type (BASETYPE)

- BASETYPE refers to a definition of baseline that characterizes the value of BASE on that row in case of multiple baseline definition

Baseline Record Flag (ABLFL)

- Baseline flag ABLFL must be set to "Y" on the row whose analysis value AVAL is used to populate BASE for that parameter and subject

Analysis Visit (AVISIT)

- Analysis visit AVISIT is used to describe the analysis visit or conceptual timepoint characterizing the row



Summary of Key BDS parameters

BDS Population Flags

- ▶ Subject-Level Population Flags: Subject-level population flags are contained in ADSL, and can be copied to BDS datasets.
e.g.: ITTFL
- ▶ Record-Level Population Flags: Record-level population flag variable names end in RFL . A record-level flag is useful when there are reasons for excluding some records within subject and parameter from the analysis.
e.g.: PPROTRFL="Y"
- ▶ Parameter-Level Population Flags: Parameter-level population flag variable names end in PFL. A parameter-level flag is useful when observed data influence whether or not a subject is considered to be in a population for reporting purposes.
e.g.: for the analysis of a given parameter, a subject is included in the analysis if
 - (1) the subject is a member of the ITT population, and
 - (2) for the given parameter, the subject has a non-missing baseline value (BASE) and at least one non-missing post-baseline value of AVAL. For the subjects who meet both conditions, the parameter-level Full Analysis Set population flag FASPFL is set to "Y" on all rows for the parameter





354.03 213.48
+0.13% +0.48%

End of Session

ADaM variable rules

Anagha Bhatkhande
Varsha Mahajan

November 2, 2017

TCS confidential

| Copyright © 2017 Tata Consultancy Services Limited

Rules for creation of rows and columns in ADaM

RULE 1 :

A parameter-invariant function of AVAL and BASE on the same row that does not involve a transform of BASE should be added as a new column.

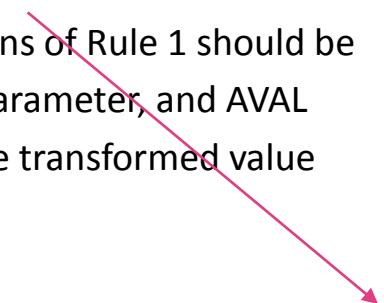
Example of Rule 1: CHG or PCHG as both are derived from AVAL and BASE on the same row and are calculated for all the rows and there is no transformation of BASE.

	USUBJID	AVISIT	PARAM	PARAMCD	AVAL	PARCAT1	BASE	PARAMTYP	DTYPE	CHG	PCHG
1	XYZ-10001		GAPS01-Overall Efficacy,type 3	GAPS0101	. GAPS01	
2	XYZ-10001	Screening	IPSS01-Incomplete emptying	IPSS0101	1 IPSS01	1				.	
3	XYZ-10001	Screening	IPSS01-Frequency	IPSS0102	1 IPSS01	1				.	
4	XYZ-10001	Screening	IPSS01-Intermittency	IPSS0103	1 IPSS01	1				.	
5	XYZ-10001	Screening	IPSS01-Urgency	IPSS0104	2 IPSS01	2				.	
6	XYZ-10001	Screening	IPSS01-Weak stream	IPSS0105	3 IPSS01	3				.	
7	XYZ-10001	Screening	IPSS01-Straining	IPSS0106	3 IPSS01	3				.	
8	XYZ-10001	Screening	IPSS01-Nocturia	IPSS0107	2 IPSS01	2				.	
9	XYZ-10001	Screening	IPSS01-Feel about Life	IPSS0108	3 IPSS01	3				.	
10	XYZ-10001	Screening	IPSS01-Total Score	IPSS0199	13 IPSS01	13 DERIVED				.	
11	XYZ-10001	Visit 5	IPSS01-Incomplete emptying	IPSS0101	0 IPSS01	1				.	
12	XYZ-10001	Visit 5	IPSS01-Frequency	IPSS0102	1 IPSS01	1				.	
13	XYZ-10001	Visit 5	IPSS01-Intermittency	IPSS0103	1 IPSS01	1				.	
14	XYZ-10001	Visit 5	IPSS01-Urgency	IPSS0104	2 IPSS01	2				.	
15	XYZ-10001	Visit 5	IPSS01-Weak stream	IPSS0105	1 IPSS01	3				.	
16	XYZ-10001	Visit 5	IPSS01-Straining	IPSS0106	1 IPSS01	3				.	
17	XYZ-10001	Visit 5	IPSS01-Nocturia	IPSS0107	2 IPSS01	2				.	
18	XYZ-10001	Visit 5	IPSS01-Feel about Life	IPSS0108	. IPSS01	3				.	
19	XYZ-10001	Visit 5	IPSS01-Total Score	IPSS0199	8 IPSS01	13 DERIVED				.	
20	XYZ-10001	Screening	SFI01-Lack of Sex Drive	SFI0101	4 SFI01	4				-5	-38.46153
21	XYZ-10001	Screening	SFI01-Ability to get and keep Erections	SFI0102	2 SFI01	2				.	
22	XYZ-10001	Screening	SFI01-Ejaculation Problem	SFI0103	1 SFI01	1				.	
23	XYZ-10001	Screening	SFI01-Total Score	SFI0199	7 SFI01	7 DERIVED				.	
24	XYZ-10001	Visit 3	SFI01-Lack of Sex Drive	SFI0101	4 SFI01	4				.	
25	XYZ-10001	Visit 3	SFI01-Ability to get and keep Erections	SFI0102	2 SFI01	2				.	
26	XYZ-10001	Visit 3	SFI01-Ejaculation Problem	SFI0103	1 SFI01	1				.	
27	XYZ-10001	Visit 3	SFI01-Total Score	SFI0199	7 SFI01	7 DERIVED				0	
28	XYZ-10001	Visit 5	SFI01-Lack of Sex Drive	SFI0101	4 SFI01	4				.	
29	XYZ-10001	Visit 5	SFI01-Ability to get and keep Erections	SFI0102	4 SFI01	2				.	
30	XYZ-10001	Visit 5	SFI01-Ejaculation Problem	SFI0103	4 SFI01	1				.	
31	XYZ-10001	Visit 5	SFI01-Total Score	SFI0199	12 SFI01	7 DERIVED				5	71.4285714
32	XYZ-10002		GAPS01-Overall Efficacy,type 3	GAPS0101	. GAPS01	.				.	

Rules for creation of rows and columns in ADaM

RULE 2 :

A **transformation of AVAL** that does not meet the conditions of Rule 1 should be added as a new parameter, and AVAL should contain the transformed value



Row	PARAM	PARAMCD	VISIT	AVISIT	AVISITN	VISITNUM	ABLFL	AVAL	BASE	CHG
1	Weight (kg)	WEIGHT	Visit -1	Screening	4	1		99	100	.
2	Weight (kg)	WEIGHT	Visit 0	Run-In	-2	2		101	100	.
3	Weight (kg)	WEIGHT	Visit 1	Baseline	0	3	Y	100	100	0
4	Weight (kg)	WEIGHT	Visit 12	Week 24	24	4		94	100	-6
5	Weight (kg)	WEIGHT	Visit 24	Week 48	48	5		92	100	-8
6	Weight (kg)	WEIGHT	Visit 26	Week 52	52	6		95	100	-5
7	Log10(Weight (kg))	L10WT	Visit -1	Screening	-4	1		1.9956	2	.
8	Log10(Weight (kg))	L10WT	Visit 0	Run-In	-2	2		2.0043	2	.
9	Log10(Weight (kg))	L10WT	Visit 1	Baseline	0	3	Y	2	2	0
10	Log10(Weight (kg))	L10WT	Visit 12	Week 24	24	4		1.9731	2	-0.0269
11	Log10(Weight (kg))	L10WT	Visit 24	Week 48	48	5		1.9638	2	-0.0362
12	Log10(Weight (kg))	L10WT	Visit 26	Week 52	52	6		1.9777	2	-0.0223

Rules for creation of rows and columns in ADaM

RULE 3 :

A function of one or more rows within the same parameter for the purpose of creating an analysis timepoint should be added as a new row for the same parameter.

STUDY	ADDOMAIN	USUBJID	SUBJID	SITEID	TRTA	TRTAN	AVISIT	AVISITN	PARAM	PARAMCD	AVAL	BASE	CHG	DTYPE	SAFFL	ABFL	SRCDOM	SRCVAR	SRCSEQ
XXXX	ADLB	XXX-001-001		1	1A		PRE-1SCREENING		CREATININE -10 (mg/L)	CREA	110.	.	.	Y		LB	LBSTRESN	LBSEQ	
XXXX	ADLB	XXX-001-001		1	1A		1SCREENING		CREATININE 1(mg/L)	CREA	109.	.	.	Y		LB	LBSTRESN	LBSEQ	
XXXX	ADLB	XXX-001-001		1	1A		1BASELINE		CREATININE 2(mg/L)	CREA	110	.	.	Y	Y	LB	LBSTRESN	LBSEQ	
XXXX	ADLB	XXX-001-001		1	1A		1BASELINE		CREATININE 2(mg/L)	CREA	109.7	109.7	AVG	Y	Y	LB	LBSTRESN	LBSEQ	
XXXX	ADLB	XXX-001-001		1	1A		1VISIT 3		CREATININE 3(mg/L)	CREA	114	109.7	4.3333	Y		LB	LBSTRESN	LBSEQ	
XXXX	ADLB	XXX-001-001		1	1A		1VISIT 4		CREATININE 4(mg/L)	CREA	102	109.7	-7.667	Y		LB	LBSTRESN	LBSEQ	
XXXX	ADLB	XXX-001-001		1	1A		1VISIT 5		CREATININE 5(mg/L)	CREA	121	109.7	11.333	Y		LB	LBSTRESN	LBSEQ	
XXXX	ADLB	XXX-001-001		1	1A		1ENDPOINT		CREATININE 99(mg/L)	CREA	121	109.7	11.333 LOCF	Y		LB	LBSTRESN	LBSEQ	

Rules for creation of rows and columns in ADaM

RULE 4 :

A function of multiple rows within a parameter should be added as a new parameter.

Rule 4 is a special case of Rule 2. The functions covered by this rule violate the second condition of Rule 1 (they are not same-row functions of AVAL), and may also violate the first and third conditions.

Let $AVAL(t)$ equal the value of CD4 cell count at post-baseline visit t , and let $VISITDY(t)$ be the planned study day of visit t .

CD4AUC (cumulative daily CD4 count over follow-up) is calculated at any given post-baseline visit as follows:

- CD4AUC at baseline visit is set to 0.
- $CD4AUC(t) = CD4AUC(t-1) + [0.5 * AVAL(t-1) + 0.5 * AVAL(t)] * [VISITDY(t) - VISITDY(t-1)]$.

CD4AUCMB (cumulative average change from baseline in daily CD4 count over follow-up) is calculated as

- $CD4AUCMB(t) = CD4AUC(t) / [VISITDY(t) - 1] - \text{baseline value of CD4 cell count}$.

Row	PARAM	PARAMCD	AVISIT	VISITDY	ABLFL	AVAL	BASE
1	CD4 (cells/mm3)	CD4	Week -1	-7		75	76
2	CD4 (cells/mm3)	CD4	Week 0	1	Y	76	76
3	CD4 (cells/mm3)	CD4	Week 2	15		128	76
4	CD4 (cells/mm3)	CD4	Week 4	29		125	76
5	CD4 (cells/mm3)	CD4	Week 8	57		191	76
6	CD4 (cells/mm3)	CD4	Week 12	85		167	76
7	CD4 (cells/mm3)	CD4	Week 16	113		136	76
8	CD4 Cumulative AUC	CD4AUC	Week 0	1	Y	0	0
9	CD4 Cumulative AUC	CD4AUC	Week 2	15		1428	0
10	CD4 Cumulative AUC	CD4AUC	Week 4	29		3199	0
11	CD4 Cumulative AUC	CD4AUC	Week 8	57		7623	0
12	CD4 Cumulative AUC	CD4AUC	Week 12	85		12635	0
13	CD4 Cumulative AUC	CD4AUC	Week 16	113		16877	0
14	CD4 Cumulative AUCMB	CD4AUCMB	Week 2	15		26	-
15	CD4 Cumulative AUCMB	CD4AUCMB	Week 4	29		38.25	-
16	CD4 Cumulative AUCMB	CD4AUCMB	Week 8	57		60.125	-
17	CD4 Cumulative AUCMB	CD4AUCMB	Week 12	85		74.4167	-
18	CD4 Cumulative AUCMB	CD4AUCMB	Week 16	113		74.6875	-

Rules for creation of rows and columns in ADaM

RULE 5 :

A function of **more than one parameter** should be added as a new parameter

$$\text{BMI} = \text{Weight(kg)} / [\text{Height(m)}]^2$$

AVISIT	AVISITN	PARAM	PARAMCD	AVAL
BASELINE	2	Height (cm)	HGT	157
PRE-SCREENING	-10	Weight (kg)	WGT	60
SCREENING	1	Weight (kg)	WGT	61
BASELINE	2	Weight (kg)	WGT	61
VISIT 3	3	Weight (kg)	WGT	60.8
VISIT 4	4	Weight (kg)	WGT	60.5
VISIT 5	5	Weight (kg)	WGT	60
BASELINE	2	BMI (kg/m ²)	BMI	24.74745
VISIT 3	3	BMI (kg/m ²)	BMI	24.66632
VISIT 4	4	BMI (kg/m ²)	BMI	24.54461
VISIT 5	5	BMI (kg/m ²)	BMI	24.34176

Rules for creation of rows and columns in ADaM

RULE 6 :

When there is **more than one** definition of **baseline**, each additional definition of baseline requires the creation of its own set of rows.

	USUBJID	AVISITN	PARAMCD	AVAL	BASETYPE	ATPTN	APERIOD	ABLFL
1	XYZ-123-001-001	0	HR		60 STUDY LEVEL	.	.	.
2	XYZ-123-001-001	1	HR		55 STUDY LEVEL	-60	1	
3	XYZ-123-001-001	1	HR		55 STUDY LEVEL	-5	1	Y
4	XYZ-123-001-001	1	HR		58 STUDY LEVEL	5	1	
5	XYZ-123-001-001	1	HR		62 STUDY LEVEL	60	1	
6	XYZ-123-001-001	2	HR		59 STUDY LEVEL	-60	2	
7	XYZ-123-001-001	2	HR		60 STUDY LEVEL	-5	2	
8	XYZ-123-001-001	2	HR		58 STUDY LEVEL	5	2	
9	XYZ-123-001-001	2	HR		57 STUDY LEVEL	60	2	
10	XYZ-123-001-001	2	HR		59 PERIOD 2	--	2	
11	XYZ-123-001-001	2	HR		60 PERIOD 2	-5	2	Y
12	XYZ-123-001-001	2	HR		58 PERIOD 2	5	2	
13	XYZ-123-001-001	2	HR		57 PERIOD 2	60	2	



354.03 213.48
+0.13% +0.48%

End of Session

OCCDS - Occurrence Data Structure

Anagha Bhatkhande

Varsha Mahajan

November 2, 2017

TCS confidential

| Copyright © 2017 Tata Consultancy Services Limited

Occurrence Data structure (OCCDS)

- Occurrence analysis is the **counting of subjects** with a **given record or term**, and often includes a **structured hierarchy** of dictionary coding categories.
- e.g. Adverse Events (AE), Concomitant Medications (CM), and Medical History (MH)
- The structure for the occurrence analysis dataset is usually one record per each record in the corresponding SDTM domain.
 - Examples of when the number of records in the analysis dataset would not match the number in SDTM include:
 - SDTM data contain screen failures but screen failures are not analyzed. In this case, the screen failure records are not needed in the analysis dataset.
 - Adverse event or concomitant medication, spans several treatment periods and needs to be counted in each. Based on the analysis need, a separate row might be required for each treatment period spanned and analyzed.
 - An adverse event needs to be analyzed along multiple coding paths. In this case, a row would be needed for each coding path analyzed. An alternate solution, if multiple coding paths are not needed together, would be to put records for each coding path into a separate analysis dataset. E.g. SMQs , period/phase per coding path structure if required by the analysis.



Occurrence Data structure (OCCDS)

OCCDS variables include

- Identifier variables
- Dictionary coding and categorization variables (MedDRA for AE and MH and WHOdrug for CM)
- Timing variables
- Indicator variables
- Occurrence flag variables
- Treatment variables
- Descriptive variables

Refer ADaM structure for OCCDS V1.0 guideline



Adobe Acrobat
Document



Thank You