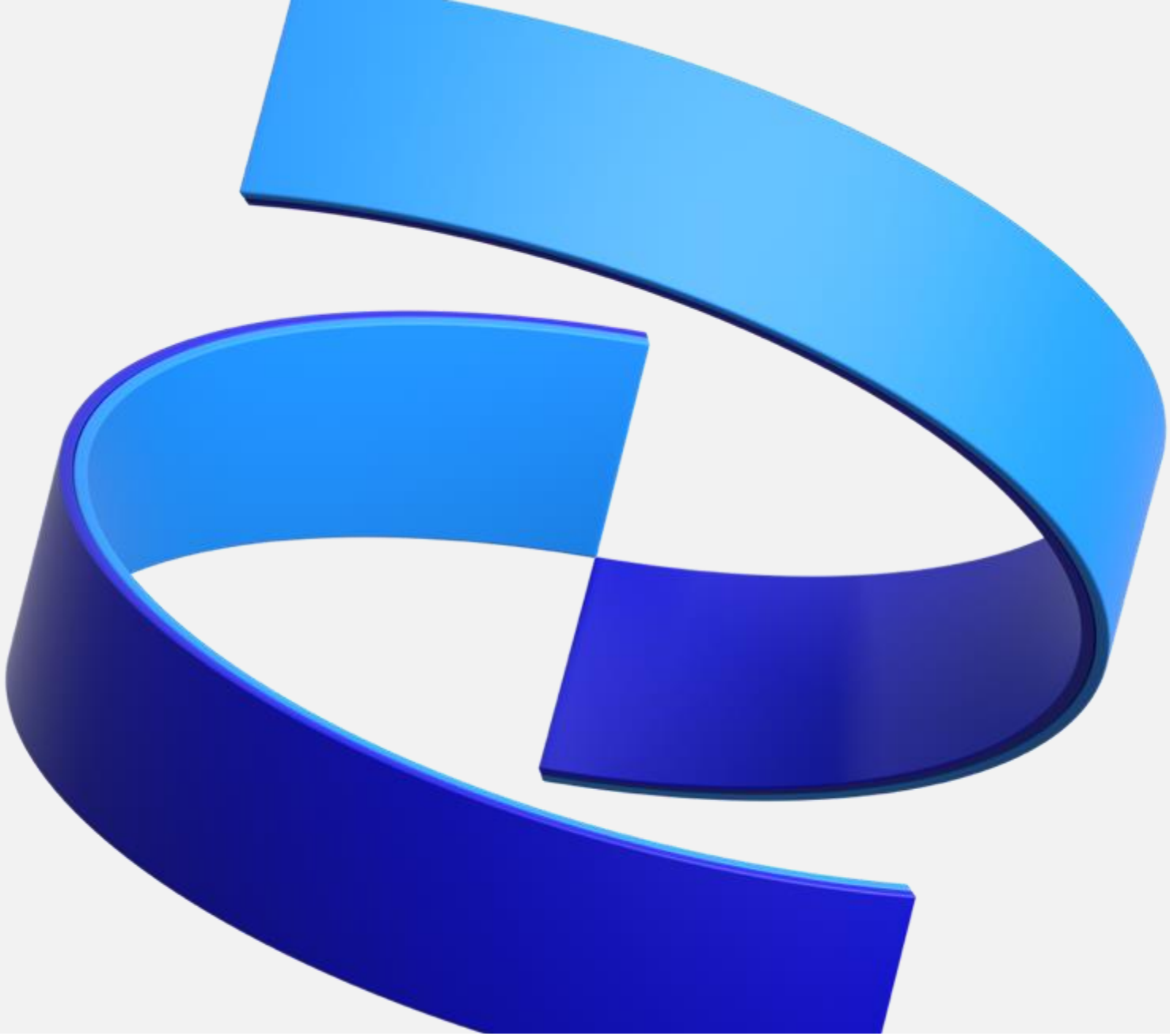
{carver}
Clinical Analysis Report and Visualization Ensemble in R



Smriti Anand
Sukalpo Saha

24-Sep-2024

Background

Purpose

Development of an open-source tool and package to enable generation of identified reports and figures for clinical review and direct inclusion in submissions to regulatory agencies

Initial Scope

To develop a package to generate interactive forest and volcano plots for adverse events and FDA Medical Queries (FMQs) analysis outputs for inclusion in submissions to the FDA

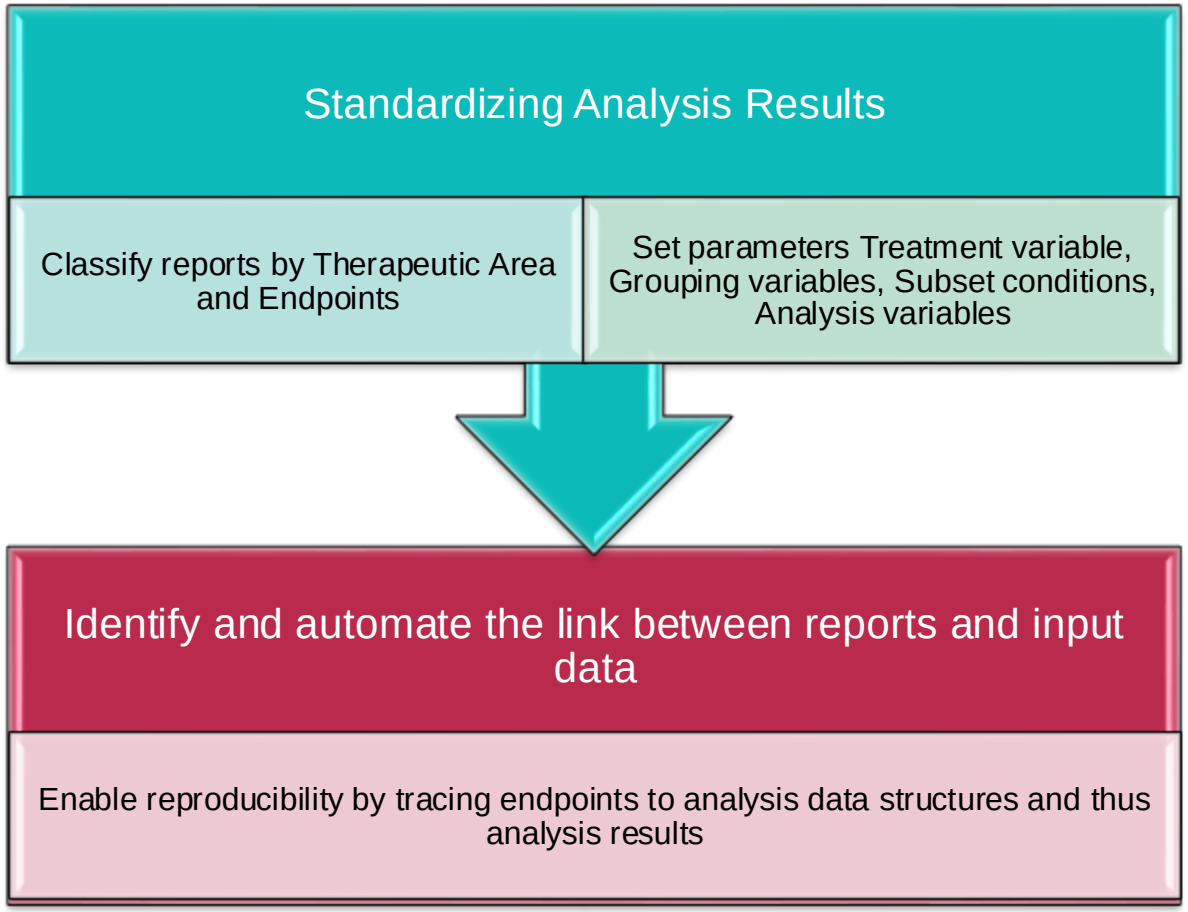
This work started in a PHUSE working group as a collaboration with the American Statistical Association and the FDA

Collaboration

Now officially recognized by COSA as being open-source projects focused on implementing or developing CDISC standards to drive innovation in the CDISC community.

Analysis Results Standards and Visualizations

- Analysis Results, when visualizations and summaries are used together, allow us to make inferences from a large amount of clinical data at once.
- Interactive graphs can go a step further by allowing the user to dive into each individual data point if required.



Prerequisites and References

Input data should be classified by domain/TA, and follow SDTM/ADaM IG standards:
- Labels, Variable Naming convention
- Code list
- Data types & Formats
- Derivation rules

Domain Specific Prerequisites; e.g., ADAE – SOC, Term, FMQ variables, Flag Variables, Grouping Variables

Referenced guidance: FDA TLF Integrated Guide, TAUGs, CDISC Standards for Adverse Event and FMQ mapping terms.

CDISC Pilot data is pre-loaded in the app for testing/exploratory purposes

Methods – Using R

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 the tidyverse suite of packages	flextable (RTRS recommended) and rmarkdown templates for tabular outputs	epitools, survival, ggsurvfit for statistical analysis

- Reproducibility of project environment using {renv}
- Unit Tests using {testthat}, High Test Coverage {covr}
- Automated CI/CD pipelines using Github Actions

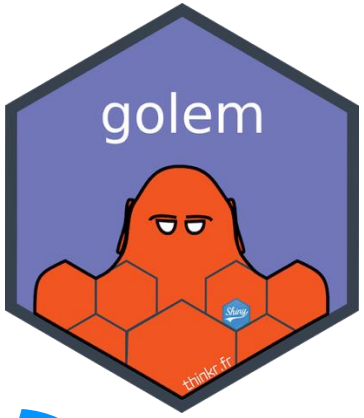
Pfizer

Shiny Interface

This package uses the {golem} framework to build the Shiny interface which consists of multiple reusable modules. Users can easily add or customize the existing modules to add or modify specific reports as per their need

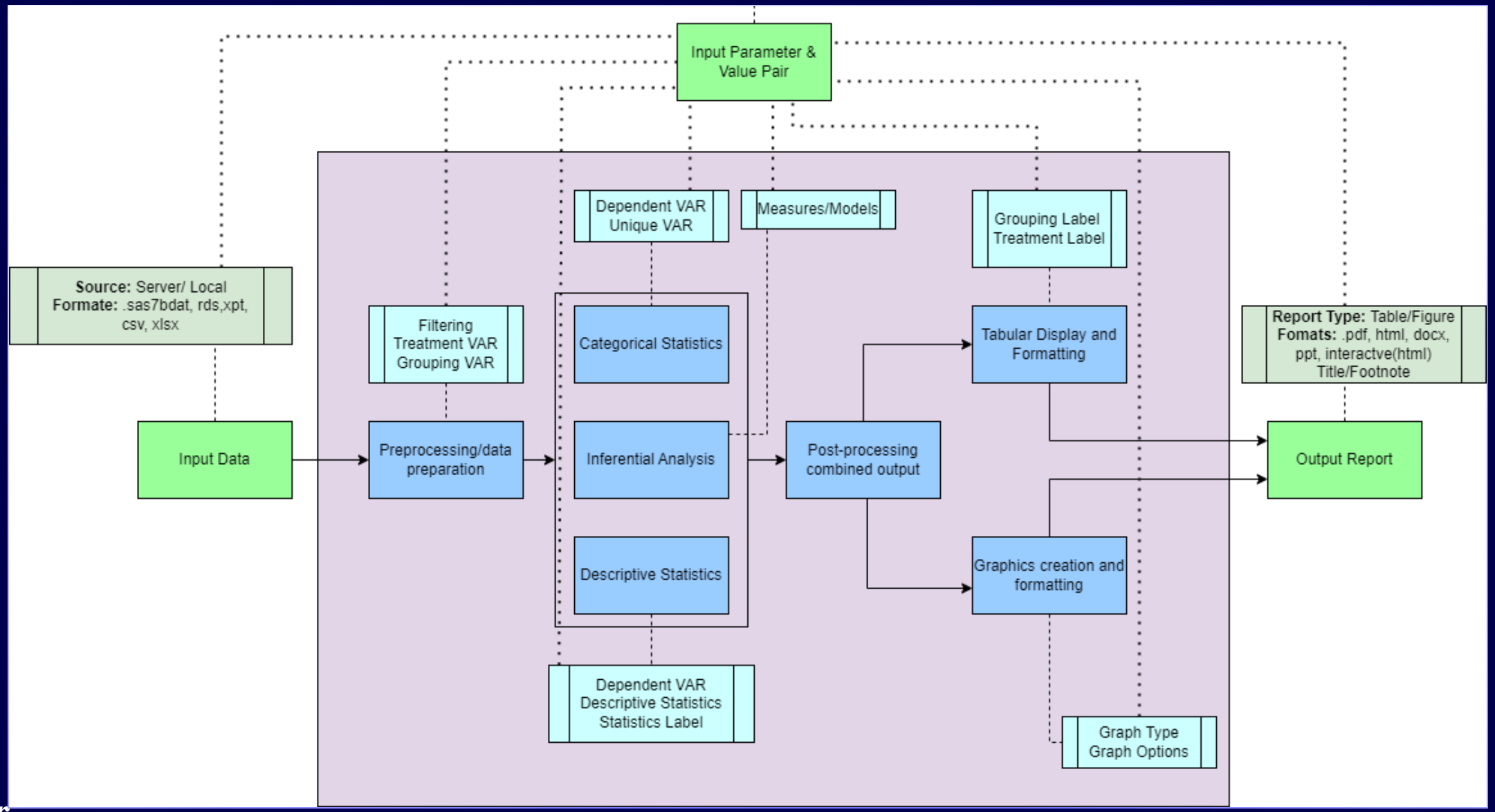
Tables and Figures can be generated interactively with customizable filters to drill down and enable interaction between multiple domains. Reports can be downloaded with intermediate datasets for QC

Shiny is easily scalable and distributable – easily deployed on R studio Connect or other options at Enterprise level; removing dependency on R installation. Can greatly speed up decision making by giving the end user direct access to all the results



Pfizer

Programming Flow



Outcomes

- Carver R package generates several safety and efficacy reports:
  - Demographic summaries
  - Adverse Events graphs and summaries (includes FMQs)
  - Survival, Time to Event Graphs
  - Vaccine TA Graphs
  - Generic Summary and Categorical Analysis
- RMarkdown templates for each report with minimal intervention required from user.
- Reports can be saved in multiple formats, along with data for QC and validation.
- The Shiny application includes interactive plots, ability to drill down to patient level and explore data on the fly.

https://github.com/pfizer-opensource/carver.git

Demonstration

Future Enhancements and Scope

Input Capture

- Save Report specific user inputs/filters for reproducibility
- Using saved inputs and outputs for QC

Additional Reports

- TA Specific Analysis
- Generic OCCDS and BDS analysis reports
- Identify reports used commonly across industry

Submission

- Bundling selected reports
- By embedding study-specific data, entire application can be used as part of submission package.

Release to CRAN for wider usage

Collaborate with variety of users

Conclusion

- Use of open-source for regulatory submissions or clinical reviews.
- Package customizable to make it company specific or use their enterprise servers (which makes data pipelines secure)
- Carver does not aim to be an exhaustive solution, but to cover the more commonly used regulatory reports.
- The scope of reports can be expanded via support and collaboration across industry.

Q & A

Contact:

smriti.anand@pfizer.com

sukalpo.saha@pfizer.com

Example Reports - Backup

Demographic Characteristics - Safety population

Gender	Ethnicity		Placebo (N=86)	Xanomeline Low Dose (N=84)	Xanomeline High Dose (N=84)	Total (N=254)
F	HISPANIC OR LATINO	Age (Years), n (%)				
		<65	1 (1.16%)	2 (2.38%)	1 (1.19%)	4 (1.57%)
		65-80	0	1 (1.19%)	0	1 (0.39%)
		>80	1 (1.16%)	1 (1.19%)	0	2 (0.79%)
		n	2	4	1	7
		Mean (SD)	74.5 (16.26)	65.5 (12.37)	56.0 (-)	66.7 (12.68)
		Median (range)	74.5 (63.0, 86.0)	62.0 (56.0, 82.0)	56.0 (56.0, 56.0)	63.0 (56.0, 86.0)
		(Q1,Q3)	(68.8, 80.2)	(56.0, 71.5)	(56.0, 56.0)	(56.0, 75.0)
		Race, n (%)				
		WHITE	2 (2.33%)	4 (4.76%)	1 (1.19%)	7 (2.76%)
		BMI				
		n	2	4	1	7
		Mean (SD)	23.1 (2.76)	27.4 (4.41)	23.5 (-)	25.6 (3.99)
		Median (range)	23.1 (21.2, 25.1)	26.5 (23.6, 33.0)	23.5 (23.5, 23.5)	24.2 (21.2, 33.0)
		(Q1,Q3)	(22.2, 24.1)	(24.0, 29.9)	(23.5, 23.5)	(23.6, 27.0)
	NOT HISPANIC OR LATINO	Age (Years), n (%)				
		<65	8 (9.30%)	3 (3.57%)	4 (4.76%)	15 (5.91%)

The denominator to calculate percentages is N, the Number of Participants in the full analysis set, within each treatment group
*n is the Number of Participants with non-missing AGE

Forest plot for Risk Ratio of Any Adverse Events

Example Reports - Backup