

Will R/Shiny provide a novel way for study teams to review clinical data?

Gustav Söderström, Early Biometrics & Statistical Innovation, Data Science & Artificial Intelligence, R&D, AstraZeneca, Gothenburg, Sweden

ABSTRACT

The open-source programming language R is experiencing increased popularity among statistical programmers in the pharmaceutical industry. Pharmaverse [1] is an excellent example, illustrating the industry's keen interest in the language where multiple companies are joining forces to create and enhance R packages for clinical reporting purposes. The ability to create interactive visualizations makes R/Shiny a popular choice for clinical and statistical programming.

R/Shiny is a tool that has the potential to streamline the work of reviewers. This paper aims to provide an overview of the tool and discuss the challenges involved in implementing it for clinical data review.

INTRODUCTION

Working as a Statistical Programmer the main responsibility does not lie in evaluating the produced Tables, Figures, and Listings (TFLs). Instead, it focuses on ensuring the correct production and representation of data. The review of clinical data can be resource and time-intensive. To expedite and simplify this process, tools like R/Shiny, which enable interactivity, have the potential to facilitate faster conclusions and clearer messaging, leading to informed decision-making.

For a long time, SAS has been the primary tool for statistical programmers. However, R has gained popularity, and many pharmaceutical companies are exploring its potential as either a replacement for or complement to SAS. A prime example of this trend is Pharmaverse[1], where multiple pharmaceutical companies collaborate to develop R packages for creation of i.e. ADaM and SDTM datasets. One significant advantage of R over SAS is its open-source nature and its capacity to create visually appealing, high-quality graphics.

R provides a package called Shiny, a web application framework that allows the creation of interactive visualizations with R code on the backend. This means that individuals with R programming skills, without the need for HTML, JavaScript, or CSS knowledge, can create web applications for reviewing clinical data. The combination of interactive data visualizations and the computational power of R makes R/Shiny a compelling tool.

REVIEW OF CLINICAL DATA TODAY

In current clinical data review processes, many outputs are presented as static visualizations, often delivered in PDF format. Reviewing these outputs is not only time-consuming but also limits the reviewer's ability to interactively explore the underlying data. To streamline the process and minimize time requirements, innovative tools are needed, and R/Shiny has the potential to be one such tool.

The R/Shiny framework offers clinical study teams the opportunity for interactive engagement with data and visualizations. This interactivity can significantly impact the interpretation of TFLs, potentially making the process less burdensome.

Moreover, the traditional approach to clinical data review often involves cumbersome manual processes. Clinicians, researchers, and data scientists may need to work through extensive printed reports or static digital documents, making it challenging to quickly identify trends, anomalies, or critical insights. R/Shiny's interactive capabilities empower users to customize their data views, apply filters, and drill down into specific data points with ease. This would not only accelerate the review process but also enhances the quality of decision-making.

R AND R/SHINY

R is an open-source programming language known for its diverse statistical and graphical techniques. It is widely used for statistical analysis and is highly extensible through various packages. R's strength lies in its ability to produce customized, high-quality plots.

Shiny, an R package, empowers users to build interactive web applications using R code. When combined with other packages, these applications can be tailored to the specific needs of end-users, such as physicians or researchers.

LEARNING R/SHINY

While mastering R and Shiny may initially seem daunting, these technologies offer a wealth of resources and learning opportunities. There are multiple ways to learn R programming, catering to a variety of learning styles. R benefits from a large online community, with resources like StackOverflow and R-specific forums being valuable sources of information. Additionally, numerous freely available online books are great resources for both beginners and experienced Shiny developers seeking more in-depth knowledge.

POTENTIAL OF R/SHINY

Using R/Shiny in clinical data review introduces an additional layer of understanding and interactivity compared to static visualizations. These applications can be customized to suit the end-user's needs or specific aspects of the data. R/Shiny also enables continuous monitoring of data as it becomes available, facilitating timely decision-making. To illustrate the possibilities of R/Shiny, refer to Figure 1, which depicts an application providing a timeline for selected subjects within a study. Such applications can assist safety delegates in reviewing adverse events and concomitant medication usage.

Another noteworthy aspect is the ease of sharing applications via cloud hosting solutions. This enables multiple collaborators to simultaneously use and share insights using the same application.

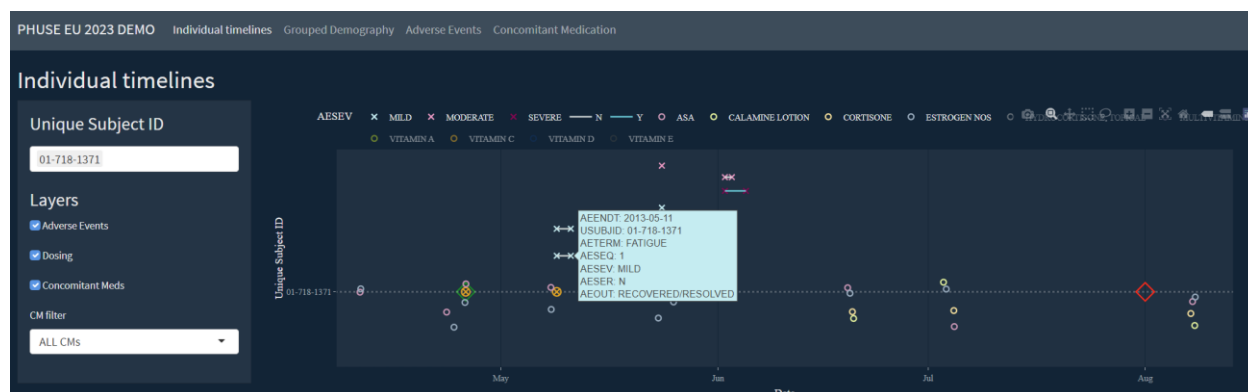


Figure 1 – Example of an R/Shiny application

CHALLENGES AND MITIGATIONS

One challenge lies in the limited level of R knowledge among Statistical Programmers. Mitigating this challenge involves offering courses and providing opportunities for statistical programmers to gain experience with R. Beginners may initially face a steep learning curve and once they grasp the syntax and fundamentals, producing R scripts and conducting analyses can be achieved within a short timeframe. Furthermore, many newcomers to the industry may already possess R knowledge from their university studies, which can be advantageous and reduce training requirements.

Another challenge involves compliance with regulatory requirements and data security standards to maintain data integrity and confidentiality. While R/Shiny does not directly address these issues, it can operate within compliant frameworks. The setup of the infrastructure plays a crucial role in ensuring compliance but falls outside the scope of this paper.

In addition to these challenges, another consideration is the seamless integration of R/Shiny applications into existing data review workflows. Successful integration requires careful planning, compatibility assessments, and potential modifications to adapt to specific clinical data review processes and requirements.

CONCLUSION

As the development of pharmaceuticals increasingly demands timely reviews, innovative solutions are essential. While there are several candidates capable of providing similar solutions to R/Shiny, further analysis is required to determine whether it is the optimal tool. However, with R now emerging as an option for producing SDTM and ADaM datasets, the potential of using the same programming language for both data creation and interactive visualization is intriguing and could become a widely adopted process in the future.

Embracing R/Shiny as a tool for clinical data review could potentially be a step toward more efficient, interactive, and collaborative data evaluation in the pharmaceutical industry. This transformative potential is promising for enhancing decision-making processes and driving pharmaceutical development forward.

REFERENCES

[1] Pharmaverse, <https://pharmaverse.org/>

RECOMMENDED READING

Mastering Shiny, Hadley Wickham, <https://mastering-shiny.org/>

CONTACT INFORMATION

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

Gustav Söderström

AstraZeneca

Pepparedsleden 1, 43183, Mölndal, Sweden

Gustav.soderstrom@astrazeneca.com

Web: <https://www.linkedin.com/in/gustav-soderstrom/>

Brand and product names are trademarks of their respective companies.