

## Using R-Shiny Capabilities in Clinical Programming to create Applications

Prakasam R, Efficacy Lifescience Analytics;  
Sujatha Kandukuri, Efficacy Lifescience Analytics;  
Tyagrajan Swaminathan, Efficacy Lifescience Analytics

### ABSTRACT

As of 2021, 399,549 clinical studies have been reported in clinicaltrials.gov. For the last 4 years, more than 30,000 studies were registered each year. Enormous amounts of data is generated from these studies and a number of pharmaceutical companies have started initiatives to make use of open-source tools like R to analyze this data. The potential of this data can be further unleashed using interactive visualization tools that can bring out intricacies of the data and patterns within it that cannot be identified with routine tabulation methods. In addition, there are possibilities with the available tools, such as R-Shiny, to integrate automation capabilities, web deployment capabilities, and Graphic user interface based interactive visualization capabilities to create a single front-end application that can enable end-user efficiency with zero license cost.

This paper presents a use-case for clinical programming where we can build a deployable, scalable, reusable application/utility built on R-Shiny for automation of patient profiles and visualizing patient information.

This application can demonstrate how we can build an integrated solution in a single application by making use of:

- Reactive Programming available in R Shiny to create a GUI based Interactive tool
- Web deployment capabilities available in Shiny framework to deploy the developed app, accessible through web-browsers
- Programming capabilities available in R Programming for automating repetitive tasks

This can trigger ideas that can be applied in different clinical avenues and encourage the migration from the traditional licensing models to the usage of opensource frameworks.

### INTRODUCTION

Data is the veritable oil of the 21st century and is making a compulsive redefinition of business operations across industries. With the huge amount of digital data available on one hand and the fast-paced growth of technology on the other, companies look forward to leveraging the combination of data and technology to gain the maximum benefit. In addition to using validated, robust, mainstream tools that have been around for decades to analyze this data, organizations are now slowly opening to the idea of using open-source tools that have been lurking in the shadows waiting for their moment of glory. Such tools have the potential to unleash the valuable opportunities hidden in the available data.

In addition to the very apparent advantage that open-source technologies offer which is a significant cost saving in licenses, they also offer the advantage of collaboratively solving bigger problems through open-source communities. They also offer the flexibility for the programmer to develop their application using various packages and libraries, that are in constant evolution with the involvement of developer communities.

With large amounts of data, automation and interactive visualization have become key players in understanding the data. Automation is necessary to handle large amounts of data with ease and visualization enables the identification of patterns and outliers that otherwise can be missed with the use of traditional tabulation methods. This paper takes an example of how to achieve the expected business outcomes – automation and interactive visualization on the available data and using open-source technology frameworks to design a web application that can provide clinical programming users with a new level of versatility in their day-to-day data analysis work. It has the potential to improve decision

making through visualization techniques and achieve improved efficiency with automation of data analysis.

As a use case, this paper explains about a Graphical User Interface (GUI), interactive web-based application called *Patient Profile Analyzer* built on R based Open-Source technologies that generates Patient Profiles and has the ability to do visual data analysis. The application integrates automation capabilities, web deployment capabilities, and GUI based interactive visualization capabilities brought together through the use of R, R Shiny, R Markdown and related R Packages.

## PATIENT PROFILE

### WHAT IS PATIENT PROFILE IN CLINICAL PROGRAMMING?

A Patient Profile is a document that gives the complete and detailed clinical details of individual patients in a drug trial and the creation of this profile is an important activity that is part clinical programming. The input for Patient Profile would generally be a SAS® Dataset or Excel files that are the SDTM domains such as DM, AE, and VS.

The patient profile output is in the form of an RTF (Word) or PDF File, typically generated using SAS® macros. The output pdf/word files contain the complete listing of each patient details fetched from the input datasets. Generally, hundreds of patient profile reports will be created during this process. The report mainly includes treatment records, demographic characteristics, adverse events, tests for specific therapeutic areas, etc.

### WHO USES PATIENT PROFILES?

Patient profiles are routinely perused by several teams within the study for different reasons. The biostatistics team uses patient profiles to understand data patterns within a study. The data management team uses it to check for missing data and outliers. The safety team uses it to verify safety information for the drug and to ensure study compliance with established protocols. In addition, the clinical operations team, medical writing team and subject narratives teams use the patient profiles as part of various submission activities.

Patient profiles are also routinely used in Data and Safety Monitoring Board [DSMB] meetings, submission narrative writing activities, dose escalation meetings, and pre-database lock data quality reviews.

### CURRENT SCENARIO OF PATIENT PROFILE CREATION:

Majority of pharmaceutical companies and/or their partner CROs use SAS® Software for clinical programming and patient profile report generation. SAS® is currently the de facto standard for all programming work and within SAS®, the relevant macros are created/edited as per the customer requirements.

### EMERGING TREND:

R Software is emerging into the clinical industry as a potential candidate to replace SAS® and other tools for an increasing number of clinical programming and data analysis activities. Several pharmaceutical companies have started using R and is now accepted as a tool by the US Food and Drug Administration as well. R can perform most of the clinical programming activities as well as SAS® and the R software's capability of being used as an automation tool and web-application (Shiny Apps) make it a more suitable choice for building new software applications for use within clinical programming.

### THE IDEA AND MOTIVATION FOR *PATIENT PROFILE ANALYZER*

The idea for the *Patient Profile Analyzer* came up when we were brainstorming about redefining the ways of improving current operations and efficiency within clinical programming. We wanted to create modular and reusable tools and also leverage the usage of opensource tools to save cost.

## PATIENT PROFILE ANALYZER

### INTRODUCTION TO THE *PATIENT PROFILE ANALYZER*

*Patient Profile Analyzer* is a utility tool developed with R and R-related packages particularly R Shiny and R Markdown. It automates the process of creating the Patient Profile documents (pdf/rtf) from Excel/SAS® input files. This tool is aimed for the process of patient profile creation with a user interface driven approach and makes the Patient Profile creation coding effort minimal.

### PATIENT PROFILE ANALYZER- FUNCTIONAL FLOW

The flowchart in Figure 1 outlines the functional flow of the analyzer.

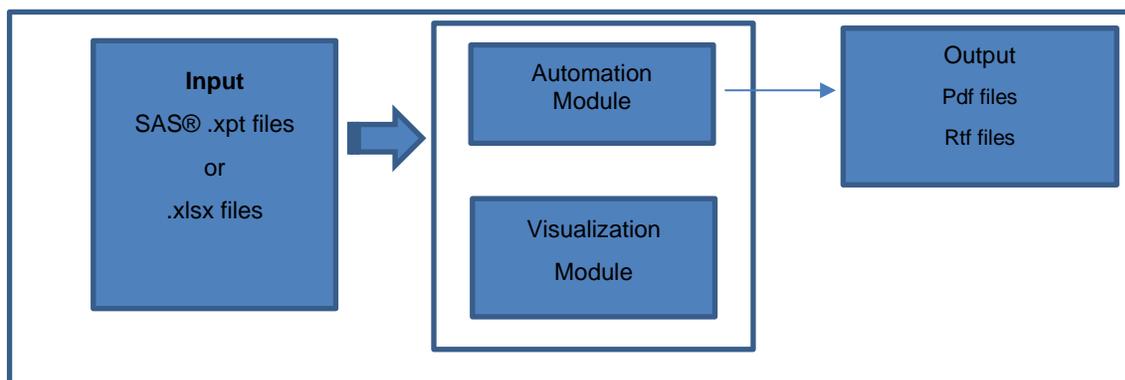
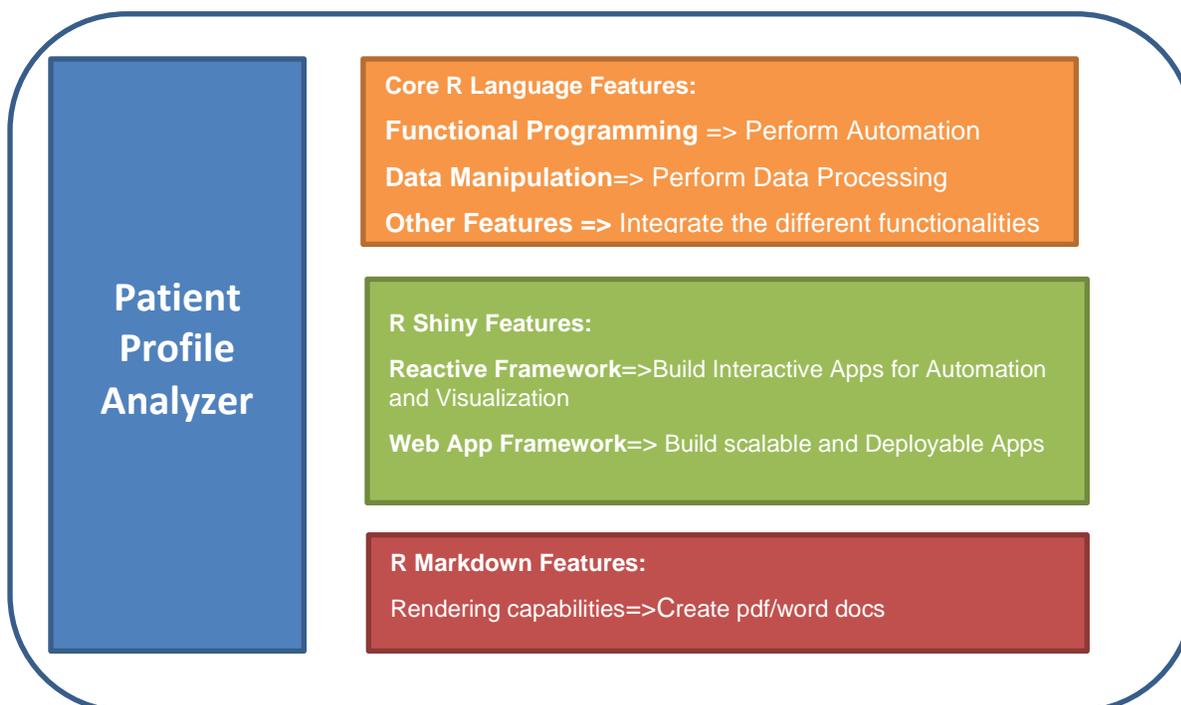


Figure 1: Functional Flow

### PATIENT PROFILE ANALYZER- TECHNICAL COMPONENTS

Figure 2 below depicts the important features of the *Patient Profile Analyzer* that are integrated in the creation of the automation framework.



## Figure 2: Patient profile analyzer components

The important technical components that make up the tool are:

- Automation Capabilities of R with its enormous number of library functions
- R Shiny's App building feature with its interactive reactive constructs
- RMarkdown's capability to generate dynamic content documents

## PATIENT PROFILE ANALYZER- TECHNICAL ARCHITECTURE

Figure 3 represents the flow of information with the analyzer. The integration of different R packages makes it possible to bring out a solution for generation of Patient Profile and do some interactive visualizations. The R Shiny App is built with its essential UI and server R components. The front end is designed and controlled by the UI module and the functional coding logic is incorporated in the server module. There is also a separate Markdown program (.rmd) that does the document generation part. The data manipulation and other coding logic is performed with R inside the App code through functions. The application is designed to run with the Shiny generated user interface to load the input files, perform interactive visualizations, apply filters, generate the files, based on the user requirement.

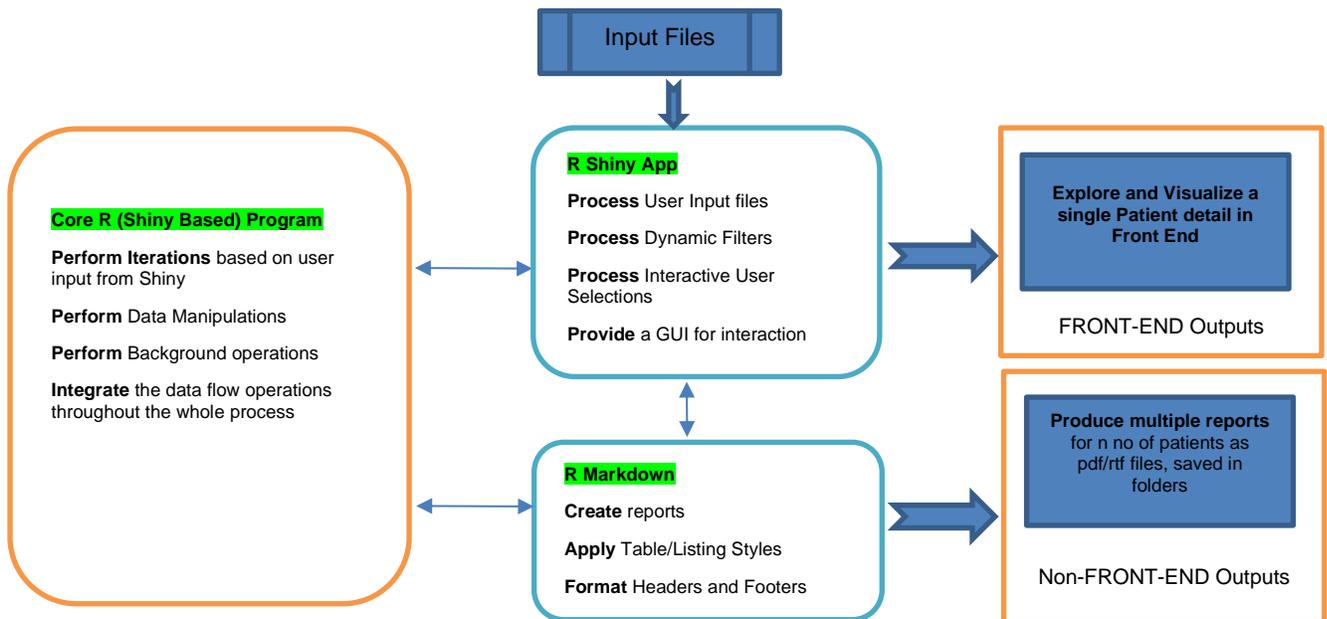


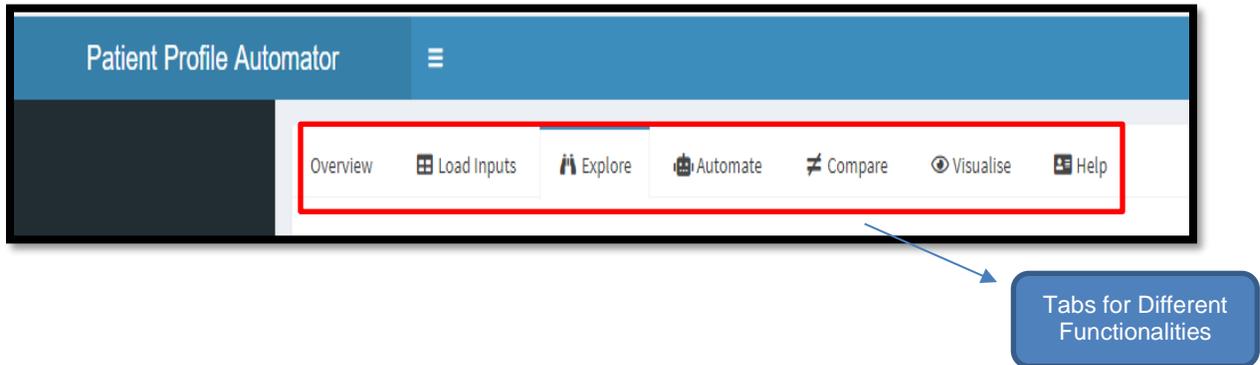
Figure 3: Flow of information representation

## PATIENT PROFILE ANALYZER WORKFLOW:

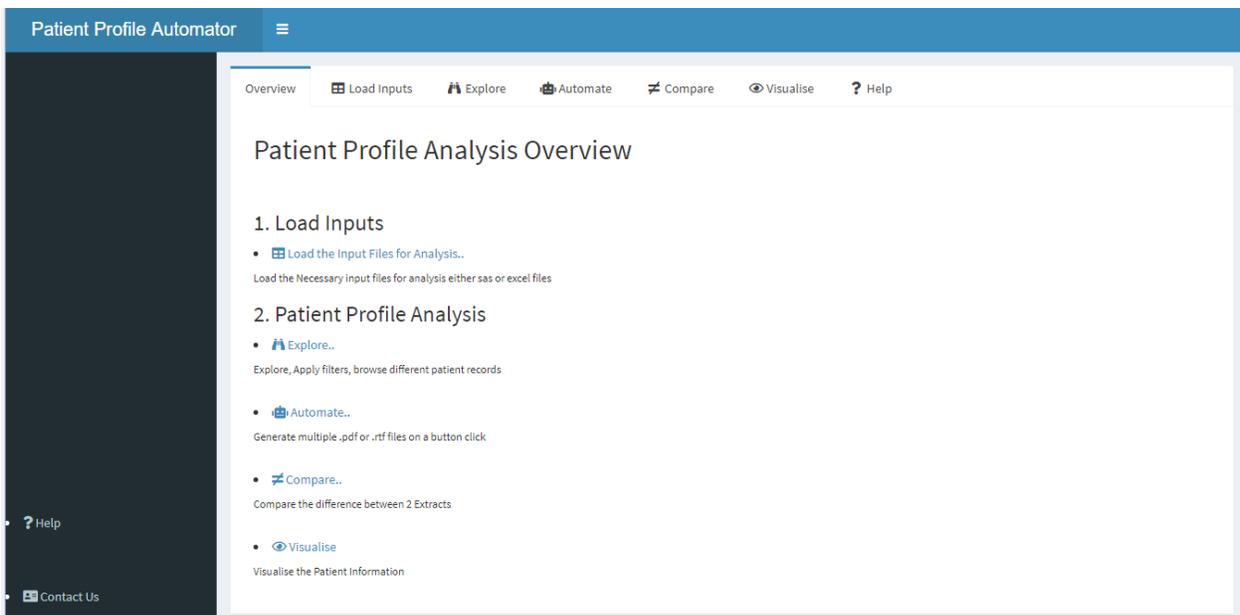
The main menu of the *Patient Profile Analyzer* shown in Figure 4 contains tabs to direct the user to the various actions possible in the app. The tabs are:

- Overview (Home tab)
- Load Inputs
- Explore
- Automate
- Visualize
- Compare

Each of the tabs are explained in detail in the rest of the paper.



**Figure 4: Patient profile analyzer features**



**Figure 5: Application layout**

The Overview page gives an outline of the different applicable functionalities

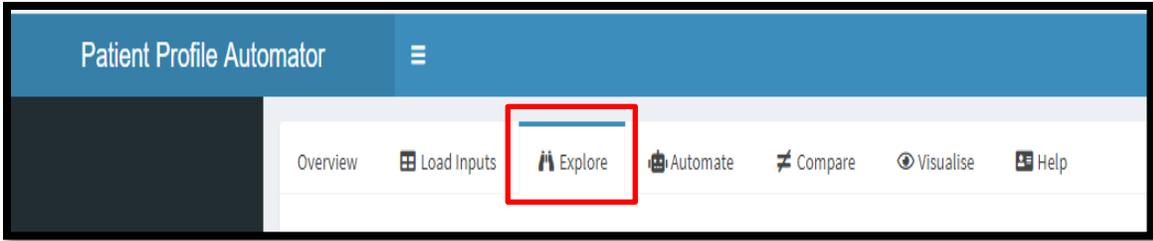
Loading Input files is the first step towards using the application, when user clicks on “Load Inputs” tab, the user can select the input files through the file browser dialog as shown in Figure 6. A preview of the files is also shown to the user.

SAS Files	Excel Files
<ul style="list-style-type: none"> <li>adae.xpt</li> <li>advs.xpt</li> </ul>	<ul style="list-style-type: none"> <li>AE.xlsx</li> <li>CM.xlsx</li> <li>DM.xlsx</li> <li>EX.xlsx</li> <li>VS.xlsx</li> </ul>

The screenshot shows the 'Patient Profile Automator' application interface. The 'Load Inputs' tab is selected and highlighted with a red box. Below the navigation bar, there is a section titled 'Choose the type of file.' with a dropdown menu currently set to 'Excel'. Below this is a file browser area showing 'Browse...' and '5 files' with an 'Upload complete' button. A callout box points to the dropdown menu with the text 'Dropdown to Choose input file type'. Another callout box points to the file browser area with the text 'File Browser to choose file'. Below the file browser, there is a note: 'Note: Use Ctrl button to choose multiple files'. At the bottom, a data preview is shown as a tibble with 11,077 rows and 7 columns. The columns are SUBJECT, AETERM, AESTDAT, AEENDAT, AEONGO\_RAM, AEOU, and AEREL. The data shows patient records with various terms and dates.

**Figure 6:Input files and ‘Load Inputs’ module**

Once the input files are loaded, user has the option to explore individual patient listings by clicking on the “Explore” Tab as shown in Figure 7 and then choose the desired Patient Id from the drop-down. The application also provides the feature to navigate to the previous and next records which makes it more interactive.



## Choose Patient..

Choose the Subject Id:



Subject	Adverse.Event.Term	Adverse.Event.Start.Date	Adverse.Event.End.Date	Is.Adverse.Event.Ongoing	Outcome.of.Adverse.Event	Causality
0011301	hypertension	04-Nov-2020	30-Nov-2020	Yes	Recovered/Resolved	Yes!Possible
0011301	Bullous dermatitis	13-Nov-2020		No	Not Recovered/Not Resolved	It is not Realated
0011301	Hypertension	01-Dec-2020	10-Jan-2021	Unknown	Recovered/Resolved	It is not Realated
0011301	Hypertension	11-Jan-2021	18-Jan-2021	Yes	Recovered/Resolved	Yes!Possible
0011301	Atrial Fibrillation	11-Jan-2021	18-Jan-2021	Yes	Recovered/Resolved	It is not Realated
0011301	Hot Flashes	18-Jan-2017	03-Mar-2018	Yes	Recovered/Resolved	Probable
0011301	Hypercalcemia	06-Mar-2017	17-Aug-2017	Yes	Recovered/Resolved	It is not Realated
0011301	Weight Gain	25-May-2017	17-Aug-2017	Yes	Recovered/Resolved	It is not Realated
0011301	Weight Gain	17-Aug-2017	12-Sep-2017	Yes	Recovered/Resolved	Yes!Possible
0011301	Hypertension	30-Jan-2018	02-Jan-2019	Yes	Recovered/Resolved	It is not Realated
0011301	Hypertension	03-Jan-2019	03-Nov-2020	Yes	Recovered/Resolved	It is not Realated

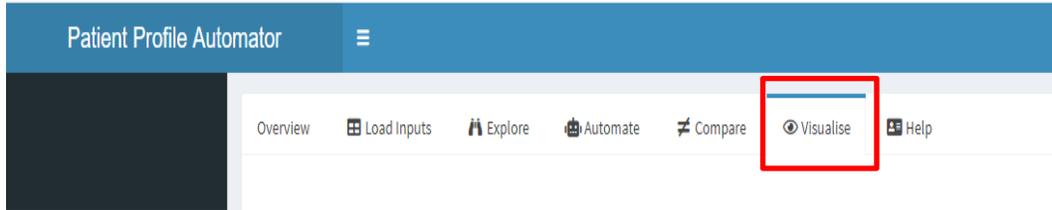
Figure 7: Explore module

Listing of the chosen subject

The results table will display for all the safety panels as separate tables. Figure 7 is just an example listing that shows the adverse event details

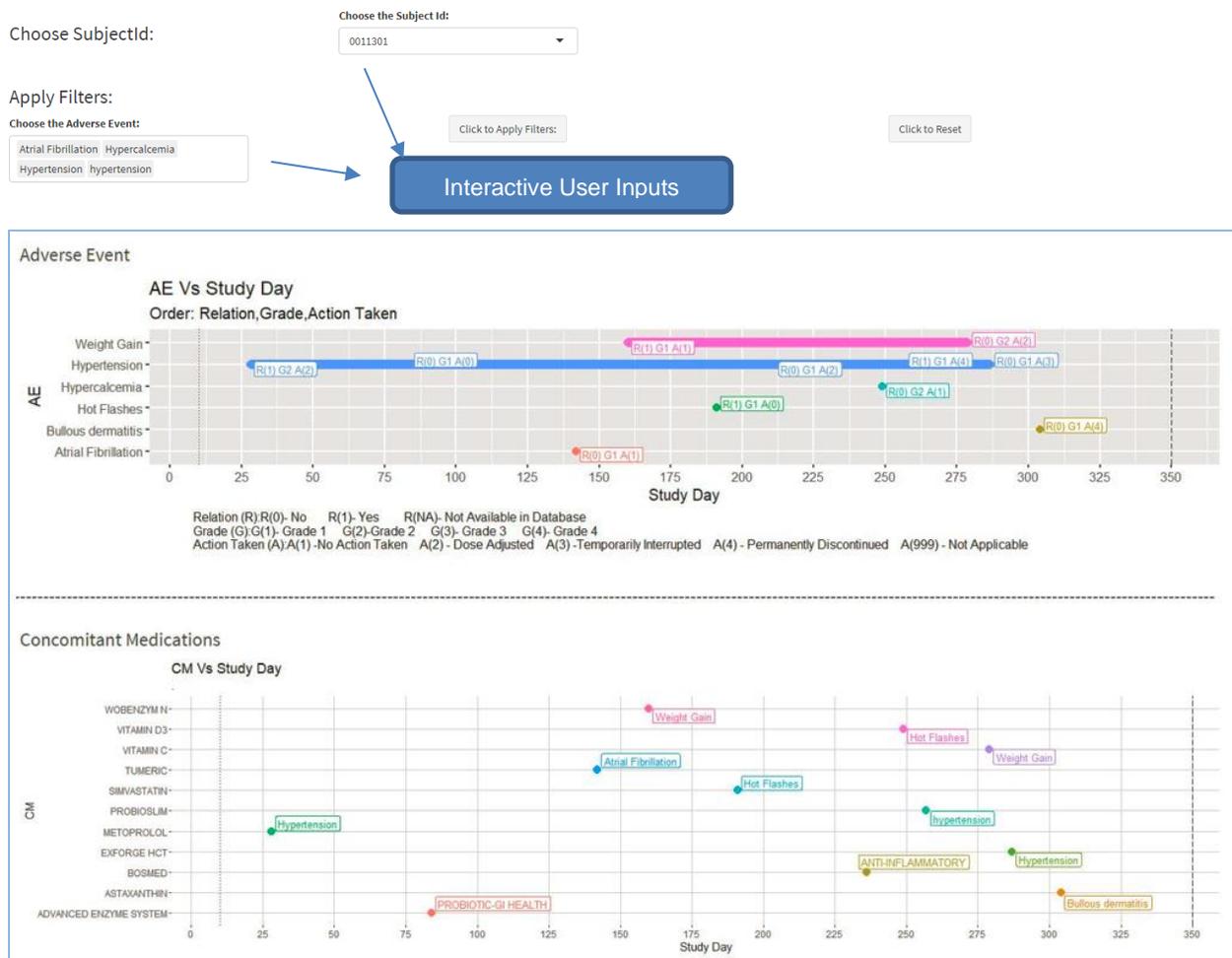
There is also a separate functionality to compare Patient Data extracts taken at 2 different time intervals and understand the difference between them. The compare module helps to achieve that by showing the difference between the data, highlighted in the front end.

As shown in Figure 8 user can click on the “Visualize” tab, choose patient and then apply the desired filters to view an interactive patient narrative as shown in Figure 9.



**Figure 8: Visualise module**

**e-Narrative Interactive plot for a Patient**



**Figure 9: Visual narrative of Adverse Event and Concomitant Medication taken by the patient over time**

The “Automate” tab allows the user to create a patient profile in PDF or RTF format as shown in Figure 10.

The screenshot shows the 'Patient Profile Automator' application. The 'Automate' tab is highlighted in the top navigation bar. Below the navigation bar, there are several tabs: Overview, Load Inputs, Explore, Automate, Compare, Visualise, and Help. The 'Automate' tab is selected and highlighted with a red box.

The main interface is divided into two sections. The top section is a configuration area with a light green background. It contains three dropdown menus: 'Select Domain' (set to 'AE'), 'Select Criteria' (set to 'AdverseEventTerm'), and 'Select Values' (containing 'hypertension', 'Bullous dermatitis', 'Atrial Fibrillation', 'Weight Gain', and 'Hot flashes'). Below these are two buttons: 'Generate rtf Files' and 'Generate pdf Files'. Three blue callout boxes point to the 'Select Domain', 'Select Criteria', and 'Select Values' dropdowns, labeled 'Choose Domain', 'Choose Filter Condition', and 'Choose Filter Value' respectively.

The bottom section is a table with a light green background. It has the following columns: Subject, AdverseEventTerm, Adverse.Event.Start.Date, Adverse.Event.End.Date, Is.Adverse.Event.Ongoing, Outcome.of.Adverse.Event, and Causality. The table contains 13 rows of data.

Subject	AdverseEventTerm	Adverse.Event.Start.Date	Adverse.Event.End.Date	Is.Adverse.Event.Ongoing	Outcome.of.Adverse.Event	Causality
0011301	hypertension	04-Nov-2020	30-Nov-2020	Yes	Recovered/Resolved	Yes/Possible
0011301	Bullous dermatitis	13-Nov-2020	NA	No	Not Recovered/Not Resolved	It is not Realited
0011301	Atrial Fibrillation	11-Jan-2021	18-Jan-2021	Yes	Recovered/Resolved	It is not Realited
0011301	Weight Gain	25-May-2017	17-Aug-2017	Yes	Recovered/Resolved	It is not Realited
0011301	Weight Gain	17-Aug-2017	12-Sep-2017	Yes	Recovered/Resolved	Yes/Possible
0011302	Hot flashes	30-May-2019	NA	No	Not Recovered/Not Resolved	Probable
0011303	Hot flashes	21-Aug-2018	NA	No	Not Recovered/Not Resolved	Probable
0011304	hypertension	21-Aug-2019	12-Nov-2019	Yes	Recovered/Resolved	It is not Realited
0011304	hypertension	13-Nov-2019	04-Feb-2020	Yes	Recovered/Resolved	It is not Realited

The second part of the screenshot shows a configuration screen for generating patient profiles. It has a light green background and the title 'Automatically Generate the Patient Profile .pdf or .rtf files'. Below the title is the text 'Based on Demographic File'. There is a text input field labeled 'For how many Patients?' with the value '5'. A blue callout box points to this field with the text 'Enter the No of files to be generated Example: 5, creates records for 5 Patients'. Below the input field are two buttons: 'Generate Pdf Files' and 'Generate rtf Files'. A blue callout box points to these buttons with the text 'Buttons to generate files'. At the bottom, there is a note: 'Note: The files will be generated based on the Demographic File. First 'n' Patient files will be generated'.

Figure 10: Automate module to generate patient profiles

**OUTPUT:**

The tool processes the input files and generates the patient profile listings of each patient. Multiple files are created based on the number of records/patients present in the demographic Input file. The generated outputs appear as shown in Figure 12.

Type of output	Sample Files produced by Tool														
Pdf output	<table border="1"> <thead> <tr> <th>Name</th> <th>Type</th> </tr> </thead> <tbody> <tr> <td>PP_0011301</td> <td>Adobe Acrobat Document</td> </tr> <tr> <td>PP_0011302</td> <td>Adobe Acrobat Document</td> </tr> <tr> <td>PP_0011303</td> <td>Adobe Acrobat Document</td> </tr> <tr> <td>PP_0011304</td> <td>Adobe Acrobat Document</td> </tr> <tr> <td>PP_0011305</td> <td>Adobe Acrobat Document</td> </tr> </tbody> </table>	Name	Type	PP_0011301	Adobe Acrobat Document	PP_0011302	Adobe Acrobat Document	PP_0011303	Adobe Acrobat Document	PP_0011304	Adobe Acrobat Document	PP_0011305	Adobe Acrobat Document		
Name	Type														
PP_0011301	Adobe Acrobat Document														
PP_0011302	Adobe Acrobat Document														
PP_0011303	Adobe Acrobat Document														
PP_0011304	Adobe Acrobat Document														
PP_0011305	Adobe Acrobat Document														
Word output	<table border="1"> <thead> <tr> <th>Name</th> <th>Type</th> </tr> </thead> <tbody> <tr> <td>0011301</td> <td>Rich Text Format</td> </tr> <tr> <td>0011302</td> <td>Rich Text Format</td> </tr> <tr> <td>0011303</td> <td>Rich Text Format</td> </tr> <tr> <td>0011304</td> <td>Rich Text Format</td> </tr> <tr> <td>0011305</td> <td>Rich Text Format</td> </tr> <tr> <td>0011306</td> <td>Rich Text Format</td> </tr> </tbody> </table>	Name	Type	0011301	Rich Text Format	0011302	Rich Text Format	0011303	Rich Text Format	0011304	Rich Text Format	0011305	Rich Text Format	0011306	Rich Text Format
Name	Type														
0011301	Rich Text Format														
0011302	Rich Text Format														
0011303	Rich Text Format														
0011304	Rich Text Format														
0011305	Rich Text Format														
0011306	Rich Text Format														

**Figure 11: Patient profile output files**

The count of files produced depends on the value entered in “For how many Patients” field

### Contents of Output:

Separate rtf/pdf files will be generated for each patient (count of files depends on the number of patients in the DM file). Each file will contain separate tables (AE, VS, EX) based on the given input content

Study: My Sample Study Study Country:India TreatmnetArm:Single Arm Study						
Subject Id :0011301 DateOfBirth:1943-09-14 Ethnic :Not Hispanic or Latino Race :WHITE						
Autor: PR						
Adverse Events (AE)						
Subject	Adverse Event Term	Adverse Event Start Date	Adverse Event End Date	Is Adverse Event Ongoing	Outcome of Adverse Event	Causality
0011301	hypertension	03-Nov-1950	29-Nov-1950	No	Recovered/Resolved	Possible
0011301	Bullous dermatitis	12-Nov-1950		Yes	Not Recovered/Not Resolved	Not Related
0011301	Hypertension	30-Nov-1950	09-Jan-1951	No	Recovered/Resolved	Not Related
0011301	Hypertension	10-Jan-1951	17-Jan-1951	No	Recovered/Resolved	Possible
0011301	Atrial Fibrillation	10-Jan-1951	17-Jan-1951	No	Recovered/Resolved	Not Related
0011301	Hot Flashes	17-Jan-1947	01-Mar-1948	No	Recovered/Resolved	Probable
0011301	Hypercalcemia	05-Mar-1947	16-Aug-1947	No	Recovered/Resolved	Not Related
0011301	Weight Gain	24-May-1947	16-Aug-1947	No	Recovered/Resolved	Not Related
0011301	Weight Gain	16-Aug-1947	11-Sep-1947	No	Recovered/Resolved	Possible
0011301	Hypertension	29-Jan-1948	31-Dec-1948	No	Recovered/Resolved	Not Related
Created on: 26-03-2022 12:52:45						
Page 1 of 11						

**Figure 12:Snippet of patient profile generated**

## PATIENT PROFILE CREATION - APPLICABLE FEATURES

While creating Patient Profiles, the below features are the most important aspects that are mostly configured via the coding.

Feature	Possible with Tool?	Comments
Table Listing cell formatting	Yes	The table styles can be customised
Column Selections	Yes	Users have the option to select /shortlist only the required columns
Column Renaming	Yes	Column names can be changed via the tool. E.g.: Patient shall be changed as Subject
Column Values Encoding	Yes	Column values shall be encoded. Ex: 0=>No; 1=>Yes.
Addition of derived variables	Partial	Partial generic variables like Age Calculation functions, Concatenator functions can be added
Header/Footer Customisation	Yes*	User can pass dynamic header content through the tool

\*Some of the features mentioned in the tool is under Work in Progress

## BENEFITS OBTAINED

In summary, the *Patient Profile Analyzer* is a reusable tool that can be used across projects.

It is deployable as a Web App and can be accessed from anywhere. It is scalable when deployed as a web application for any number of users. The deployment process in Shiny is very easy and user friendly, in comparison to other Web App frameworks This tool can also be used as a local app that runs on the local network.

In most cases, minimal coding knowledge is needed as there is no need to edit the code often (except for very customized functionalities). The user-friendly framework allows easy interaction and greater productivity for a clinical programmer.

## LIMITATIONS/CHALLENGES

The *Patient Profile Analyzer* is currently built as a proof-of-concept and not a complete automation suite. In some cases, there may be a need to do some enhancements. As of now, it is designed only for a simple patient profile analysis. Enhanced features would require additional design changes and R knowledge would be required to make any customized changes to the tool.

Thorough testing of the tool should be carried out for making it a fully validated tool that can be used for submission activities. Additional styling needs to be applied on the output files to manage formatting issues such as text wrapping.

## CONCLUSION

R and other open-source technologies have been evolving and maturing over the past few decades. It is currently at a stage where several organizations including pharmaceutical companies are ready to adopt these technologies to reduce licensing costs for tools that are well established in the industry. Such open-source technologies can help build systems that can be a replacement for the current tools and packages while still maintaining the original processes and methodologies.

The "*Patient Profile Analyzer*" tool discussed in this paper is one such example of a tool created in R that can be used for automating patient profiles and for viewing patient profiles interactively. It uses many of the packages available with R such as Shiny, RMarkdown, etc. In addition, it offers the scope of deploying the application on the web so that multiple users can access it seamlessly. The initial proof-of-concept demonstrated here can easily be expanded to create a validated system that can be used across the clinical programming team to improve productivity of the team.

## REFERENCES

<https://www.clinicaltrials.gov/>

<https://shiny.rstudio.com/>

<https://rmarkdown.rstudio.com/>

## ACKNOWLEDGMENTS

Many thanks to the Efficacy Leadership Team, who were very encouraging in the process of making this tool. Our heartfelt thanks to Manohar Sure from Efficacy whose ideas for the enhancement of the tool were very apt. And finally, our gratitude to the entire '*Patient Profile Analyzer*' Team who made the tool possible!

## CONTACT INFORMATION

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

Prakasam R  
Efficacy Lifescience Analytics  
[prakasam.r@efficacy.com](mailto:prakasam.r@efficacy.com)

Sujatha Kandukuri  
Efficacy Lifescience Analytics  
[sujatha.kandukuri@efficacy.com](mailto:sujatha.kandukuri@efficacy.com)

Tyagrajan Swaminathan  
Efficacy Lifescience Analytics  
[tyagrajan.swaminathan@efficacy.com](mailto:tyagrajan.swaminathan@efficacy.com)