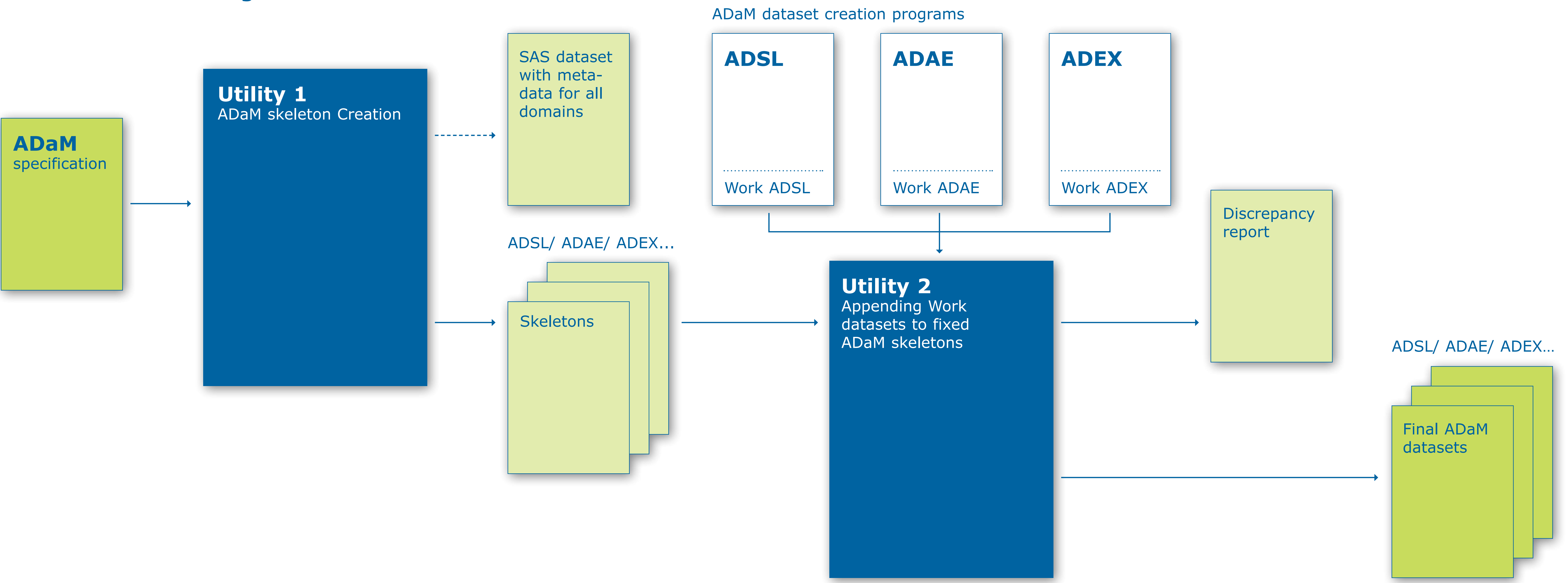


Ferring Pharmaceuticals, Copenhagen, Denmark.

Process overview diagram



Purpose

CDISC has laid strict Analysis Data Model Implementation Guideline (ADaMIG), specifying ADaM standard dataset structures and variables, including naming conventions, labels, types, formats etc.

A programming solution is required to effectively adhere to the standard without the risk of making mistakes and duplication of work. This poster describes a method which programmatically uses the ADaM specification to shape the structure of the final ADaM datasets.

- Challenges**
- Assigning variable attributes
 - Implementing changes
 - Maintaining standards
 - Governing sequence of the variables

- Solution**
- This ADaM standardization technique is based on two utilities which are used in pairs.
- The Utility 1 produces empty SAS datasets(ADaM skeletons) in a STANDARD library from the ADaM specification which is stored as a Microsoft Word file.
 - The Utility 2 appends, typically the last Work. dataset created during the development of the ADaM datasets for each domain to the respective ADaM skeleton created by Utility 1. Thus ensuring adherence to the ADaM specification.

1. ADaM specification

Subject level ADaM dataset specification

4. Analysis Dataset Specifications

4.1 ADSL - Subject Level Analysis Dataset

3 Analysis Dataset Metadata

Dataset Name	Dataset Description	Dataset Location	Dataset Structure	Key variables of Dataset	Class of Dataset	Documentation
ADSL	Subject demographics	./ADSL	One record per subject randomized	USUBJID	ADSL	SAP

Analysis Variable Metadata for ADSL

Dataset Name	Variable Name	Variable Label	Variable Type	Display Format	Codelist/ Controlled Terms	Source/ Derivation
ADSL	STUDYID	Study Identifier for the Study	Char	\$20		
ADSL	USUBJID	Unique Subject Identifier	Char	\$30		
ADSL	SITEID	Study Site Identifier	Char	\$30		
ADSL	COUNTRY	Country	Char	\$3		
ADSL	ARM	Description of Planned Arm	Char	\$200		
ADSL	ARMN	Description of Planned Arm (N)	Num			
ADSL	ACTARM	Actual Arm	Char	\$200		

Mandatory variables section

3.5 Mandatory Variables

All datasets must include the following mandatory variables:

Variable Name	Variable Label	Variable Type	Display Format	Codelist/ Controlled Terms	Source/ Derivation
STUDYID	Study Identifier	Char	\$20		
SUBJID	Subject Identifier for the Study	Char	\$20		
USUBJID	Unique Subject Identifier	Char	\$30		
SITEID	Study Site Identifier	Char	\$30		
COUNTRY	Country	Char	\$3		
ARM	Description of Planned Arm	Char	\$200		
ARMN	Description of Planned Arm (N)	Num			
ACTARM	Actual Arm	Char	\$200		

2. Utility 1

Reads in the ADaM specification document and creates 2 types of datasets:

- Analysis variable metadata for all domains available in the specification
- Empty SAS datasets (Skeletons) with the same structure as the Final ADaM dataset for each domain

Analysis variable metadata dataset

VIEWTABLE: ANALYSIS_VARIABLE_METADATA										
DOMAIN	DATASET_LABEL	SORTVAR	LENGTH	VARNAM	LABEL	TYPE	FORMAT	MANDREL	VARNUM	
1	ADSL	Subject Level Analysis Dataset	USUBJID	\$20	STUDYID	Study Identifier	Char	\$20	Y	1
2	ADSL	Subject Level Analysis Dataset	USUBJID	\$20	SUBJID	Subject Identifier for the Study	Char	\$20	Y	2
3	ADSL	Subject Level Analysis Dataset	USUBJID	\$30	USUBJID	Unique Subject Identifier	Char	\$30	Y	3
4	ADSL	Subject Level Analysis Dataset	USUBJID	\$30	SITEID	Study Site Identifier	Char	\$30	Y	4
5	ADSL	Subject Level Analysis Dataset	USUBJID	\$3	COUNTRY	Country	Char	\$3	Y	5
6	ADSL	Subject Level Analysis Dataset	USUBJID	\$200	ARM	Description of Planned Arm	Char	\$200	Y	6
7	ADSL	Subject Level Analysis Dataset	USUBJID	8	ARMN	Description of Planned Arm (N)	Num		Y	7
8	ADSL	Subject Level Analysis Dataset	USUBJID	\$200	ACTARM	Actual Arm	Char	\$200	Y	8

Skeleton dataset ADSL

SAS - [VIEWTABLE: Subject Level Analysis Dataset]

Variable	Type	Length	Format	Label
STUDYID	Text	20	\$20.	Study Identifier
SUBJID	Text	20	\$20.	Subject Identifier for the Study
USUBJID	Text	30	\$30.	Unique Subject Identifier
SITEID	Text	30	\$30.	Study Site Identifier
COUNTRY	Text	3	\$3.	Country
ARM	Text	200	\$200.	Description of Planned Arm
ARMN	Number	8		Description of Planned Arm (N)
ACTARM	Text	200	\$200.	Actual Arm
ACTARMN	Number	8		Actual Arm (N)
ITTLFL	Text	1	\$1.	Intent-To-Treat Population Flag
PASPL	Text	1	\$1.	Full Analysis Set Population Flag

Skeleton dataset ADSL properties

SAS - [Standard:Adsl Properties]

Column Name	Type	Length	Format	Label
STUDYID	Text	20	\$20.	Study Identifier
SUBJID	Text	20	\$20.	Subject Identifier for the Study
USUBJID	Text	30	\$30.	Unique Subject Identifier
SITEID	Text	30	\$30.	Study Site Identifier
COUNTRY	Text	3	\$3.	Country
ARM	Text	200	\$200.	Description of Planned Arm
ARMN	Number	8		Description of Planned Arm (N)
ACTARM	Text	200	\$200.	Actual Arm
ACTARMN	Number	8		Actual Arm (N)
ITTLFL	Text	1	\$1.	Intent-To-Treat Population Flag
PASPL	Text	1	\$1.	Full Analysis Set Population Flag

3. Utility 2

Utility 2 (Skeleton to ADaM)

Utility 2 is called at the end of each ADaM dataset creation program:

- Appends last Work.dataset created in the program to the ADaM skeleton dataset (created by Utility 1) for the specified domain
- Sorts the data according to sorting key available in the Analysis variable metadata (created by Utility 1)
- Creates a discrepancy report

Calling Utility 2 (SAS code)

```
* work.ADSL ;
proc sql;
create table ADSL as
select a.*, case
when b.COMPLFL = '1' then 'N'
else b.COMPLFL
end as COMPLFL, b.DISCOT, b.REAS, b.REASPE, b.REASCAT,
RFKSTDT
as TRTSDT format = date9.,
from ADSL as a left join DS_OUT as b
on a.USUBJID = b.USUBJID
order by USUBJID
;
quit;

%UTILITY2 ;
%UTILITY2(inds = WORK.ADSL,
adsm_outds = ADAM.ADSL);

proc datasets solist;
delete _1 _2 ds_out vs/ newtype = DATA;
run;
quit;
```

Last Work.dataset

SAS - [Work:Adsl Properties]

Column Name	Type	Length	Format	Label
W_TMP_	Number	8	DATES9.	
W_TMP1_	Number	8	DATES9.	
W_T	Number	8		
W_WEIGHTBL	Number	8		
ITTLFL	Text	1		
SITEID	Text	30		
AGEU	Text	8		
ASEX	Text	6		
ARACE	Text	50		
AETHNIC	Text	22		
STUDYID	Text	20		
SUBJID	Text	20		
USUBJID	Text	30		

Work.dataset properties

SAS - [Work:Adsl Properties]

Column Name	Type	Length	Format	Label
W_TMP_	Number	8	DATES9.	
W_TMP1_	Number	8	DATES9.	
W_T	Number	8		
W_WEIGHTBL	Number	8		
ITTLFL	Text	1		
SITEID	Text	30		
AGEU	Text	8		
ASEX	Text	6		
ARACE	Text	50		
AETHNIC	Text	22		
STUDYID	Text	20		
SUBJID	Text	20		
USUBJID	Text	30		

Final ADSL dataset

SAS - [LE:Adsl Properties]

Study Identifier	Subject Identifier for the Study	Unique Subject Identifier	Study Site Identifier	Country	Description of Planned Arm	Description of Planned Arm (N)	Actual Arm	Actual Arm (N)	Intent-To-Treat Population Flag
1	FE0001234	8101F001	FE0001234_8101F001	8101	JPN	TaA	3	TaA	3 Y
2	FE0001234	8101F002	FE0001234_8101F002	8101	JPN	TaB	1	TaB	1 Y
3	FE0001234	8101F003	FE0001234_8101F003	8101	JPN	TaA	3	TaA	3 Y
4	FE0001234	8101F004	FE0001234_8101F004	8101	JPN	TaC	2	TaC	2 Y
5	FE0001234	8101F005	FE0001234_8101F005	8101	JPN	TaB	1	TaB	1 Y
6	FE0001234	8101F487	FE0001234_8101F487	8101	JPN	TaC	2	TaC	2 Y
7	FE0001234	8102F007	FE0001234_8102F007	8102	JPN	TaC	3	TaC	3 Y
8	FE0001234	8102F008	FE0001234_8102F008	8102	JPN	TaA	3	TaA	3 Y
9	FE0001234	8102F009	FE0001234_8102F009	8102	JPN	TaB	1	TaB	1 Y
10	FE0001234	8102F010	FE0001234_8102F010	8102	JPN	TaB	1	TaB	1 Y

Final ADSL properties

SAS - [Adam:Adsl Properties]

Column Name	Type	Length	Format	Label
STUDYID	Text	20	\$20.	Study Identifier
SUBJID	Text	20	\$20.	Subject Identifier for the Study
USUBJID	Text	30	\$30.	Unique Subject Identifier
SITEID	Text	30	\$30.	Study Site Identifier
COUNTRY	Text	3	\$3.	Country
ARM	Text	200	\$200.	Description of Planned Arm
ARMN	Number	8		Description of Planned Arm (N)
ACTARM	Text	200	\$200.	Actual Arm
ACTARMN	Number	8		Actual Arm (N)
ITTLFL	Text	1	\$1.	Intent-To-Treat Population Flag
PASPL	Text	1	\$1.	Full Analysis Set Population Flag
PASPL	Text	1	\$1.	Safety Population Flag

4. Discrepancy report

Utility 2 creates a HTML report if:

- Error in structure
 - a) Miss-match in variable type
 - b) Length of the variable in the work dataset exceeds that of the corresponding variable in the skeleton

→ Output data will NOT be created

- Variables existing in the skeleton, but not in the work dataset
- Can occur during development, but not for the final data

Discrepancy report

Output data adam.adae not created due to the following critical errors in data structure

Variable	Type	Length	Type	Length	Structure Error
AESEV	C	8	C	10	Length of variable in input data exceed that of the skeleton. Variable content might be truncated. Please confirm and re-specification

Variable(s) listed below was not found in the input data work.adae_m. Do add it them to your input data, or remove it them from the specification.

Variable	Type	Length
STUDYID	C	20
USUBJID	C	30

5. When to run the utilities

- **Utility 1** is run only when there is a change in the ADaM specification document
- **Utility 2** is run every time the ADaM dataset creation program is run

Conclusions

- Programmatically using the ADaM specification
- No need to remember which variables to keep or drop
- Ease of governing the sequence of variables
- Ease of maintaining the standards
- All the variables and their attributes are listed in SAS dataset containing the Analysis variable metadata
- Easy to check for inconsistencies and make necessary corrections
- Can be used by the sponsor when checking for structural inconsistencies in the ADaM datasets delivered by a CRO