



2023
EUROPE
INTERCHANGE
COPENHAGEN | 26-27 APRIL



Analysis Results Standards Guidelines and Implementation using R shiny

Jayashree Vedanayagam, Applied Statistician, Data Science, Princeps Technologies Inc
Smriti Anand, Statistical Programmer, GPD-SDSA, Pfizer

Meet the Speakers



Jayashree Vedanayagam

Title: Applied Statistician

Organization: Princeps Technologies Inc

Area of expertise is ADaM & TLF Clinical Analysis and Programming. Worked on TAs like Arthritis, Influenza, Lung Cancer; SME in LB, PK/PD & IS domains. Currently working on developing standard solutions based on open-source technologies for Clinical Analysis and regulatory submissions.



Smriti Anand

Title: Statistical Programmer

Organization: Pfizer Healthcare

I'm a Statistical Programmer at Pfizer. Currently focused on using open-source technologies to add value to Clinical Analysis and Submissions. I'm relatively new to the field, but I like innovating new solutions to existing problems and I'm usually trying to find a way to use R in everything.



Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the authors and do not necessarily reflect the official policy or position of CDISC.*
- *All views and opinions on this presentation are those of the authors, and do not necessarily reflect the views and opinions of the organizations they represent.*
- *The author(s) have no real or apparent conflicts of interest to report.*



Agenda

1. Introduction to Analysis Results Standards
2. Inputs, Data Standards and Analysis
3. Introduction to R and R Shiny
4. A&R Application Operations
5. Additional Features
6. Testing & Advantages of Shiny App
7. Future Enhancements and Conclusion



Introduction to Analysis Results standards

Focus Point for Analysis Results Standards

Deriving Rules from Available CDISC Standards

Common analysis results



Focus Point for Analysis Results Standards

1. The aim of creating analysis results standards is to enable identifying and automating the link between the report and the input data.
2. Categorization of report to TA or related safety & efficacy clinical endpoints is essential
3. Trace of endpoint to standards analysis data structure and in turn to the analysis reports enables reproducibility.
4. Classification of Reports based on Data Domains, Data structures, Summary Structure, Statistical Analysis, Graphical analysis
5. Identification of Treatment variable, Population variable, Grouping variable, Subset Variable, Analysis Variable



Deriving Rules from Available CDISC Standards

- SDTM and ADaM Standards
 - Mapping results to SDTM domain Models
 - Mapping results to ADaM data structures
 - Annotation of report templates to variable type and Category
- TAUGs
 - Endpoint and analysis reports
 - Required Collection Variables
 - Domain Details
 - Assessments



Common Analysis Results

1. Some analysis results discussed in CDISC are for Adverse Event and PK/PD
2. Safety domains analysis results
3. Generic summary and graphical analysis
4. Efficacy domain analysis results
5. Frequent TA specific analysis results



Inputs, Data Standards and Analysis

Prerequisite CDISC Standards

Prerequisite Reporting Standards



Prerequisite CDISC Standards

Following are the prerequisite for the Implementation & Generation of Reports using the standardized application

1. The Input data should follow the SDTM / ADaM IG standards
 1. Labels
 2. Variable Naming convention
 3. Code list
 4. Data types & Formats
 5. Derivation rules
2. Clear Classification of TA and Domains

Prerequisite Reporting Standards

1. Summary Table Prerequisite

1. Grouping Columns: Subgroup Variables
2. Grouping Rows: By Group Variables
3. Variable types vs Statistics:
 1. Categorical - Frequency/Percentages
 2. Continuous – Descriptive statistics

2. Graphical Prerequisite

1. Type of Graph vs Type of analysis
 1. Bar/line – discrete or descriptive analysis
 2. Step Plot / Histogram – Time point / duration-based analysis / Trend analysis
 3. Box Plot/Error Bar – Descriptive analysis
 4. Scatter Plot – Individual/Actual datapoint analysis

3. Domain Specific Prerequisite

1. ADLB – eDISH – ALT, AST, BILI ; Flag Variable; Analysis Variables
2. ADAE – SOC, Term, SMQ/FMQ variables, Flag Variables, Grouping Variables
3. ADDS – Epoch, DS phase variables, Reason Variable

4. Statistics Specific Prerequisite

1. Ratio – pair wise analysis



Introduction to R and R shiny

Using R to build reports

Harnessing the power of R Shiny Dashboard

Program Flow

Using R to build reports

Being open-source, R can be used widely without license restrictions, and hands access to a lot of packages suited for reporting purposes. Some that we used are:

- Powerful data visualisations – ggplot2, plotly, cowplot, etc
- Data wrangling and tidying with tidyverse suite of packages.
- Flextable, Rmarkdown for formatted and clear outputs
- Epitools, survival, for statistical analysis
- Shiny, shinywidgets for dashboard

R functions and data can be packaged, which makes any project or application shareable and has potential for collaboration.

The program behind this application was built on several standardised functions as blocks



R Shiny Dashboard

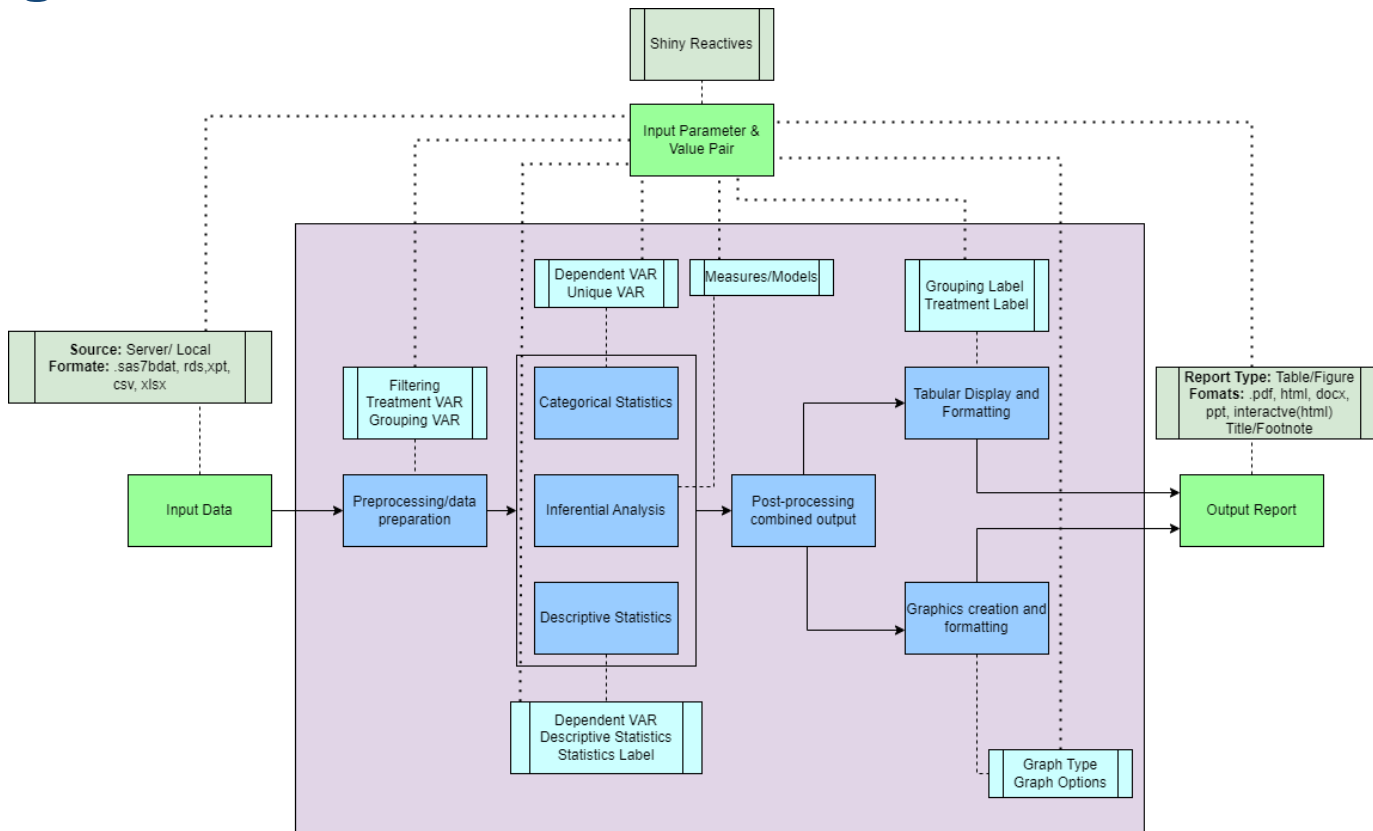
All the R programs in this tool run in a R Shiny application.

- ‘Shiny’ is a package within R that lets you build interactive dashboards, to dynamically process user inputs *without dealing with the code at all*.
- Building a web-based shiny dashboard only requires knowledge of R and is not reliant on expertise in HTML, Javascript and other Web tools.
- Hosted on R studio Connect or other options at Enterprise level; removing dependency on R installation.
- Shiny is built on reactive expressions – which take user inputs and process the code in real time.

Shiny is easily scalable and distributable – it can greatly speed up decision making by giving the end user direct access to all the data visualizations and interpretation.



Program Flow:





A&R App Operation

Dashboard User Interface

Available Reports

Demonstration:

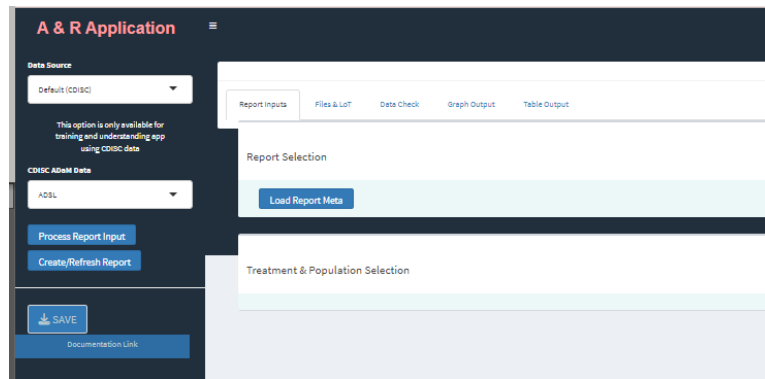
- Step Plot

- ADSL Summary Table and Panel

Dashboard User Interface

Report Creation:

- ADaM dataset sourced from Server or CDISC example data
- Glimpse of dataframe (Data Check)
- Data-driven Report Selection from List of tables (LoT)
- Report Inputs tab - User input given according to selected report
- Graph Output, Table Output tabs to display the result and make formatting changes if required.

A detailed view of the 'Report Selection' form. It features a 'Load Report Meta' button at the top left. Below it is a 'Therapeutic Area' dropdown menu with 'Any' selected. To the right is a 'Report Name' dropdown menu with 'Generic' selected. Below the 'Therapeutic Area' is a 'Report Type' dropdown menu with 'Table' selected, and a list of options: 'Table' and 'Figure'. To the right of the 'Report Type' is a 'Report Number' dropdown menu with '0.1' selected. At the bottom right is a 'Report Description' dropdown menu with 'Generic' selected.

Current Available Reports

Summary Tables:

- Generic Summary Tables:
 - Categorical Analysis
 - Summary/Numeric Analysis
- Demographics (ADSL) Table
- Adverse Events (ADAE) Summary and Risk Statistics
- Disposition/Discontinuation Summary (ADDS)

Visualizations:

- Generic Graphs:
 - Bar
 - Line Plot
 - Step Plot
 - Box Plot
- Demographics (ADSL) panel
- Forest Plot (ADAE) of Risk Difference/Ratio
- Volcano Plot (ADAE) of Risk Difference/Ratio
- Event Analysis (ADAE)
- eDISH (ADLB) Plot

Example – Generic Step Plot of Event Start Date

Report Selection

Load Report Help

Therapeutic Area: Any

Report Name: Generic

Report Type: Figure

Report Number: 0.1

Type of Graph: STEP

Report Description: Generic

Treatment & Population Selection

Treatment Variable: TRTA

Treatment Sort Variable: TRTA

Population Filter: Overall Population

Identify Fields

Subsetting Condition:

Sub-Group Variables:

By-Group Variables: AEDECOD

Display Treatment Total? + - *

Display Subgroup Total? + - *

Display Treatment % Count? *

Summary Statistics Fields

Dependent/Analysis Variable: ASTDT

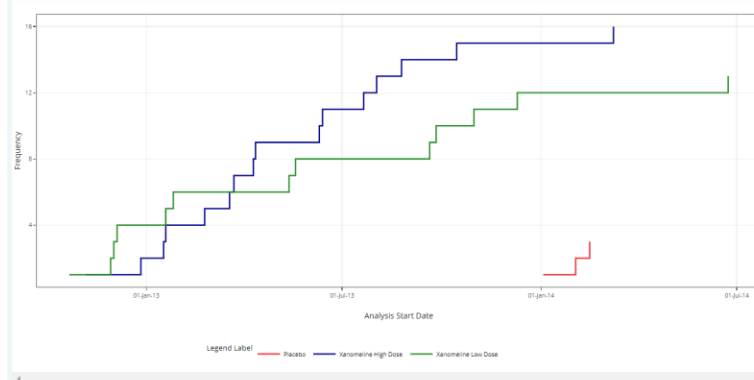
Statistics Displayed: Cumulative Count

Report Fields

Statistics Label:

By-Group Label:

Figure 3.0
Adverse Event Analysis: Summary of Analysis Start Date by Dictionary-Derived Term for Overall Population
AEDECOD = APPLICATION SITE ERYTHEMA



n is the number of subjects with non-missing ASTDT
PRIZER CONFIDENTIAL Table Generation: 10Apr2023 (2204)Output File: Generic

AEDECOD

APPLICATION SITE ERYTHEMA

Example – Demographics Panel Plot and Summary

Report Selection

Load Report Meta

Therapeutic Area
Any

Report Name
adsl_001

Report Type
Table

Report Number
1.1

Report Description
Summary report of Demography data

Treatment & Population Selection

Treatment Variable
TRToLA

Treatment Sort Variable
TRToLAN

Population Filter
Overall Population

Subsetting Condition
USBJID='1'

Sub Group Variables

Descriptive Statistics
N,Range,Meansd,Median,JQR

Display Treatment Total?
 Y N

Analysis Variables
AGEGR1,AGE-S,SEX,RACE,ETHNIC

By Group Variables

Descriptive Statistics Label
N,Range,Mean (SD),Median,Interquartile Range

Display Subgroup Total?
 Y N

Analysis Variables Label
Age Group,Age,Sex,Race,Ethnicity

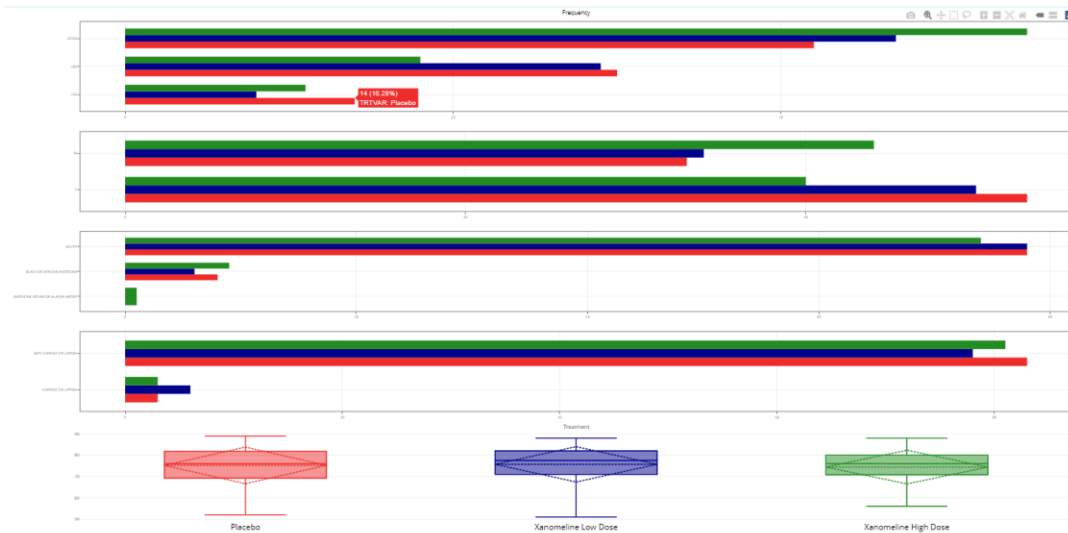
By Variable Label

Treatment Legend Color
firebrick2,blue4,forestgreen,gold,magenta3,aquamarine1

Display Treatment N Count?
 Y N

Example – Demographics Panel Plot and Summary

		Placebo (N=66)	Xanomeline Low Dose (N=64)	Xanomeline High Dose (N=64)	Total (N=194)
Age Group	<65	14 (6.23%)	8 (9.52%)	11 (13.10%)	33 (12.99%)
	65-80	42 (63.84%)	47 (55.94%)	55 (65.48%)	144 (54.60%)
	>80	30 (44.88%)	29 (34.52%)	18 (21.43%)	77 (29.31%)
Age	N	66	64	64	194
	Range	(52.0,89.0)	(51.0,88.0)	(54.0,88.0)	(51.0,89.0)
	Mean (SD)	75.2 (8.59)	75.7 (8.29)	74.4 (7.99)	75.1 (8.25)
	Median	76.0	77.5	76.0	77.0
	Interquartile Range	12.5	11.0	9.2	11.0
Sex	F	53 (81.63%)	50 (59.21%)	40 (47.62%)	143 (56.30%)
	M	33 (50.37%)	34 (40.48%)	44 (52.38%)	111 (43.70%)
Race	WHITE	78 (96.70%)	78 (92.88%)	74 (88.10%)	230 (90.55%)
	BLACK OR AFRICAN AMERICAN	8 (9.30%)	6 (7.14%)	9 (10.71%)	23 (9.06%)
	AMERICAN INDIAN OR ALASKA NATIVE	0	0	1 (1.18%)	1 (0.39%)
Ethnicity	HISPANIC OR LATINO	3 (3.49%)	6 (7.14%)	3 (3.57%)	12 (4.72%)
	NOT HISPANIC OR LATINO	81 (96.51%)	78 (92.84%)	81 (94.43%)	242 (95.28%)





Additional Features

Report metadata

User Guide

Saving and Bundling Reports

Input Capture

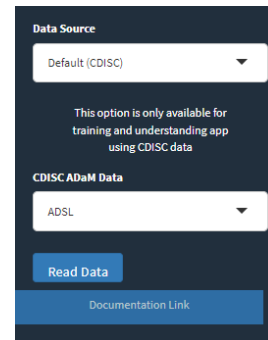
Report Metadata

The app uses a metadata file which contains all the reports corresponding to their domains, with standard report names. This can be modified on a study-wise basis by users, thus giving control of what reports are required in a particular submission.

This allows the process of report selection to be data driven as well.

User Guide

Application User Guide is a paged HTML document created using Markdown and pagedown directly linked in the application's sidebar.

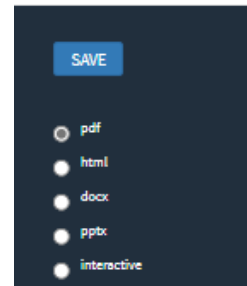


The screenshot shows a dark-themed user interface. At the top, there is a section titled "Data Source" with a dropdown menu currently set to "Default (CDISC)". Below this, a message states: "This option is only available for training and understanding app using CDISC data". Underneath, there is another section titled "CDISCADaM Data" with a dropdown menu set to "ADSL". At the bottom of the form, there is a blue button labeled "Read Data" and a link labeled "Documentation Link".

Saving and Bundling Reports

All outputs created in the application can be saved to corresponding Server location or downloaded locally, in various formats.

The application also allows the user to automatically bundle selected reports into one combined file in HTML, PDF or Word format, which can then be saved or downloaded.



Capturing Inputs and Outputs

From the app, the user can save all the inputs, outputs as well as the final processed datasets for a certain report. This has a lot of benefit in Quality Check process, and it allows for reproducibility of the same result by different users.



Testing & Advantages of Shiny App

How to test Shiny app

Advantages of using Shiny Application



How to test Shiny app

1. Shiny Test
 - a) Unit Tests
 - b) Server Function Tests
 - c) Snapshot-based test
2. Shiny Load Test
 - a) Record
 - b) Replay
 - c) Analyze
3. UAT Testing
 - a) Comprehensive UAT testing by study programmers/ statisticians
 - b) Production study data



Advantages of using Shiny Application

1. Reuse or Repurpose
2. Reproducible
3. Shared as Package
4. Published locally or globally
5. Less Maintenance
6. User Friendly
7. Programming Friendly
8. Embed Data
9. Scalable



Future Enhancements and Conclusion



Future Enhancements

1. Animated Plot & 3D plot
2. Application R Package
3. Shiny dashboard with data for eSub
4. Efficacy and TA specific analysis
5. OCCDS and BDS Generic Templates



Conclusion

- This application is aimed at providing standardized TLF guidance and implementation platform via an interactive interface that enables the user to generate standard results for regulatory submissions or in-house data monitoring
- This interactive dashboard can become an option for future regulatory submissions – since we can embed study data and metadata into the Shiny application. Having more benefits in areas like real world data, data monitoring, decision making and sharing trial success in public domains makes this a key avenue to explore further.



Thank You ! Questions?

