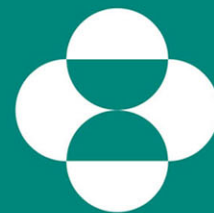


Proof of Concept (PoC) Development of Safety Visual Analytics (SVA) Using R Shiny and SAS



MERCK

INVENTING FOR LIFE

25 March 2019

Presented By: Rinki Jajoo, Yong Zhu

Biostatistics and Research Decisions Sciences (BARDS)

Agenda

- Problem Statement
- Objective of SVA team
- Challenges with Current tools
- SVA Design, Architecture and its Benefits
- A Sneak Peak Into the Tool
- Challenges and Solution
- Conclusion

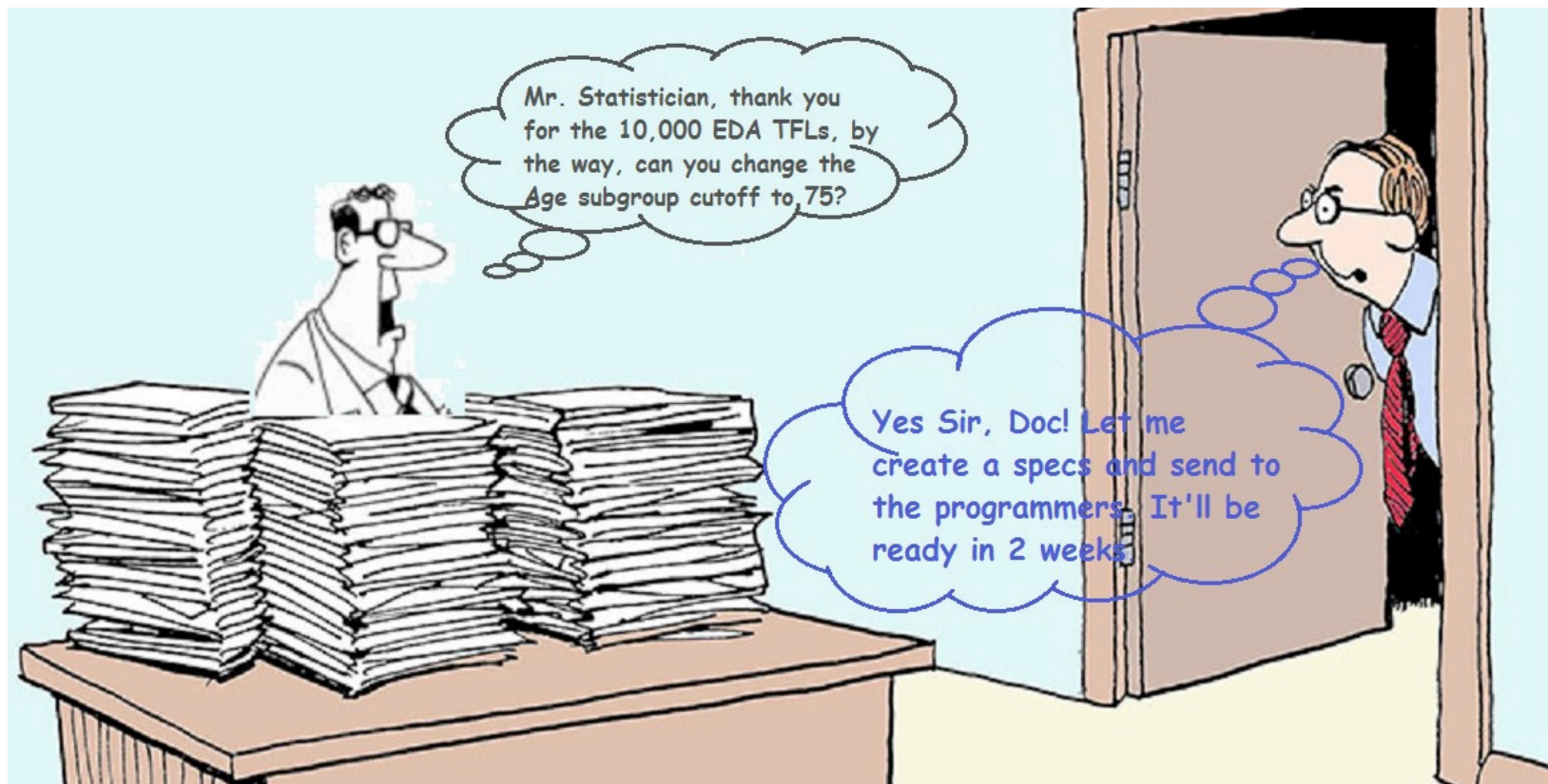


The greatest value of a picture is
when it forces us to notice what we
never expected to see.

— *John Tukey* —

AZ QUOTES

The greatest value of a picture is when it forces us to notice what we never expected to see.



Problem/Opportunity Statement

- Lack of interactive & dynamic tools with statistical methods required for aggregated monitoring and signaling
- Efficient review of aggregated data using sophisticated and interactive graphics (heatmap, volcano plot etc)
- Provide meaningful drill down or link to related safety data points (parameters)
 - E.g, Volcano plot to Rainfall plot and KM plot (time to event analysis)
- Alleviate the cumbersome burdens that safety physicians suffer every day during various reviews of ongoing clinical trial data.

Objective of Safety Visual Team

- Build a dynamic and interactive visual tool to align with ongoing aggregated safety evaluation (Aggregate Safety Assessment Plan - AgSAP)
- Enable the blinded and unblinded aggregate safety monitoring for Clinical Safety Risk Management (CSRM) physicians
- Maintain compliance (SDLC), confidentiality of ongoing trial and ability to compute large scale applications efficiently

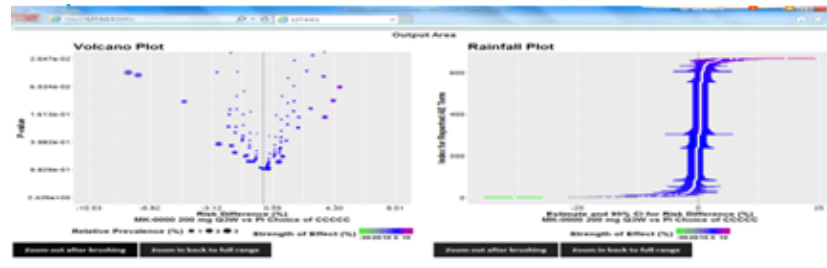
Yet Another Tool?

- Currently various systems are utilized across organizations to visualize data static as well as dynamic. (SAS, Python, R Shiny, SPSS, Matlab, Jreview, Spotfire etc)
- JReview supports interactive medical monitoring (Study level monitoring)
- No current dynamic or interactive tool to support aggregated data visualization, that enable both simple and sophisticated statistical evaluation in regulated environment

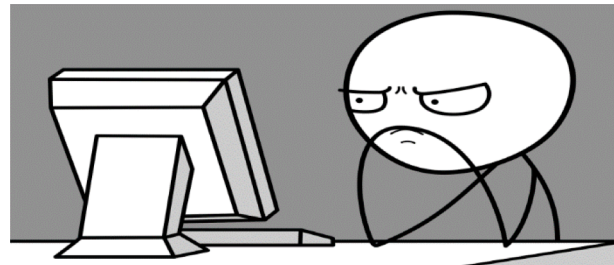
Challenges with The Current Tools

- Limited flexibility to modify the underlying statistical methods
- Requires tool specific data structure
- Lack traceability, reproducibility and validation of the tool
- Pre-canned reports methods cannot be modified.
- Expensive for larger consumption
- Not intuitive, too many parameters or selection (may not be relevant with their projects)

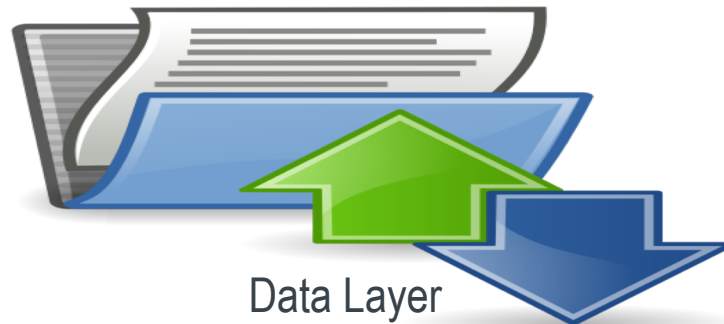
SVA Infrastructure: Multi-Layer Approach



Visual Layer



Computing Layer



Data Layer

Design of Safety Visual Analytic (SVA) Tool

- SVA three tier systems
 1. Data Visual layer (R Shiny)
 - R is an open source software, freely available made it easy to use.
 - Shiny is an R package that makes it easy to build interactive web apps straight from R
 2. Computational layer (SAS on Unix Platform)
 - Execute statistical algorithm at run time by pulling the data from data source layer.
 - Revert information after computing on SAS to R Shiny
 3. Data source layer (ADaM or SDTM datasets on Unix Platform)
 - Source data and computed data for visualization

SVA: Benefits of Multi-Layer Architecture

- Ability to update the Visual layer without impacting the other areas of the application
- Scalable by adding multiple server, allowing to deploy variety of database instead of being locked to one technology
- Reliable and ease of maintenance (change to one **layer** has minimal impact on other layers)
- Easy to adapt the ongoing technological updates

VACS – User Interface (1/2)

Blinded OAS

Unblinded OASE

Welcome to Visual Analytics Clinical Safety (VACS)

Introduction

VACS is an interactive visualization tool developed by WS#4@CSS, BARDS, via a real-time, dual computing infrastructure using SAS and R-shiny. VACS is aimed for automating various purposes of reviews of clinical trial safety data at program level to help timely decision-making on the contents of final safety reporting.

SVA Development Team

Role	Name
Sponsors	Ray Bain & Jose Vega
Team Leads	Sammy Yuan, Nathan Li, Yong Zhu & Rinki Jajoo
Core Developers	Nathan Li, Yong Zhu & Peter Hu
IT Supporters	Daniel Woodie & Elizabeth Arnold
Maintainer	Nathan Li & Yong Zhu

What Would You Like to Do While Waiting for An Analysis Result?

Depending on the instance of run-time environment, it may take 10-30 seconds for R-shiny to deliver an analysis result

- ☒ Enjoy some clips of music
- ☐ Wait quietly

VACS: User Interface (2/2)

Blinded OASE

Unblinded OASE

Safety Profile Plots

KM Plot

AE Bar Plot

Forest Plot

Frequentist-CMA

Bayesian-CMA

Welcome to Visit Clinical Safety (VACS)

Introduction

VACS is an interactive visualization developed by WS#4@CSS, BARC, time, dual computing infrastructure SAS and R-shiny. VACS is aimed at automating various purposes of clinical trial safety data at program level to help timely decision-making on the contents of final safety reporting.

Development Team

Role	Name
Sponsors	Ray Bain & Jose Vega
Team Leads	Sammy Yuan, Nathan Li, Yong Zhu & Rinki Jajoo
Core Developers	Nathan Li, Yong Zhu & Peter Hu
IT Supporters	Daniel Woodie & Elizabeth Arnold
Maintainer	Nathan Li & Yong Zhu

What Would You Like to Do While Waiting for An Analysis Result?

Depending on the instance of run-time environment, it may take 10-30 seconds for R-shiny to deliver an analysis result

- ☒ Enjoy some clips of music
- ☐ Wait quietly

VACS: User Defined Parameters

AE Profile Plot

Introduction

AE Profile Plot often refers to Volcano Plot and Rainfall Plot, which displays between-treatment comparison results for all AEs under investigation in one plot. Volcano plot is a scatter plot of transformed p-values against the point estimate of risk difference between 2 treatment groups from a randomized study, in the sense that a safety concern should not be determined the size of p-values alone but taken the size of effect into consideration as well. Rainfall plot is a horizontal bar plot displaying the confidence intervals ordered by the size of risk difference to show potentially the most harmful and protective effects of the investigational product.

Input Area

Analysis Population Specifications

List of Randomized Study ID

0000-045

Click box below to select exactly 2 treatments for comparison (The latter one selected is considered as the control)

MK-0000 200 mg Q3W PI Choice of CCCCC

AE Specifications

AE Term Level

Preferred Term

Narrow down AE Scope (select 1 or more or none)

- ☐ SAE
- ☐ Treatment-related AE
- ☐ Grade 3-5 AE
- ☐ AE leading to study disposition event
- ☐ AEOsI
- ☐ AE leading to death

Optional FDR Adjustment

- ☒ No FDR adjustment
- ☐ BH-FDR
- ☐ double-FDR

Choices of endpoints

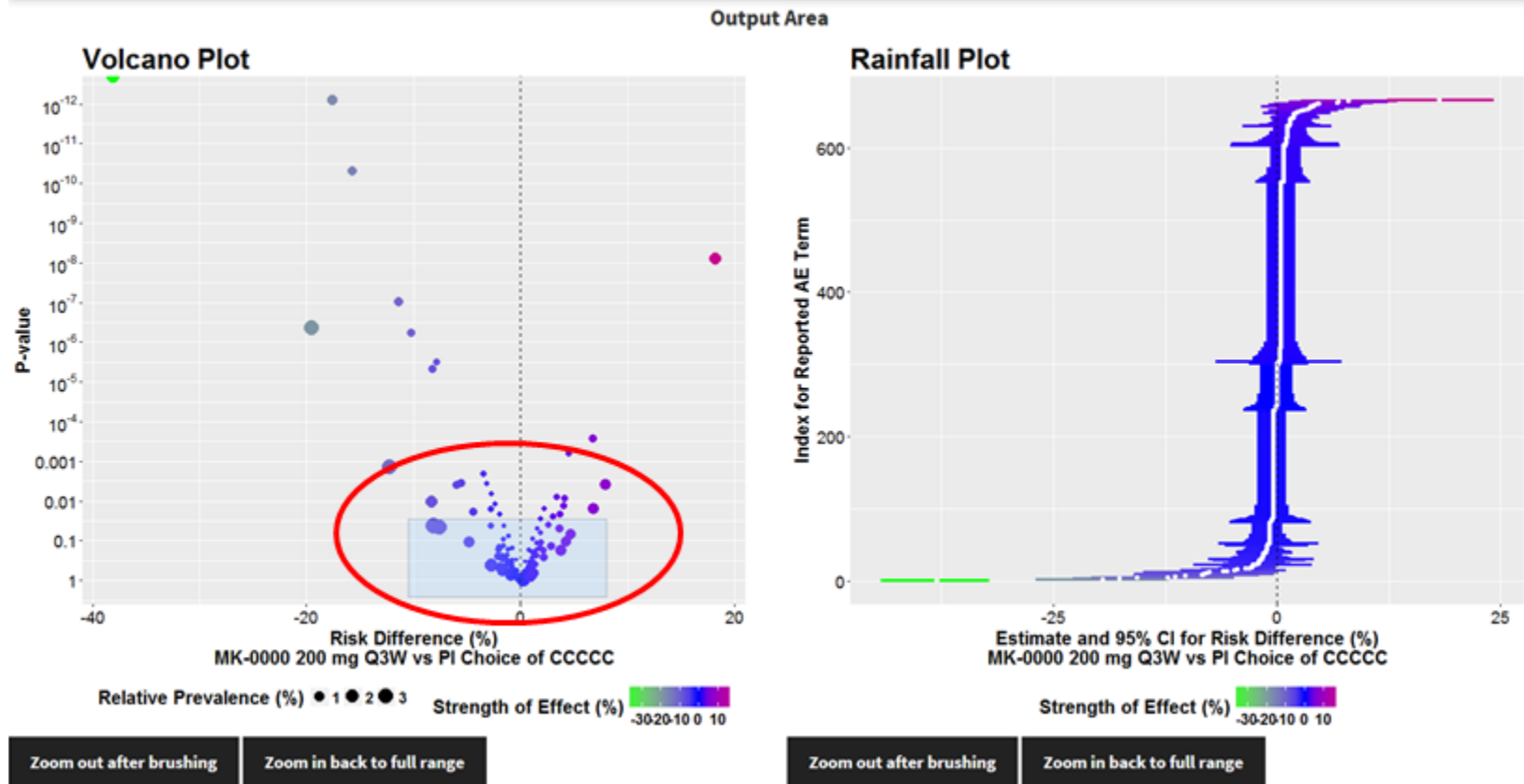
- ☒ Binary
- ☐ Time to event

Risk Metric Specifications

- ☒ Risk Difference (RD)
- ☐ Risk Ratio (RR)

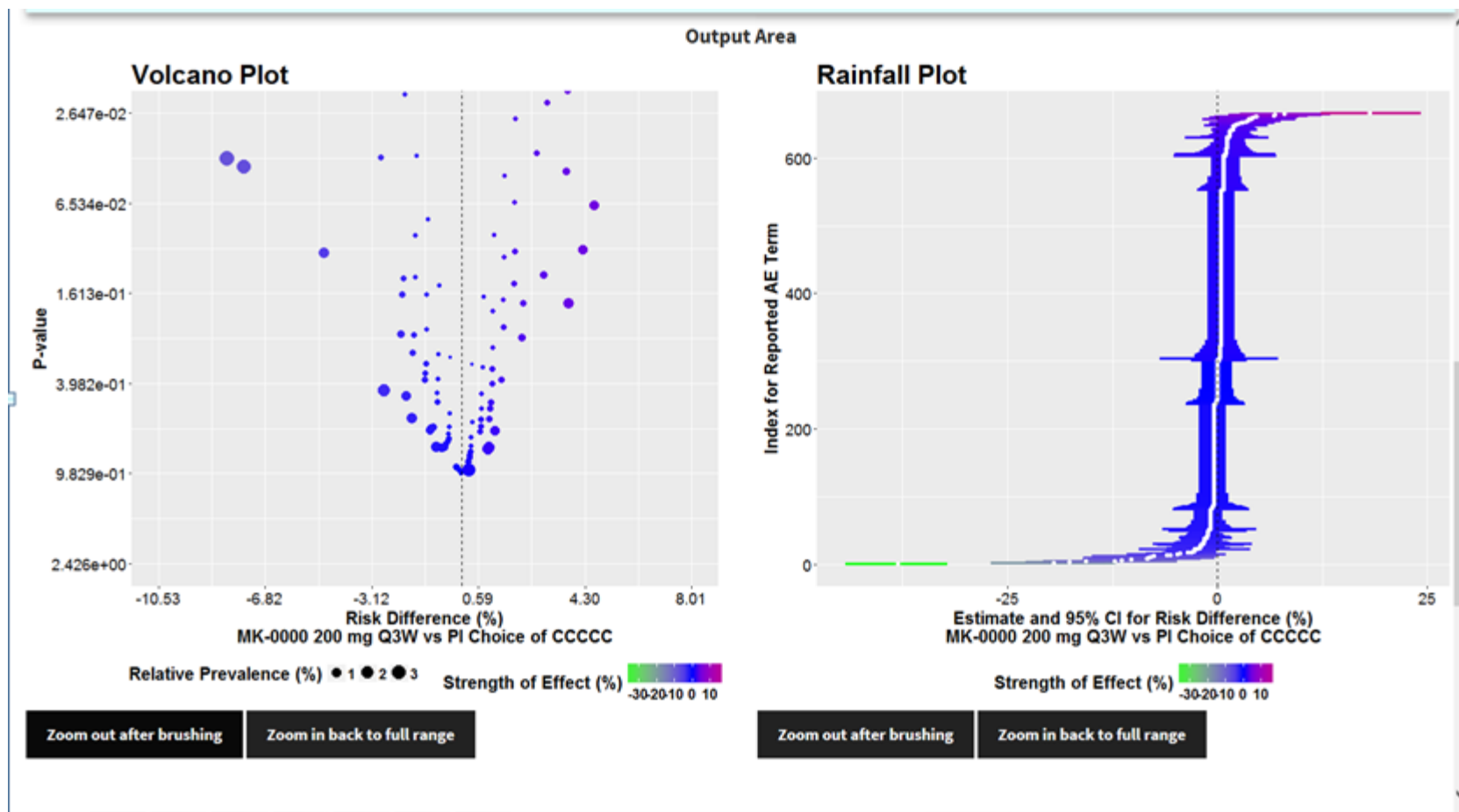
Submit your specifications

VACS: Safety Profile



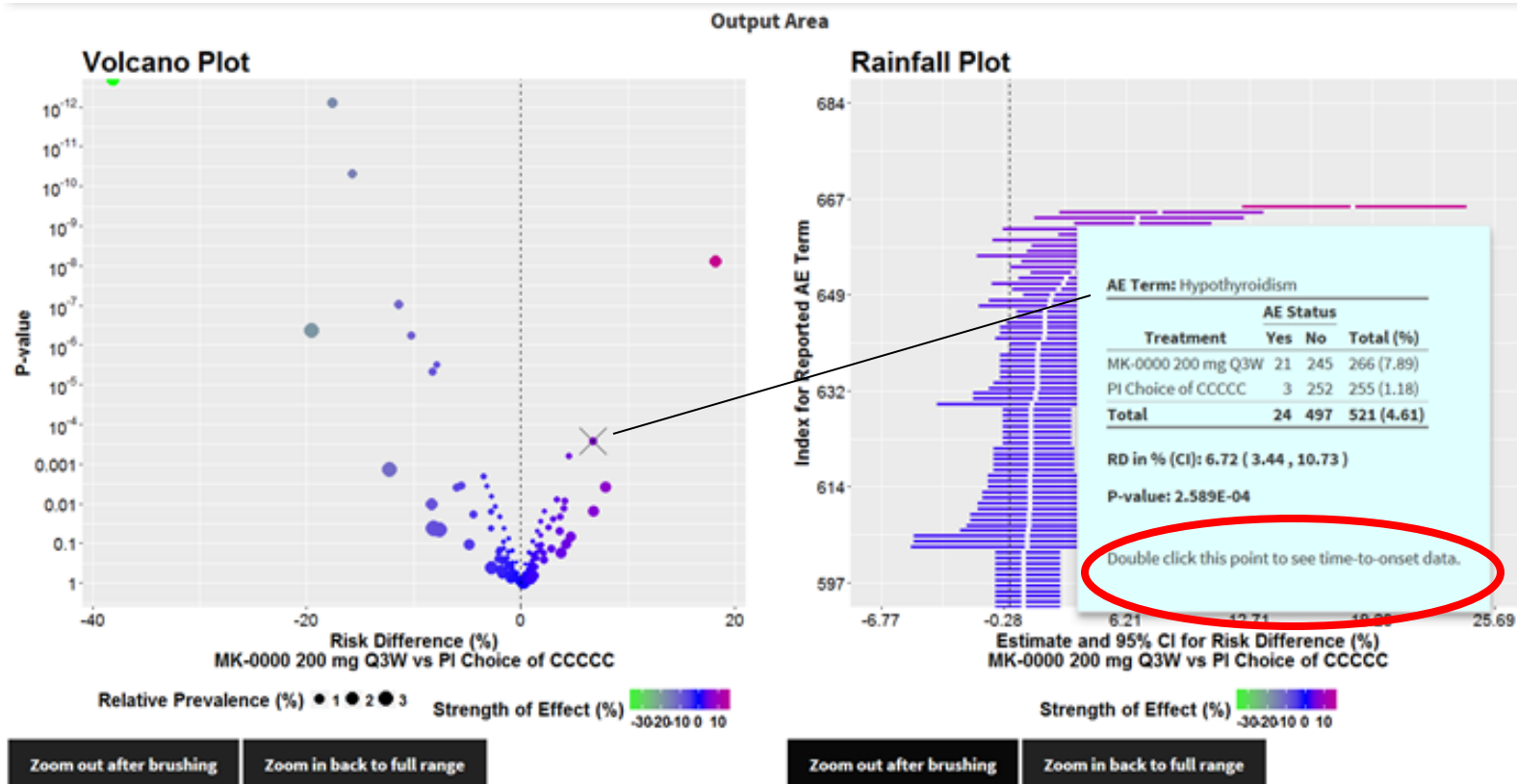
**Developed by Peter Hu & Yong Zhu

VACS: Zoom In...



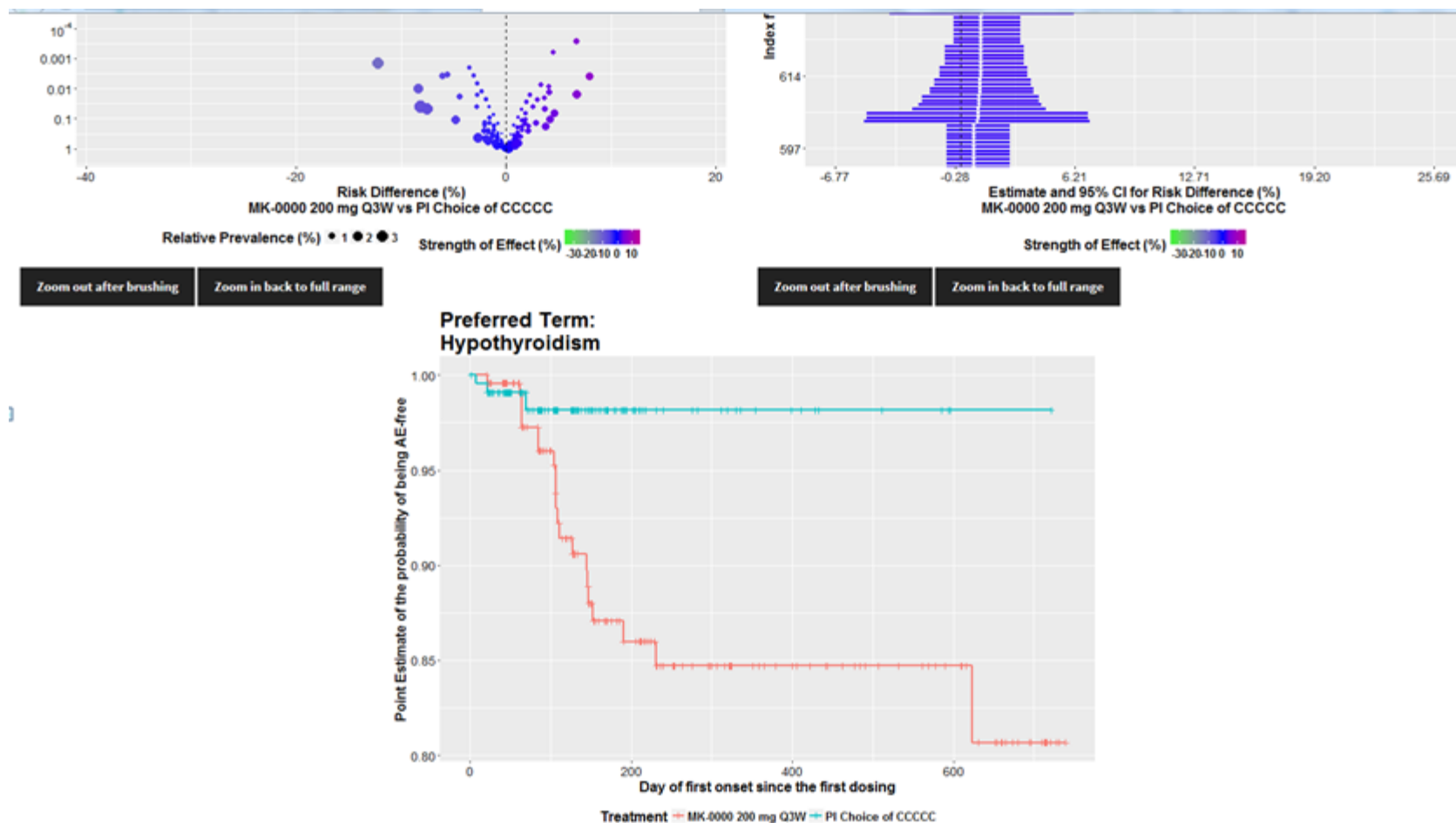
**Developed by Peter Hu & Yong Zhu

VACS: Drill Down



**Developed by Peter Hu & Yong Zhu

VACS: Drill Down



**Developed by Peter Hu & Yong Zhu

VACS: Compare Multiple Indication for An Adverse Event of Special Interest

Choose a Dataset:

AEOSI By Indications

Set a Comparison

Indication 18

Indications (the comparison must be chosen)

Indication 1 Indication 2 Indication 3
Indication 4 Indication 5 Indication 6
Indication 7 Indication 8 Indication 9
Indication 10 Indication 11 Indication 12
Indication 13 Indication 14 Indication 15
Indication 16 Indication 17 Indication 18

Order AEOSI based on following rules

Population

Indication 1

Ordering

Default

Cutoff of Relative Risk

0 2 6

0 1 2 3 4 5 6

Incidence Rate Size

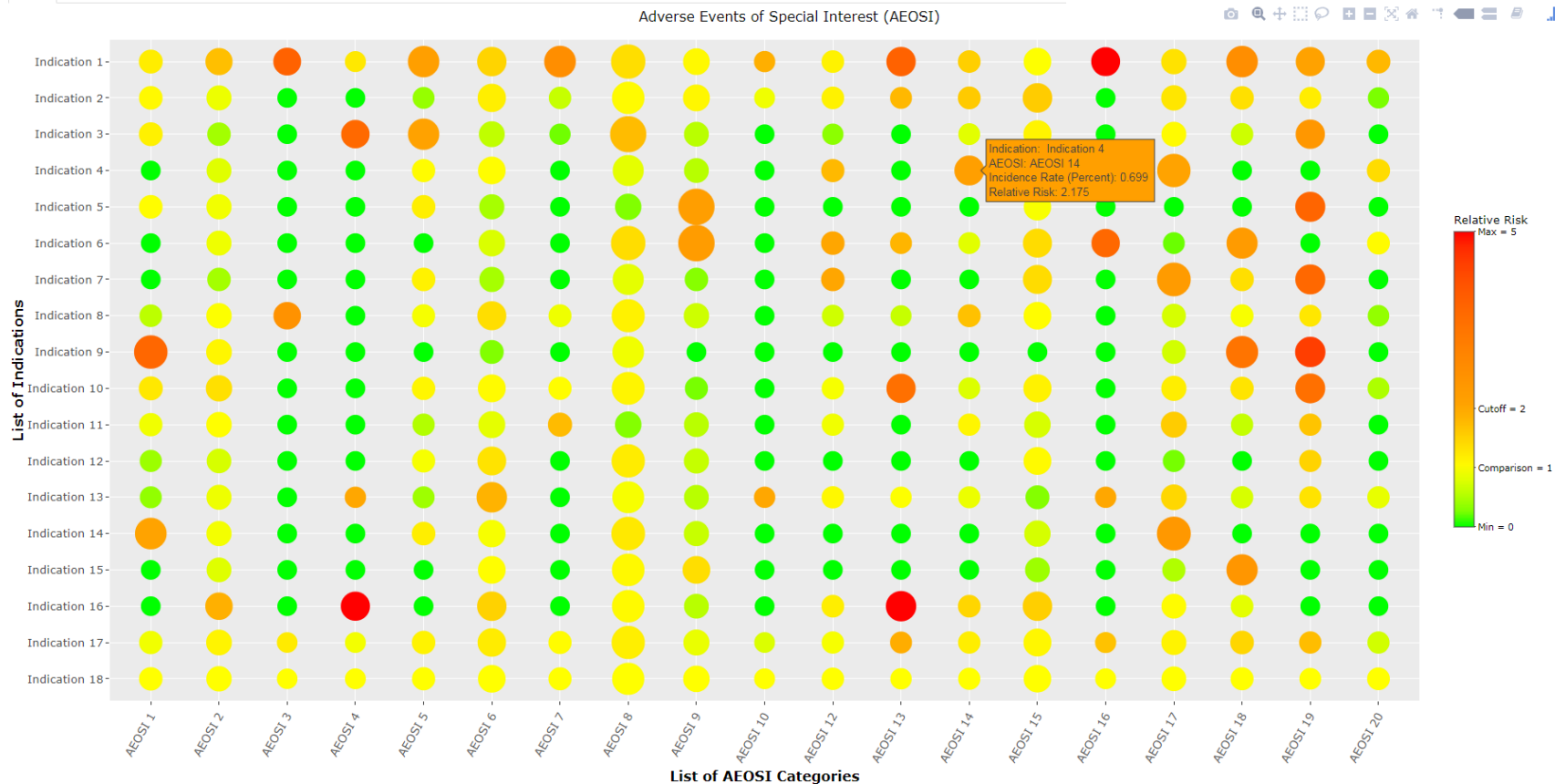
1 6 11 15

1 3 5 7 9 11 13 15

Color Choice

Colorful

Heat Map: Multiple Indication



Developed by Nathan Li; Ellen Snyder

Which AEOSI Has Larger Incidence Rate

Choose a Dataset:

AEOSI By Indications ▼

Set a Comparison

Indication 18 ▼

Indications (the comparison must be chosen)

Indication 1 Indication 2 Indication 3
Indication 4 Indication 5 Indication 6
Indication 7 Indication 8 Indication 9
Indication 10 Indication 11 Indication 12
Indication 13 Indication 14 Indication 15
Indication 16 Indication 17 Indication 18

Order AEOSI based on following rules

Population

Indication 18 ▼

Ordering

By Incidence Rates ▼

Cutoff of Relative Risk

0 2 6

0 1 2 3 4 5 6

Incidence Rate Size

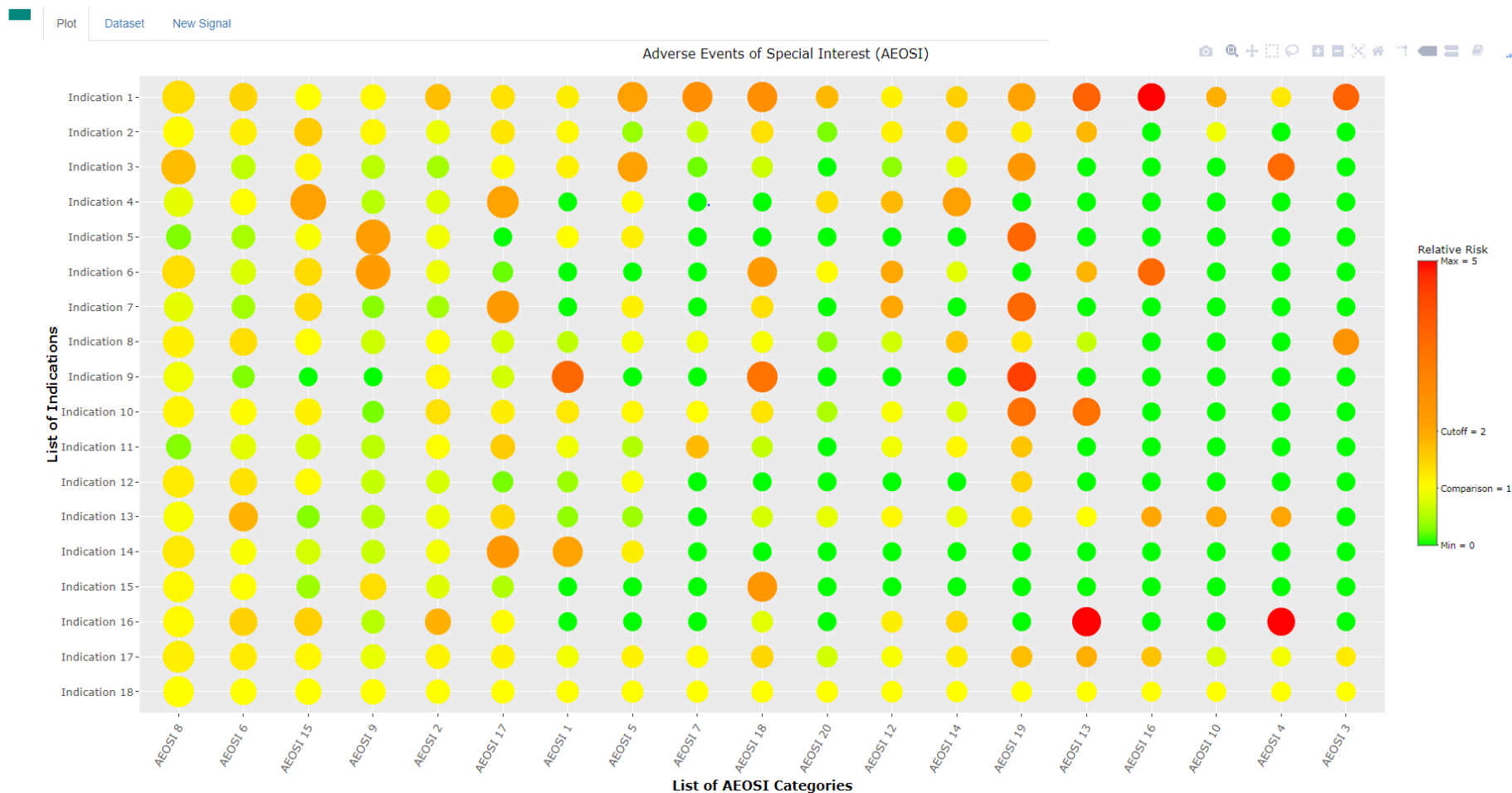
1 6 11 15

1 3 5 7 9 11 13 15

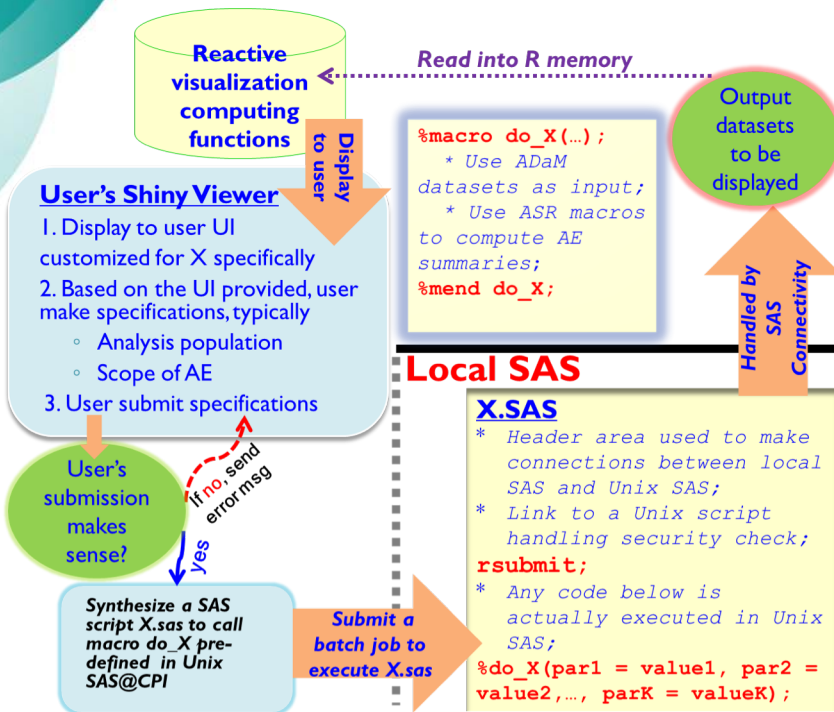
Color Choice

Colorful ▼

Heat Map: Sorted by Highest Incidence Rate



Developed by Nathan Li; Ellen Snyder

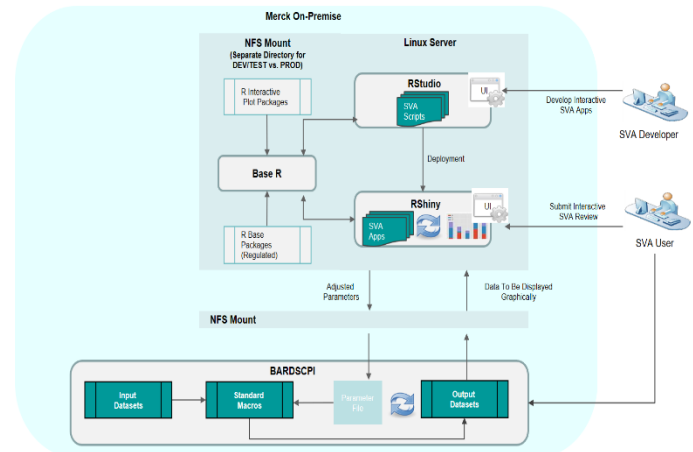


Proof Of Concept: Concept to Design of Safety Visual Analytics

Safety Visual Analytic Architecture

Concept

Design and implementation



Challenges & Solution(1/2)

- How to meet regulatory compliance?
 - Designed a divide-and-conquer mechanism to separate statistical computing and data visualization
 - Dataset computing handled by Unix SAS via calling standard SAS macros in Standard Knowledge Libraries used for analysis and reporting.
 - R-shiny is used for data visualization
- How to maintain data confidentiality and prevent from data being overly 'cooked'?
 - No data manipulation functionality provided to end user
 - Comprehensive user-interface (UI) provided to allow user making various specifications adherent to the chosen analysis module
 - UIs for specifications irrelevant to the chosen analysis module not provided
 - Each analysis module linked only to its own analysis specification UIs

Challenges & Solutions (2/2)

- How to enable communication between R-shiny (Linux) and SAS (Unix)?
 - Developed R package (rCPI) to enable communication between R-Shiny and SAS
 - The rCPI package created to interact with data layer (CPI) will enable to
 - Run SAS macros on CPI from the Shiny server
 - Read data from CPI
 - Create log files of every event
- How to assess and accommodate all the computing needs?
 - Based on user parameter selection enable real time computing
 - rCPI package enables transferring UI parameter to SAS for computing and returning output to user interface

Conclusion

- SVA system combines R-shiny as visualization tool and SAS as the computation tool.
- The design of SVA system enables us to achieve the goal of interactive graphs and computing on fly, which is based on the three-layer architecture structure.
- All the computations are performed by SAS with existing validated macros to guarantee the validity, reproducibility, traceability and compliance of all analyses.
- R is self-sufficient for any analyses but lack of fully validated functions and compliance make R not appropriate for analyses that needs full validations. In the ideal case, we hope to use R and shiny for both computation and visualization and is a long-term goal

Author(s) & Contributors

❖ Rinki Jajoo

- ❖ Associate Director –Clinical Safety Statistical Programming; Merck & Co. Inc

❖ Yong Zhu

- ❖ Associate Principle Scientist- Statistical Programmer Clinical Safety Statistical Programming; Merck & Co. Inc

❖ Sammy Yuan

- ❖ Principle Statistician- Clinical Safety Statistics Merck & Co. Inc

❖ Nathan Li

- ❖ Senior Scientist- Statistician Clinical Safety Statistics Merck & Co.Inc

❖ Daniel Woodie

- ❖ Sr. Specialist Engineer, Development & Integration Merck IT group.

Questions?

Thank You