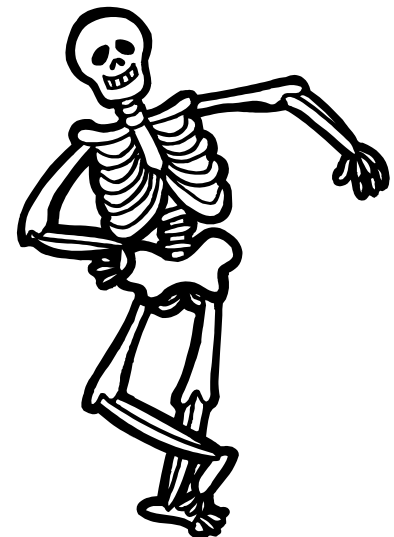




Ferring Pharmaceuticals

Getting your skeletons out of the closet



Parag Naithani

PhUSE SDE Copenhagen, May 28 2013

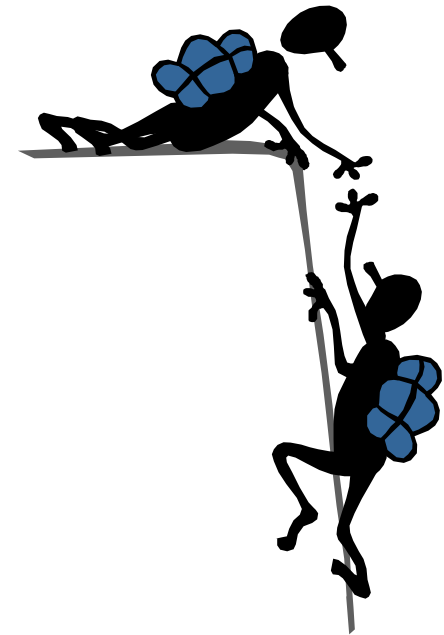
Purpose



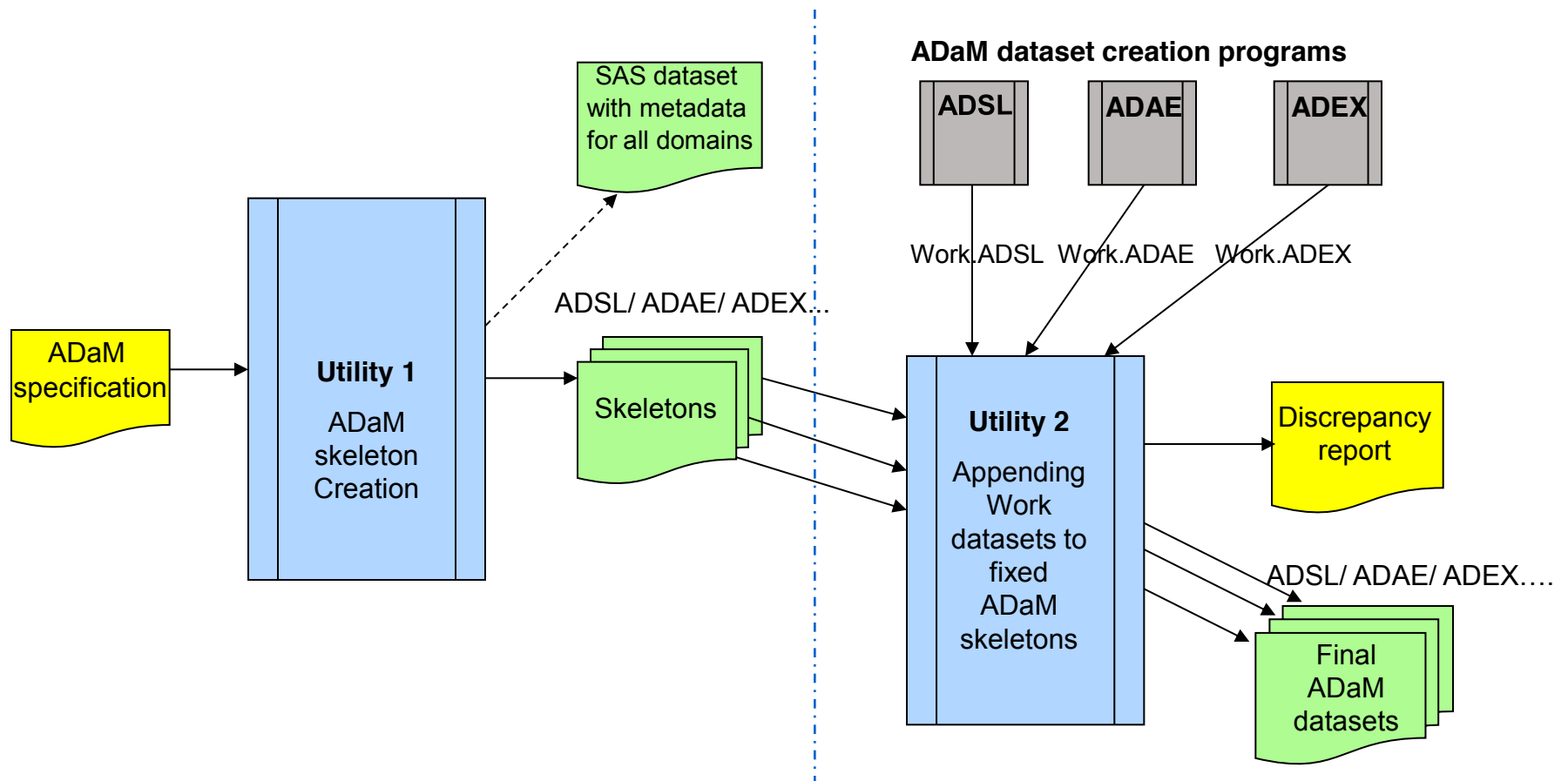
- Strict Analysis Data Model Implementation Guideline (ADaMIG) specifying ADaM standard
- Need for a Programming solution
- Programmatically use the ADaM specification document to shape the structure of the final ADaM datasets

Challenges

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



Process overview diagram



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 randomised	USUBJID	ADSL	SAP

Analysis Variable Metadata for ADSL

Dataset Name	Variable Name	Variable Label	Variable Type	Display Format	<u>Codelist/</u> Controlled Terms	Source/ Derivation
ADSL	STUDYID	Study Identifier	Char	\$20		
ADSL	SUBJID	Subject Identifier for the Study	Char	\$20		
ADSL	USUBJID	Unique Subject Identifier	Char	\$30		
ADSL	SCREENNO	Screening Number	Char	\$20		
ADSL	FASFL	Full Analysis Set Population Flag	Char	\$1	Y, N	

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		

Utility 1 (Spec to Skeleton)

- Reads in the ADaM specification document
- Creates 2 types of datasets:
 - ➔ Analysis variable metadata for all domains available in the specification
 - ➔ Empty SAS datasets (Skeletons) with same structure as Final ADaM dataset for each domain



Analysis variable metadata dataset

VIEWTABLE: ANALYSIS_VARIABLE_METADATA										
	DOMAIN	DATASET_LABEL	SORTVAR	LENGTH	VARNAM	LABEL	TYPE	FORMAT	MANDFL	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	COUNTR	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

The screenshot shows the SAS software interface. The title bar reads "SAS - [VIEWTABLE: Subject Level Analysis Dataset]". Below the title bar is a menu bar with the following options: File, Edit, View, Tools, Data, Solutions, Window, and Help. A toolbar is located below the menu bar, containing various icons for file operations (such as Open, Save, Print) and data manipulation (such as Sort, Filter, and View). The main window displays a table with the following columns:

	Study Identifier	Subject Identifier for the Study	Unique Subject Identifier	Study Site Identifier	Country	Description of Planned Arm	Description of Planned Arm (N)

On the left side of the window, there is a "Contents of" pane showing a "Library" icon and a "File Short" icon.

Skeleton dataset ADSL properties

SAS - [Standard.Adsl Properties]

File Edit View Tools Solutions Window Help

Find column name:

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)
 ITTFL	Text	1	\$1.	Intent-To-Treat Population Flag
 FASFL	Text	1	\$1.	Full Analysis Set Population Flag

Explorer

Contents of 'Stan'

Adex

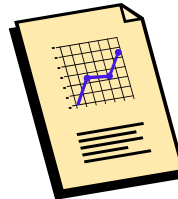
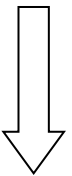
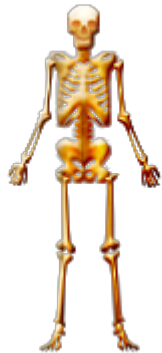
Adsl

Skeleton

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 an discrepancy report



Calling Utility 2

```
* work.ADSL ;
proc sql;
  create table ADSL as
  select a.*, case
    when b.COMPLFL = '' then 'N'
    else b.COMPLFL
  end as COMPLFL, b.DISCDT, b.REAS, b.REASSPE, b.REASCAT,
    RFXSTDT          as TRTSDT  format = date9. , RFXENDT) as TRTEDT  format = date9.
  from _ADSL as a left join  DS_OUT as b
  on a.USUBJID = b.USUBJID
  order by USUBJID
  ;
quit;

*UTILITY 2 ;
%Utility2(inds = WORK.ADSL,
          adam_outds  = ADAM.ADSL);

proc datasets nolist;
  delete _ : __ : ds_out vs/ memtype = DATA;
run;
quit;
```

Last Work.dataset

3LE: Work.AdsI]													
Tools Data Solutions Window Help													
                 													
	TMP	_TMP1_	I	WEIGHTBL	ITTFI	SITEID	AGEU	ASEX	ARACE	AETHNIC	STUDYID	Subject Identifier for the Study	USUBJID
1	12NOV2012	28JAN2013	1	63	Y	8131	YEARS	Female	Asian	Not Hispanic or Latino	FEXXXX 1234	8131F181	FEXXXX1234_8131F181
2	12OCT2012	06JAN2013	2	49.5	Y	8102	YEARS	Female	Asian	Not Hispanic or Latino	FEXXXX 1234	8102F007	FEXXXX1234_8102F007
3	09NOV2012	07FEB2013	3	50.1	Y	8103	YEARS	Female	Asian	Not Hispanic or Latino	FEXXXX 1234	8103F014	FEXXXX1234_8103F014
4	26OCT2012	15JAN2013	4	49.2	Y	8105	YEARS	Female	Asian	Not Hispanic or Latino	FEXXXX 1234	8105F025	FEXXXX1234_8105F025
5	28NOV2012	19FEB2013	5	47.3	Y	8124	YEARS	Female	Asian	Not Hispanic or Latino	FEXXXX 1234	8124F141	FEXXXX1234_8124F141

Work.dataset properties

SAS - [Work.Ads1 Properties]

File View Tools Solutions Window Help

Explorer
Contents of 'Work'

- Ads1
- Analysis_va...
- Formats
- Sasmacr

General Details Columns Indexes Integrity Passwords

Find column name:

Column Name	Type	Length	Format	Label
12% _TMP_	Number	8	DATE9.	
12% _TMP1_	Number	8	DATE9.	
12% I	Number	8		
12% WEIGHTBL	Number	8		
A% ITTFL	Text	1		
A% SITEID	Text	30		
A% AGEU	Text	8		
A% ASEX	Text	6		
A% ARACE	Text	50		
A% AETHNIC	Text	22		
A% STUDYID	Text	20		
A% SUBJID	Text	20		Subject Identifier for the Study
A% USUBJID	Text	30		

Final ADSL dataset

[LE: Adam.AdsI]										
Tools Data Solutions Window Help										
                  										
	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	FEXXXX 1234	8101F001	FEXXXX1234_8101F001	8101	JPN	Trt A	3	Trt A	3	Y
2	FEXXXX 1234	8101F002	FEXXXX1234_8101F002	8101	JPN	Trt B	1	Trt B	1	Y
3	FEXXXX 1234	8101F003	FEXXXX1234_8101F003	8101	JPN	Trt A	3	Trt A	3	Y
4	FEXXXX 1234	8101F004	FEXXXX1234_8101F004	8101	JPN	Trt C	2	Trt C	2	Y
5	FEXXXX 1234	8101F005	FEXXXX1234_8101F005	8101	JPN	Trt B	1	Trt B	1	Y
6	FEXXXX 1234	8101F487	FEXXXX1234_8101F487	8101	JPN	Trt C	2	Trt C	2	Y
7	FEXXXX 1234	8102F007	FEXXXX1234_8102F007	8102	JPN	Trt C	2	Trt C	2	Y
8	FEXXXX 1234	8102F008	FEXXXX1234_8102F008	8102	JPN	Trt A	3	Trt A	3	Y
9	FEXXXX 1234	8102F009	FEXXXX1234_8102F009	8102	JPN	Trt B	1	Trt B	1	Y
10	FEXXXX 1234	8102F010	FEXXXX1234_8102F010	8102	JPN	Trt B	1	Trt B	1	Y
11	FEXXXX 1234	8102F011	FEXXXX1234_8102F011	8102	JPN	Trt A	3	Trt A	3	Y

Final ADSL properties

SAS - [Adam.Adsl Properties]

File View Tools Solutions Window Help

Explorer

Contents of 'Adam'

Adsl

Find column name:

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)
ITTFL	Text	1	\$1.	Intent-To-Treat Population Flag
FASFL	Text	1	\$1.	Full Analysis Set Population Flag
SAFFL	Text	1	\$1.	Safety Population Flag

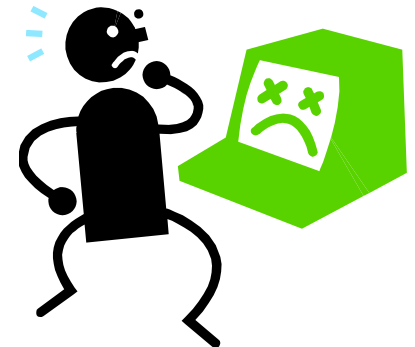
Discrepancy report

Utility 2 creates a HTML report [if](#):

1. 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

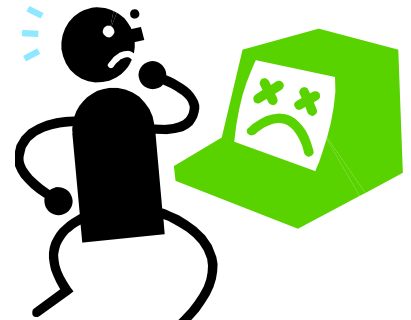


Discrepancy report , continued

Utility 2 creates a HTML report [if](#):

2. 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

	Skeleton		Input data		
	standard.adae		work.adae_m		
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 the 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.

Skeleton		
standard.adae		
Variable	Type	Length
STUDYID	C	20
USUBJID	C	30

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



Conclusion



- Programmatically using the ADaM specification document
- No need to remember which variables to keep or drop
- Ease of governing the sequence of variables
- Ease of maintaining the standards

Conclusion, continued



- 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
- Used by the sponsor for checking the structural inconsistencies in the ADaM datasets delivered by the CRO

Questions?



Thank you!