

AUTOMATION OF STANDARD ADAM VARIABLES THROUGH AN R PACKAGE – GenAdam

Mahesh Divakaran (Data Analyst II)

(mahesh.divakaran@genproresearch.com)

Mohin Mohanan (Data Analyst II)

(mohin.mohanan@genproresearch.com)

Denny Sebastian (Data Analyst I)

(denny.sebastian@genproresearch.com)

Agenda

R in Clinical Research

ADaM: ADSL and BDS Background

Adoption of R for ADaM Datasets

GenAdam

Examples of GenAdam Output

R in Clinical Research



Open Source



Integration



Wide Library
Support



Statistical
Computing and
Libraries



Data Visualization



Interactive Apps
(Shiny)



Great Community
Support



IN GENERAL BIOSCIENCE AND
ACADEMIA, R HAS BUILT OVER
YEARS ITS POSITION OF ONE
OF THE INDUSTRY STANDARDS



R IS INCREDIBLY FLEXIBLE



THE COLLECTION OF
PACKAGES NOW MAKE IT VERY
EASY TO MANIPULATE DATA
FRAMES AND TO PERFORM
TRANSFORMING OPERATIONS



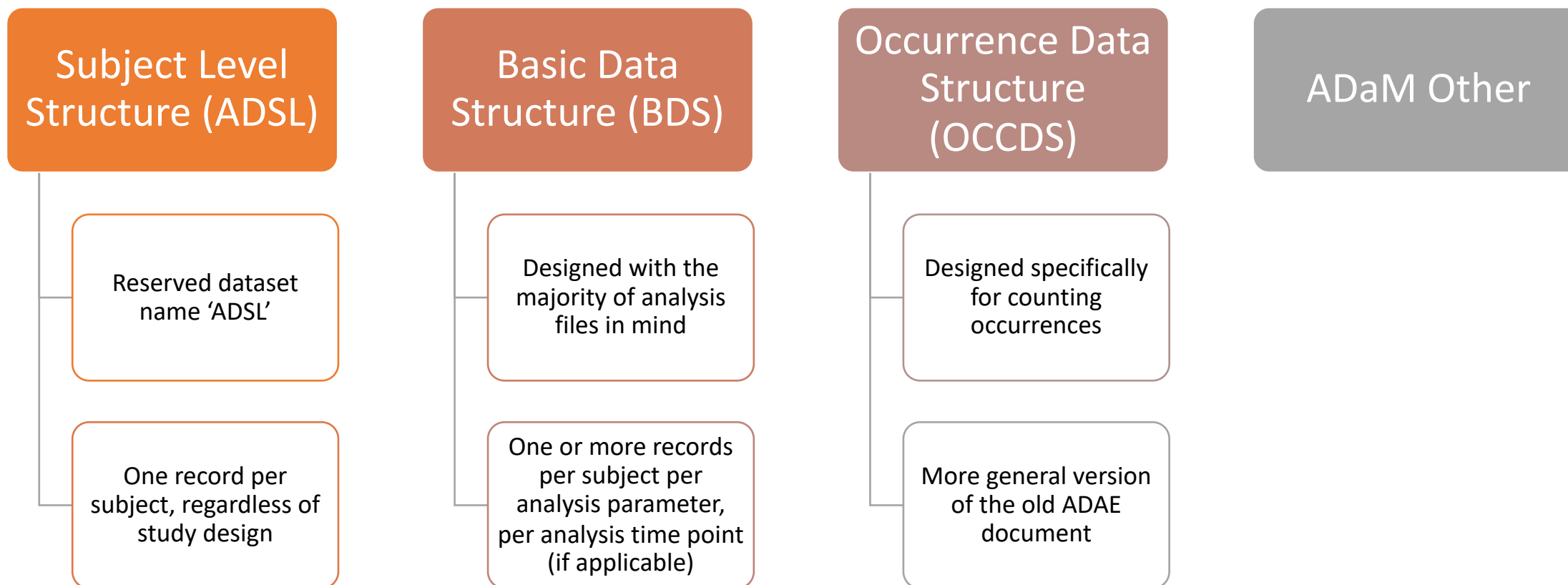
R CAN BE AN EXTREMELY
POWERFUL TOOL TO CREATE
TABLES, LISTINGS AND GRAPHS



PHARMACEUTICAL
COMPANIES, CROS AND EVEN
FDA DO USE R "INTERNALLY".
BUT THEY RESIST (OR
HESITATE) TO USE IT IN
SUBMISSIONS (TO FDA).

ADaM: ADSL and BDS Background

Analysis Data Model (ADaM) : provide a framework to statistical programmers and statisticians that enables analysis of the data



Adoption of R for ADaM Datasets

Commonly R is utilized for exploratory analysis, routine data reading, and interpretation tasks but not for creating TFL or CDISC datasets. With the extensive library support, we can quickly do the data wrangling with minimal code. We have the following benefits and challenges while creating datasets using R.



Benefits

- Free and open-source
- Reproducible research
- Extremely easy data wrangling
- Can be easily extended
- Less code
- Community support



Challenges

- Validation
- Packages do not follow a typical SDLC
- Legacy data
- R support

GenAdam

R package

Effective alternative in creating ADaM
datasets

Single function call

Visualizing dataset

Download .sas7bdat

GenAdam setup

Installation

- `install.packages("<source_package_path>/GenAdam_0.1.0.tar.gz", repos = NULL, type = "source")`

List of Package Dependencies

- `data.table`
- `shiny`
- `shinythemes`
- `shinyjs`
- `tidyverse`
- `ggplot2`
- `haven`
- `shinyWidgets`
- `DT`
- `reshape2`
- `openxlsx`
- `tidyr`

GenAdam - Functions

We have 4 functions to create the 4 ADaM datasets in the first phase:

ADSL

ADVS

ADLB

ADQS

The function call for ADSL dataset creation:: ***adsl(folder_path, varlist, metadata)***

- Here the function takes three input arguments;
 - folder_path : Folder path of input datasets (Required)
 - varlist : List of variables to be output (Optional)
 - metadata : Metadata/ Codelist filename (Optional)

Similarly we have other functions as well,

- ***advs(folder_path, advs_var, metadata)***
- ***adlb(folder_path, adlb_var, metadata)***
- ***adqs(folder_path, adqs_var, metadata)***

5 Main Functions of GenAdam

sdtm_data

The sdtm_data function displays the dataset in R shiny. We will be able to filter, search and explore the data in the interactive visualization feature of R shiny and datatable (DT).

```
sdtm_data(folder_path,file_name)
```

adsl

The adsl() function creates the AD_SL dataset with the standard variables. This function will work only with single-period simple to moderate studies, which is a package limitation.

adv

The advs() function creates the ADVS dataset with the standard variables and predecessor variables from AD_SL. Here we develop a list of standard variables associated with the ADVS dataset.

adlb

The adlb() function creates the ADLB dataset with the standard variables and predecessor variables from AD_SL. Here we develop a list of standard variables associated with the ADLB dataset.

adqs

The adqs() function creates the ADQS dataset with the standard variables and predecessor variables from AD_SL. Here we develop a list of standard variables associated with the ADQS dataset.

Example output of ADSL

```
adsl("<folder_path>", "metadata.xlsx")
```

GenAdam

Showing Dataset: ADSL

[Download](#)

Show entries

Search:

	STUDYID	SUBJID	USUBJID	AGE	AGEU	AGEGR1	AGEGR2	SEX	RACE	RANFL	ARM	ACTARM	TRTSDT
1	TRID-02	801-010	TRID-02-801-010	50	YEARS	45-65	At least 35 y old	M	ASIAN	N			
2	TRID-02	801-019	TRID-02-801-019	37	YEARS	18-45	At least 35 y old	M	BLACK OR AFRICAN AMERICAN	Y			2020-10-02
3	TRID-02	801-020	TRID-02-801-020	44	YEARS	18-45	At least 35 y old	F	WHITE	N			
4	TRID-02	801-021	TRID-02-801-021	48	YEARS	45-65	At least 35 y old	M	BLACK OR AFRICAN AMERICAN	N			
5	TRID-02	801-025	TRID-02-801-025	50	YEARS	45-65	At least 35 y old	F	ASIAN	N			
6	TRID-02	801-030	TRID-02-801-030	59	YEARS	45-65	At least 35 y old	M	WHITE	Y	EM-7.5mg BID	EM-7.5mg BID	2020-08-20

Example output of *ADVS*

```
advs("<folder_path>", "metadata.xlsx")
```

GenAdam

Showing Dataset: ADVS

Download

Show 20 entries

Search:

	STUDYID	SUBJID	USUBJID	ADT	ATM	ADTM	PARAMCD	PARAM	ABLFL	BASE	CHG	PCHG	TRT01P	TRT01PN
							/					All		
41	TRID-02	801-019	TRID-02-801-019	2020-10-02	11:51	2020-10-02T11:51	PULSE	Pulse Rate		68	7	10.2941176470588		
42	TRID-02	801-019	TRID-02-801-019	2020-10-02	14:51	2020-10-02T14:51	PULSE	Pulse Rate		68	-4	-5.88235294117647		
43	TRID-02	801-019	TRID-02-801-019	2020-10-02	14:51	2020-10-02T14:51	PULSE	Pulse Rate		68	10	14.7058823529412		
44	TRID-02	801-019	TRID-02-801-019	2020-09-18	12:17	2020-09-18T12:17	PULSE	Pulse Rate		68	-14	-20.5882352941176		
45	TRID-02	801-019	TRID-02-801-019	2020-10-01		2020-10-01	PULSE	Pulse Rate		68	-3	-4.41176470588235		
46	TRID-02	801-019	TRID-02-801-019	2020-10-01	13:26	2020-10-01T13:26	PULSE	Pulse Rate		68	-6	-8.82352941176471		
47	TRID-02	801-019	TRID-02-801-019	2020-10-01	13:40	2020-10-01T13:40	PULSE	Pulse Rate		68	-3	-4.41176470588235		
48	TRID-02	801-019	TRID-02-801-019	2020-10-01	13:54	2020-10-01T13:54	PULSE	Pulse Rate	Y	68	0	0		
49	TRID-02	801-019	TRID-02-801-019	2020-10-02	11:51	2020-10-02T11:51	PULSE	Pulse Rate		68	-3	-4.41176470588235		



THANK YOU

Any Questions ?