Paper PD08

SAIL.R – Journey Towards R Software in Johnson & Johnson. A Training Perspective

Yannick Vandendijck, Johnson & Johnson, Beerse, Belgium
Annemie Van Dijck, Johnson & Johnson, Beerse, Belgium
Daniel Hofstaedter, Johnson & Johnson, Raritan, United States
Katherine Wilde, Johnson & Johnson, High Wycombe, United Kingdom
Mia Qi, Johnson & Johnson, Raritan, United States
Rawan Rupnow, Johnson & Johnson, Raritan, United States

ABSTRACT

SAIL.R is a comprehensive cross-functional initiative within Johnson & Johnson aimed at implementing R for diverse purposes. The primary goal is to facilitate the development of both static and interactive analyses for data exploration of clinical data and the delivery of production outputs using R software. Multiple streams have been established. Key streams include submissions, reporting framework and standards, interactive visualization and reports, and change management. One of the key areas of focus within the change management stream is the training of statisticians and statistical programmers in effectively utilizing R.

The main focus of this paper is to describe the R training program designed for statisticians within Johnson & Johnson on the use of R for exploring, summarizing, and visualizing clinical data. We provide insights into how the initiative aims to empower statisticians to harness the full potential of R by detailing the training framework, methodologies, contents and materials utilized.

INTRODUCTION

R is a free software environment for statistical computing and graphics [(1)]. Over the years, R has gained substantial popularity among pharmaceutical companies for data analysis, visualization, and modeling. Pharmaverse was established as a connected network of companies and individuals working to promote collaborative development of curated open-source R packages for clinical reporting usage in pharma. Adopting shared solutions in this post-competitive space should ultimately ease regulatory review, resulting in bringing new treatments to patients faster [(2)].

In 2015, the FDA issued a Statistical Software Clarifying Statement in which it is stated that the FDA does not require the use of any specific software for statistical analyses [(3)]. However, the software package(s) used for statistical analyses should be fully documented in the submission, including version and build identification. Further, as noted in the FDA guidance, E9 Statistical Principles for Clinical Trials [(4)], "The computer software used for data management and statistical analysis should be reliable, and documentation of appropriate software testing procedures should be available."

Within Johnson & Johnson, SAIL.R is a comprehensive cross-functional initiative with the goal to implement R for the development of both static and interactive analyses for data exploration and the delivery of production outputs of clinical data. The vision is to have R-based submissions by 2025, and to transition to use R for interactive analysis, visualization and exploration as standard practice. To achieve this vision, the primary goal is to establish a fully implemented R environment for developing static/interactive analysis for data exploration and production output delivery.

In the next section we provide a brief overview of the SAIL.R streams. Additionally, we provide details on the R training framework designed for statisticians which is aimed at using R for exploring, summarizing, and visualizing clinical data. Details will be provided on the training strategy, the curriculum, training implementation and challenges. Example materials will be shown. Finally, we describe differences between the R training implemented for statisticians and statistical programmers within Johnson & Johnson.

SAIL.R INITIATIVE

Implementation of R software in a big and complex organization with many stakeholders requires a broad framework to ensure a smooth and successful implementation. For this reason, the SAIL.R initiative encompasses various streams, each with specific goals. A brief overview of the four key streams within the SAIL.R initiative is provided below.

REPORTING FRAMEWORK AND STANDARDS

The Reporting Framework and Standards stream plays a crucial role in defining and establishing the core standards to be utilized for the creation of production outputs. In order to ensure the highest level of safety and compliance, these standards are aligned with the FDA's standards. As part of the implementation of R, this stream is implementing quality control measures such as reviewing and analyzing results from scripts written in R and SAS®; script validation by running newly developed standard template R scripts on real study data; and the development of a script catalog for standard R TLGs (Tables, Listings, and Graphs).

SUBMISSION

The Submission stream plays a crucial role in ensuring the smooth and successful submission of R developed programs and outputs. To achieve this, testing methodologies are employed to identify any necessary adjustments or additions to current submission programs, tools, and guidelines. These findings are then discussed with stakeholders where practical implementations are explored to enhance the submission process. The stream provides technical and strategic submission support for selected early-adopter studies that use R developed programs and outputs. This is done by providing briefing book support, pre-submission communication with regulatory divisions (e.g., mock-submission that was done before the real submission to already collect feedback from regulatory authorities before the real submission packages were created), and independently reviewing submission packages (assuming the role of a regulatory reviewer).

INTERACTIVE VISUALIZATION AND REPORTS

The Interactive Visualization and Reports stream focuses on optimizing and improving existing visualization applications while also developing new ones using R to enhance data visualization for exploratory purposes as well as the production of outputs for clinical study reports or interim safety analyses. One of the primary goals is to develop a library of interactive data visualization applications (i.e. using RShiny) that can support the streamlined production of static TLGs. This will enable efficient and effective data visualization and reporting. This stream aims to provide users with intuitive and interactive visualization tools that enable them to effectively analyze and interpret data.

CHANGE MANAGEMENT

The Change Management stream plays a crucial role in driving the adoption and implementation of R within the organization. Change champions are appointed within therapeutic areas to promote the use of R and showcase its benefits. Regular meetings are organized to facilitate information sharing between colleagues who are using R in their work. These meetings cover various topics, including small and interesting training topics, sharing experiences, and providing updates on processes related to R. Effective communication is also a key aspect of the change management stream, with a focus on updates from the SAIL.R initiative. Additionally, the stream is responsible for providing R training to ensure that statisticians and statistical programmers have the necessary skills to effectively use R in their daily work.

R TRAINING

An overview is provided on the R Training framework. The main focus is on the R training enrolled for statisticians at Johnson & Johnson on the use of R for exploring, summarizing, and visualizing clinical data. An important goal is to provide statisticians the R toolbox to perform QC on analysis datasets, tables, listings and figures. In the next sections we will discuss the strategy, curriculum, materials, implementation and examples of the R training.

STRATEGY

When developing a training program, it is crucial to have a clear understanding of the intended audience. Important questions that should be discussed are:

- What skills do they need to be trained on?
- What skills do they already have?
- What type of learners are they?
- What are the competing priorities for the audience?
- Are they eager to learn or is there already a change saturation?

Secondly, another important consideration is selecting the trainers who will deliver the training. These trainers should possess a deep understanding of the subject matter and have good communication and teaching skills. They should be able to engage and effectively convey information to the audience. Important questions that should be discussed are:

- Do you have a pool of statisticians that are experienced in R?
- Are these also experienced trainers, or do they need additional training to become an effective teacher?

- Will the trainers develop the training materials?
- What are the competing priorities of the trainers?
- How will you support questions during/after the training?

Finally, determining the environment in which the training will take place is essential. This includes considering the physical location, necessary equipment or technology, and any specific requirements or constraints. Important questions that should be discussed are:

- Do you already have a working R Environment for training? Will you need to utilize a vendor? Will you use R-Studio Desktop?
- How will you ensure a consistent environment for trainers and the audience?
- Do you want to use proprietary data for training and/or practice?
- Traditional Classroom Approach or Leveraging Modern Technology

By considering answers to the key questions above, a training program can be developed that effectively meets the needs of the audience, will help tailor the training content and ensures a successful learning experience.

The statistics group within Johnson & Johnson is a very diverse group with statisticians only experienced in SAS without any R knowledge and vice versa (R experience in college or prior job vs. decades with SAS). The competing priorities of the audience, as in many organizations, is the need to continue the day-to-day job (statistical design support, submission support, other trainings, still need to use SAS in ongoing projects, etc.).

Younger colleagues with a lot of R experience are available as trainers for both developing the training materials and providing the trainings. Practice sessions were planned with each trainer to guide them in becoming an effective teacher. The competing priorities of a trainer are the same as the audience. Trainers were selected from different therapeutic areas and groups (e.g. early and late development) to develop a network of R experts (subject matter experts) to whom other colleagues can reach out to for R related questions.

Training was provided in the statistical programming and computing environment that is used by all Johnson and Johnson statisticians and statistical programmers for clinical data. RStudio within this environment was used for the training. To keep the training closely relatable to the job, all data examples were on SDTM and ADaM datasets. The clinical test data provided by the CDISC organization was used for this purpose (available at https://github.com/cdisc-org/sdtm-adam-pilot-project).

CURRICULUM

The curriculum of the R training is split into 2 parts. Part 1 focuses on an introduction to R and covers all the basics needed for exploring, summarizing, visualizing of clinical data and performing QC on TLGs. Part 2 requires basic R knowledge and the focus is on training of R tools for more automation and interactive exploration of clinical data. Part 1 is split into 8 different topics, whereas Part 2 consists of 2 different topics. In the sections below example training material and code are provided for some of these topics.

Part 1

- Topic 1: R and RStudio:
  - Setup RStudio in the statistical programming and computing environment used at Johnson & Johnson
  - RStudio interface
  - Navigating in RStudio
- Topic 2: R Programming Basics
  - Intro to R objects (vectors, factors, matrices, data frames, tibbles, etc.)
  - R packages
  - How to load data
- Topic 3: Data Manipulations 1
  - Introduction to *dplyr* [5] package
  - Filter, select, rename, summarize, group by, etc. functions
  - How to join data
- Topic 4: Data Manipulations 2
  - Pivoting data from long to wide, and vice versa
  - Working with dates and times
  - Compare datasets
  - Introduction to *tidyltg* [6] package to output tables and listings
- Topic 5: Data Visualization 1
  - Introduction to *ggplot2* [7] package
  - The most important geom functions
  - Building a good-looking figure

    - Changing colors, axis, labels, etc.
    - Plotting in multiple panels
- Topic 6: Data Visualization 2
    - More example of figures using *ggplot2*
        - Bar graphs
        - Boxplots
        - Line plots
        - Forest plots
    - Introduction to *tidyltg* [6] package to output graphs
- Topic 7: Modeling 1
    - Inference for binary variables
        - One-proportion tests, estimates and confidence intervals
        - Comparing two proportions: Fisher's exact test, Chi-square test, Cochran-Mantel-Haenszel test
        - Logistic regression
    - Inference for continuous variables
        - Two-sample t-test
        - Wilcoxon rank-sum test
        - ANOVA
        - Linear regression
- Topic 8: Modeling 2
    - Introduction to the *mmrm* [8] package
    - Survival analysis

Part 2

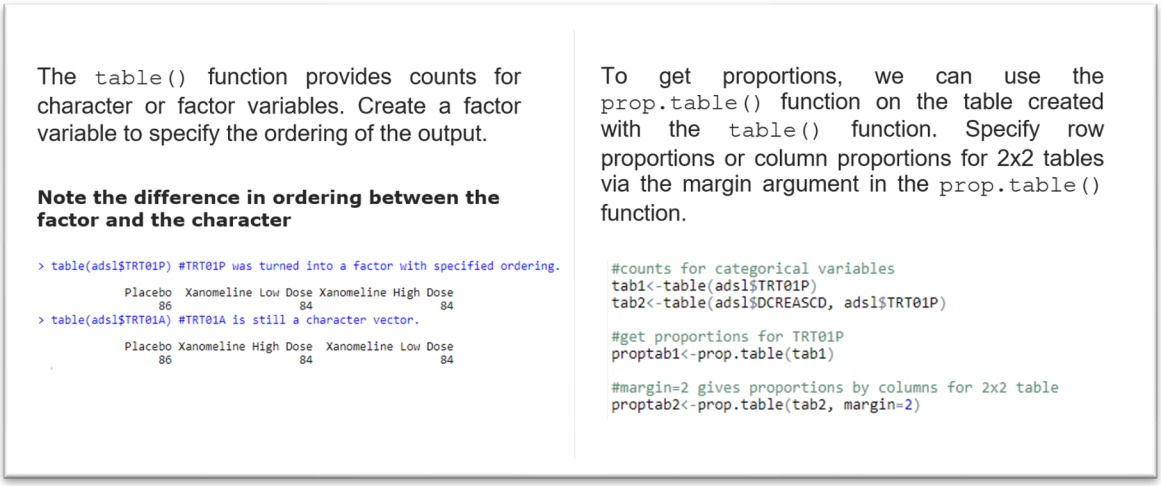
- Topic 9: R Markdown and Interactive plots
    - Introduction to R Markdown [9]
    - Introduction to *plotly* [10] package for interactive plotting
- Topic 10: Examples of internal innovative tools created using R Markdown

MATERIALS

The following training materials are created for each topic of the training. The training materials were developed to provide quick wins for the audience, ensure relatability to their job, and keep the audience actively engaged. All training materials were made available to the participants before the training started.

- **Presentation:** Slide decks were created to serve multiple purposes: The topic was introduced and learning goals were specified; the demo R code (see below) was provided with explanations; possible background information is provided; for some specific functions R vs SAS code was provided to highlight important differences (see below).

Figure 1: Example slide showing the use of *table()* and *prop.table()* function.

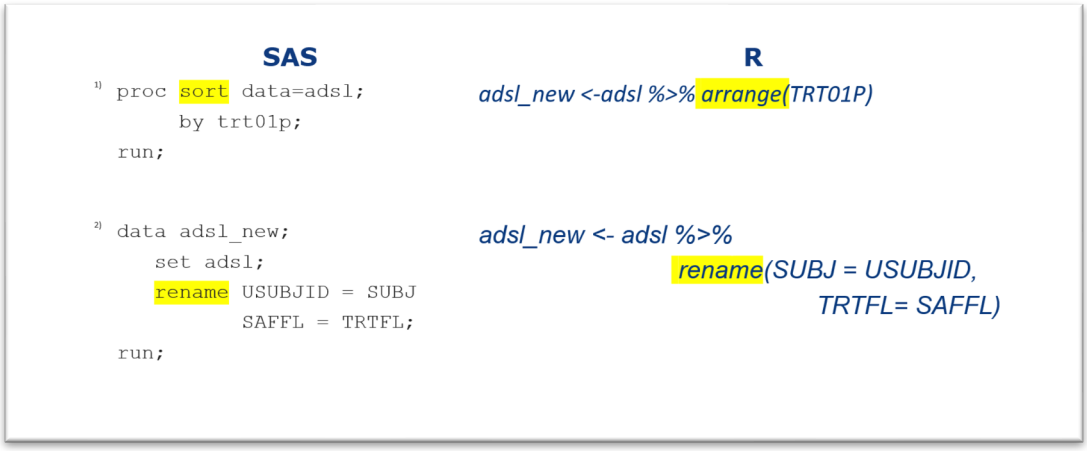


- **Demo R Code:** In most cases, the slides were used to introduce the topic and the learning goals. However, the major part of the training was done directly in RStudio showcasing demo code. The demo code contained all code used for a specific topic and included clear comments to facilitate easier understanding for trainees

and support the use of the code in their own projects. The participants were encouraged to run the demo code alongside with the trainer to reinforce learning and to keep them engaged during a training.
A high emphasis was set on using the "?" function to showcase how to easily find more information on functions (e.g., which arguments are needed in a function, what are the defaults, etc.). Two examples of demo code are presented further in this paper.

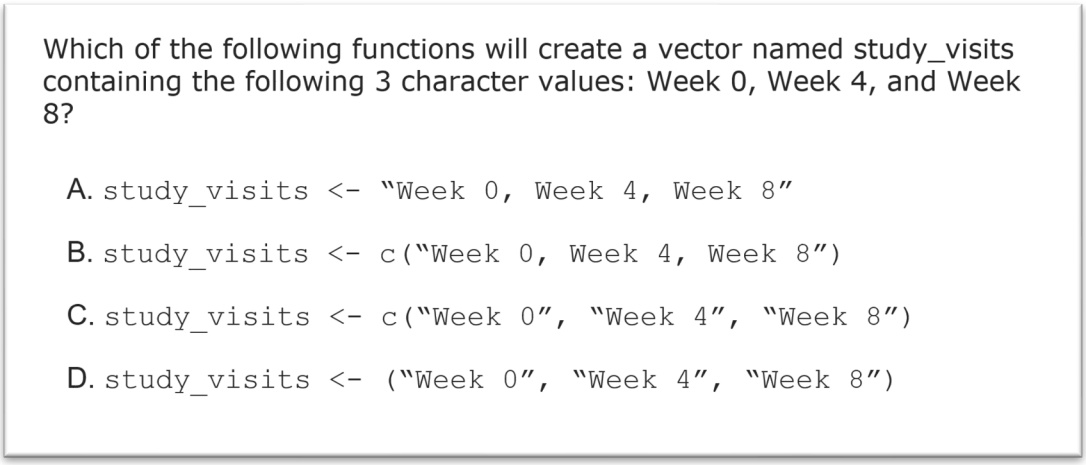
- **R vs SAS:** For some specific functions R vs SAS code was provided to highlight important differences. An example is presented in Figure 2 highlighting the difference in position of the new and old name in the *rename()* function between R and SAS.

Figure 2: Example slide showing the *arrange()* and *rename()* function in the dplyr package in comparison to SAS counterparts.



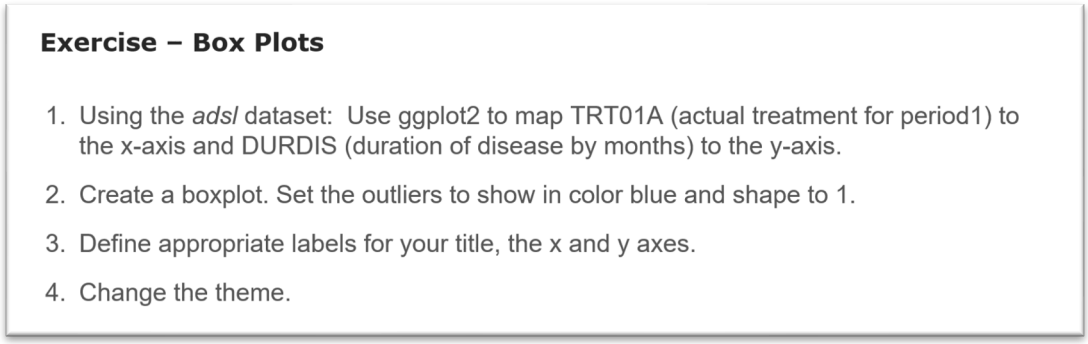
- **Interactive Polls:** Three to four class polls were included for each topic (example provided in Figure 3). Class polls are highly valuable during training sessions as they encourage active participation and engagement from participants. By incorporating polls, trainers can effectively assess the understanding on specific topics or concepts. This real-time feedback allows trainers to address any knowledge gaps. Furthermore, polls promote interaction amongst participants.

Figure 3: Example class poll.



- **Exercises:** Two to three practical exercises were planned for each topic to provide hands-on experience and reinforce the knowledge and skills learned. Trainees can compare their solutions with or learn from the solutions provided by the training team.

Figure 4: Example of a practical exercise.



- **Video recording:** For each topic, a video recording was made available as well. This allowed participants to review the content and material covered during the training at their own pace. This can be particularly helpful for complex or detailed topics, as participants can revisit the video to reinforce their understanding and refresh their memory. A video recording also serves as a valuable resource for those who were unable to attend the training session. Further, a video recording can be used as a training resource for new employees or team members.

IMPLEMENTATION

Practical implementation of the training will involve several key components. Firstly, it was decided to have online "live" classes. The choice to conduct the training online was driven by the international nature of the organization.

To accommodate participants' varying schedules, multiple classes per topic were offered. This flexibility allowed individuals to subscribe to the sessions that align with their availability, ensuring that maximum participation was achieved. Four to six classes were organized per topic with a maximum of 30 possible registrants per class.

To effectively manage the registration process and track participant progress, a registration system was implemented. This registration process ensures that all participants were accounted for and allowed for timely communication and updates regarding the training.

Each training session followed a consistent and standardized structure. This approach ensured that participants comprehend and followed the material in a systematic manner. It aids in building a solid foundation and understanding of the topics covered.

Each lesson was planned to take 90 minutes, with the first 60 minutes covering sections of the slides and showing the demo code with detailed explanation. The last 30 minutes were allocated for exercises. All classes for a single topic were provided over a span of 2-3 weeks before moving on to the next topic. This allowed participants to have some time between different topics and to delve into each subject without feeling overwhelmed. The total duration of training was approximately 6 months.

EXAMPLE 1: DATA MANIPULATIONS

In topic 3 (Data Manipulations 1) the *dplyr* package is introduced. Frequently used functions, such as filter(), select(), mutate(), rename(), summarize(), group by(), etc. are introduced. Initially, all of these functions are explained separately. Finally, they are combined to demonstrate an example in which:

- adsl dataset is filtered to include only ITT patients;
- a BMI variable is calculated from the available HEIGHTBL and WEIGHTBL variables;
- several summary statistics are calculated by treatment group.

Here is the example demo code that is explained in detail to the trainees:

```
tbl = adsl %>%
  filter(ITTFL == "Y") %>%                          # Filter ITT population
  select(USUBJID, TRT01A, HEIGHTBL, WEIGHTBL) %>%   # This step is not required
  mutate(BMIn = WEIGHTBL / (HEIGHTBL / 100) ^ 2) %>%  # Calculate BMI variable
  group_by(TRT01A) %>%                              # Summary stats by treatment
  summarise(                                        # Calculate summary stats
    n = n(),
    nHeight = sum(!is.na(HEIGHTBL)),
    AvgHeight = mean(HEIGHTBL, na.rm = TRUE),
```

```
    nWeight = sum(!is.na(WEIGHTBL)),
    AvgWeight = mean(WEIGHTBL, na.rm = TRUE),
    nBMI = sum(!is.na(BMIn)),
    AvgBMI    = mean(BMIn, na.rm = TRUE)
  )
```

EXAMPLE 2: DATA VISUALIZATION

In topic 6 (Data Visualization 2) several different types of graphs using *ggplot2* package are introduced. One type of graph that is introduced is a line plot. To do this, the line plot is build-up gradually with each step adding a new feature to the graph. The following demo code are explained to the trainees in steps:

1. From the dataset of interest filter the ITT patients and the parameter of interest. Create a factor variable of treatment group.

```
temp01 = adqsadas %>%
  filter(PARAM == "Attention/Visual Search Task",
         ITTFL == "Y", ANL01FL == "Y") %>%
  mutate(TRTPf = factor(TRTP, c("Placebo", "Xanomeline Low Dose",
                                "Xanomeline High Dose")))
```

2. Calculate summary statistics by treatment group and visit (exclude missing values).

```
summ_data = temp01 %>%
  filter(!is.na(AVAL)) %>%
  group_by(TRTPf, AVISIT, AVISITN) %>%
  summarise(n = n(),
            mean = mean(AVAL),
            sd   = sd(AVAL),
            se   = sd(AVAL)/sqrt(n()))
```

3. Generate plot 1: basic ggplot call with adding points and lines (explain the *group* input).

```
ggplot(data = summ_data,
       aes(x = AVISITN, y = mean, group = TRTPf)) +
  geom_point(size = 2.5) +
  geom_line()
```

4. Generate plot 2: add different colors, line types and shape per treatment group.

```
ggplot(data = summ_data,
       aes(x = AVISITN, y = mean,
           col = TRTPf, linetype = TRTPf, shape = TRTPf)) +
     geom_point(size = 2.5) +
     geom_line()
```

5. Generate plot 3: Add error bars to the graph.

```
ggplot(data = summ_data,
       aes(x = AVISITN, y = mean, ymin = mean - se, ymax = mean + se,
           col = TRTPf, linetype = TRTPf, shape = TRTPf)) +
  geom_point(size = 2.5) +
  geom_line() +
  geom_errorbar(width = 1)
```

6. Generate plot 4: Make use of position_dodge() to avoid overlapping error bars.

```
ggplot(data = summ_data,
       aes(x = AVISITN, y = mean, ymin = mean - se, ymax = mean + se,
           col = TRTPf, linetype = TRTPf, shape = TRTPf)) +
  geom_point(size = 2.5, position = position_dodge(width = 2)) +
  geom_line(position = position_dodge(width = 2)) +
  geom_errorbar(width = 1, position = position_dodge(width = 2))
```

7. Generate plot 5: Add more details to the plots by changing the x-axis , y-axis, add manual colors for the treatment groups and a horizontal line.

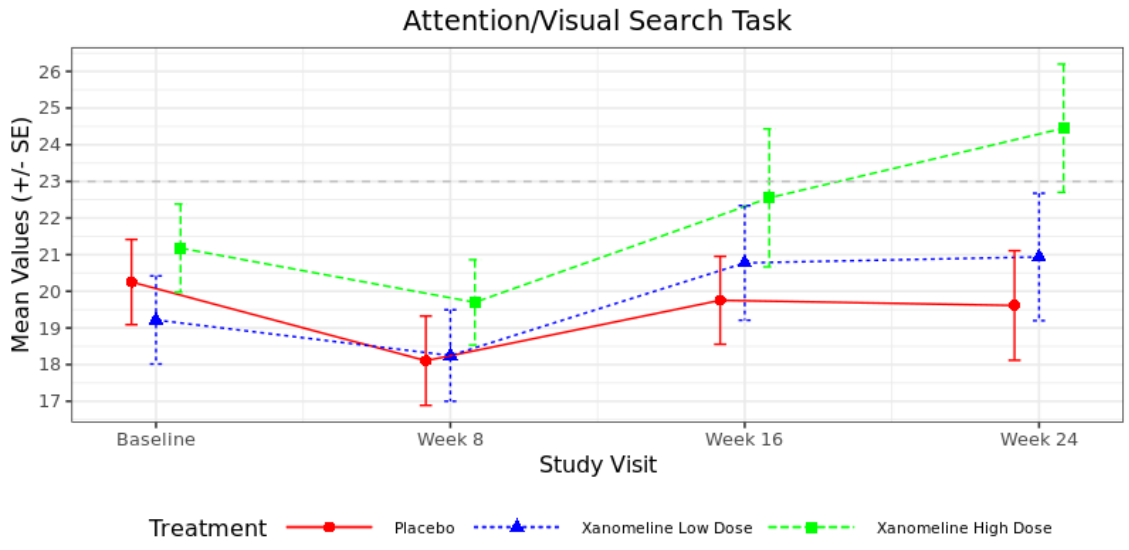
```
ggplot(data = summ_data,
       aes(x = AVISITN, y = mean, ymin = mean - se, ymax = mean + se,
           col = TRTPf, linetype = TRTPf, shape = TRTPf)) +
  geom_point(size = 2.5, position = position_dodge(width = 2)) +
  geom_line(position = position_dodge(width = 2)) +
  geom_errorbar(width = 1, position = position_dodge(width = 2)) +
  scale_x_continuous(breaks = c(0, 8, 16, 24),
                     labels = c("Baseline", "Week 8", "Week 16", "Week 24")) +
  scale_y_continuous(breaks = seq(15, 30, 1)) +
  scale_color_manual(values = c("red", "blue", "green")) +
  geom_hline(yintercept = 23, color = "grey", linetype = 2)
```

8. Generate plot 6: Add more details to the plots by adding appropriate labels for axes, title, and changing the theme.

```
ggplot(data = summ_data,
       aes(x = AVISITN, y = mean, ymin = mean - se, ymax = mean + se,
           col = TRTPf, linetype = TRTPf, shape = TRTPf)) +
  geom_point(size = 2.5, position = position_dodge(width = 2)) +
  geom_line(position = position_dodge(width = 2)) +
  geom_errorbar(width = 1, position = position_dodge(width = 2)) +
  scale_x_continuous(breaks = c(0, 8, 16, 24),
                     labels = c("Baseline", "Week 8", "Week 16", "Week 24")) +
  scale_y_continuous(breaks = seq(15, 30, 1)) +
  scale_color_manual(values = c("red", "blue", "green")) +
  geom_hline(yintercept = 23, color = "grey", linetype = 2) +
  labs(x = "Study Visit", y = "Mean Values (+/- SE)",
       title = "Attention/Visual Search Task",
       color = "Treatment", shape = "Treatment", linetype = "Treatment") +
  theme_bw(base_size = 12) +
  theme(plot.title = element_text(hjust = 0.5),     # this centers the title
        legend.position = "bottom",                 # legend position
        legend.text     = element_text(size = 8),   # legend text size
        legend.key.width = unit(2, "cm")             # length line in legend
```

9. The resulting figure is:

Figure 5: Resulting figure from example 2 data visualization (line plot).



STATISTICIAN VS STATISTICAL PROGRAMMERS

The above sessions aimed to describe the R training program provided for statisticians at Johnson & Johnson on the use of R for exploring, summarizing, and visualizing clinical data. Here, we briefly describe differences between the R training implemented for statisticians and that for statistical programmers.

The statistical programing team consisting of approximately 300 programmers, located across the United States, Europe, and Asia, is mainly responsible for programming ADaM analysis datasets and creating statistical summaries (TLGs) to support formal analysis of clinical studies. These programmers have extensive experience in clinical trial work and a strong understanding of programming concepts. Their primary expertise lies in programming using SAS, and their experience with R is often limited or non-existent. They are often faced with strict timelines for generating TLFs at database locks, submissions, and health authority requests.

The levels of R proficiency are different between the statistician and the statistical programing teams. Before the R Training, approximately 59% of the statisticians indicated that they used R more than SAS, and approximately 50% felt comfortable using R in their current work. In contrast, less than 20% of the statistical programmers were confident in using R. The list below emphasizes differences in focus between both groups:

- The first four topics (R and RStudio; R Programming Basics, Data Manipulations 1 & 2) were very similar across the statistician and the statistical programing teams. These topics include all building blocks to accomplish basic tasks that are needed for both teams.

- For the statistical programing team a detailed training was provided on the statistical reporting package (*tidytlg* package) as creating TLGs according to predefined standards is a core part of their daily work. Whereas for the statisticians, that was only touch-upon at a high level.

- For statisticians a greater emphasis was placed on data visualization, as they require a big skillset to create informative graphs to support clinical teams and internal decision making. Further, as can be expected, statisticians received more detailed trainings on statistical modeling using R.

- A training session was conducted for statistical programmers on creating functions in R to automate and customize programming code. Also, a comparison to SAS macros was included in their training.

- For statisticians, specific trainings were provided on the use of R Markdown and interactive plots using the *plotly* package. This was done since many statisticians use R Markdown reports and interactive graphs for effective communications with their clinical teams for internal decision making.

CONCLUSION

This paper provides insights into the framework employed for training statisticians in the effective use of R. The training strategy, contents, materials and its implementation are described. Training on the effective use of R is a continuous process. Training on how to use R for study design, power calculations, simulations, advanced plotting will be offered in the future.

A “live” classroom approach with internal trainers was chosen for the R training for statisticians. Another approach could be through an online learning platform. This could, for example, be chosen when an organization lacks sufficient in-house R expertise, does not have training expertise, or is unable to spare the time or resources for preparing and conducting the training.

REFERENCES

1. https://www.r-project.org/ (2024)
2. https://pharmaverse.org/ (2024)
3. Statistical Software Clarifying Statement. FDA (2015). Available at https://fda.report/media/109552/Statistical-Software-Clarifying-Statement-PDF.pdf
4. ICH harmonised tripartite guideline: statistical principles for clinical trials (E9) (1988). Available at http://www.fda.gov/Drugs/GuidanceCompliance RegulatoryInformation/Guidances/default.htm
5. Wickham H, François R, Henry L, Müller K (2022). dplyr: A Grammar of Data Manipulation. https://dplyr.tidyverse.org, https://github.com/tidyverse/dplyr.
6. Masel N, Haesendonckx S, Alexandra Papadopoulou P, Wang S, Miller E, Kosiba N, Ceney A (2023). tidytlg: Create TLGs using the 'tidyverse'. https://pharmaverse.github.io/tidytlg/main/
7. Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.
8. Sabanes Bove D, Dedic J, Kelkhoff D, Kunzmann K, Lang B, Li L, Stock C, Wang Y, James D, Sidi J, Leibovitz D, Sjoberg D (2023). mmrm: Mixed Models for Repeated Measures. https://openpharma.github.io/mmrm/.
9. https://rmarkdown.rstudio.com/ (2024)
10. Sievert C (2020). Interactive Web-Based Data Visualization with R, plotly, and shiny. Chapman and Hall/CRC, Florida.

ACKNOWLEDGMENTS

The authors acknowledge the R Training teams across statisticians and statistical programmers.

RECOMMENDED READING

- https://psiaims.github.io/CAMIS/ : CAMIS (Comparing Analysis Method Implementations in Software) is a cross-industry PHUSE DVOST working group. An important goal of this project is to demystify conflicting results between software and to help ease the transitions to new languages by providing comparison and comprehensive explanations.
- https://pharmaverse.org/ : Pharmaverse is a connected network of companies and individuals working to promote collaborative development of curated open-source R packages for clinical reporting usage in pharma, in a space where previously we would only ever have worked in silos on our own closed source and often duplicative solutions. Adopting shared solutions in this post-competitive space should ultimately ease regulatory review, resulting in bringing new treatments to patients faster.
- https://www.openstatsware.org/ : openstatsware is a scientific working group of the American Statistical Association (ASA) Biopharmaceutical section (BIOP) and a European Special Interest Group (SIG) sponsored by Statisticians in the Pharmaceutical Industry (PSI) and the European Federation of Statisticians in the Pharmaceutical Industry (EFSPI). One of their primary goals is to engineer selected R-packages to fill in gaps in the open-source statistical software landscape.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Author Name: Yannick Vandendijck
Company: Johnson & Johnson
Email: yvanden2@its.jnj.com