

Overview Projects Timeline to Arrange Statistical Programmer Resource via R Shiny

KuenHung Lin, Pharmaceutical Co., Ltd, Taipei, Taiwan

ABSTRACT

For the clinical trial industry, it is always full of competitive tasks in a company. Hence, one of the most important duty of line manager is overview multiple projects timeline to arrange members' resource. Although, there are several application packages to manage project timeline. Most of those applications can't be customizable. This article provides an interactive web application to visualize projects timeline via R Shiny which can be customized by statistical programmer who has experience in R. Hence, many customizable information can be added in the application, such as descriptive statistics for resource per month and making key points stand out freely.

INTRODUCTION

Timelines are a fundamental tool for effective clinical project management, finding use at all levels of study execution from participant-level tracking to study-wide oversight [1]. Unfortunately, the process of creating timelines is a tedious one that sees many timeline users resort to highly manual tools that lack rigorous data integration and flexibility of customization, like spreadsheets and project management systems. Hence, we would like to introduce an interactive web application to visualize projects timeline via R Shiny which can be customized by statistical programmer who has experience in R. The choice of the specific programming language and packages is considered because of the existing R knowledge from statisticians and statistical programming [2].

The rest of this article is organized as follows. First, we use one example to describe the motivation of this article. Second, we introduce the R packages used and the instruction of proposed web application. Besides, we apply a examples to illustrate the output of proposed web application. Finally, some concluding remarks are given at the end of this article.

FRAMEWORK AND MOTIVATION

In this section, we use below example to explain the motivation of this article. Even among tools that offer a more robust, programmatic approach like SAS and R, the output is typically static with limited ability for end users to explore the data in real time [1]. Figure 1 shows it is difficult to clearly display span of one year project timeline in a static R output while minor tick is displayed for each month.

Furthermore, it is not convenient to extend or update timeline without interactive platform. To overcome above limitation and inconvenience, we would like to illustrate proposed web application in the next section.

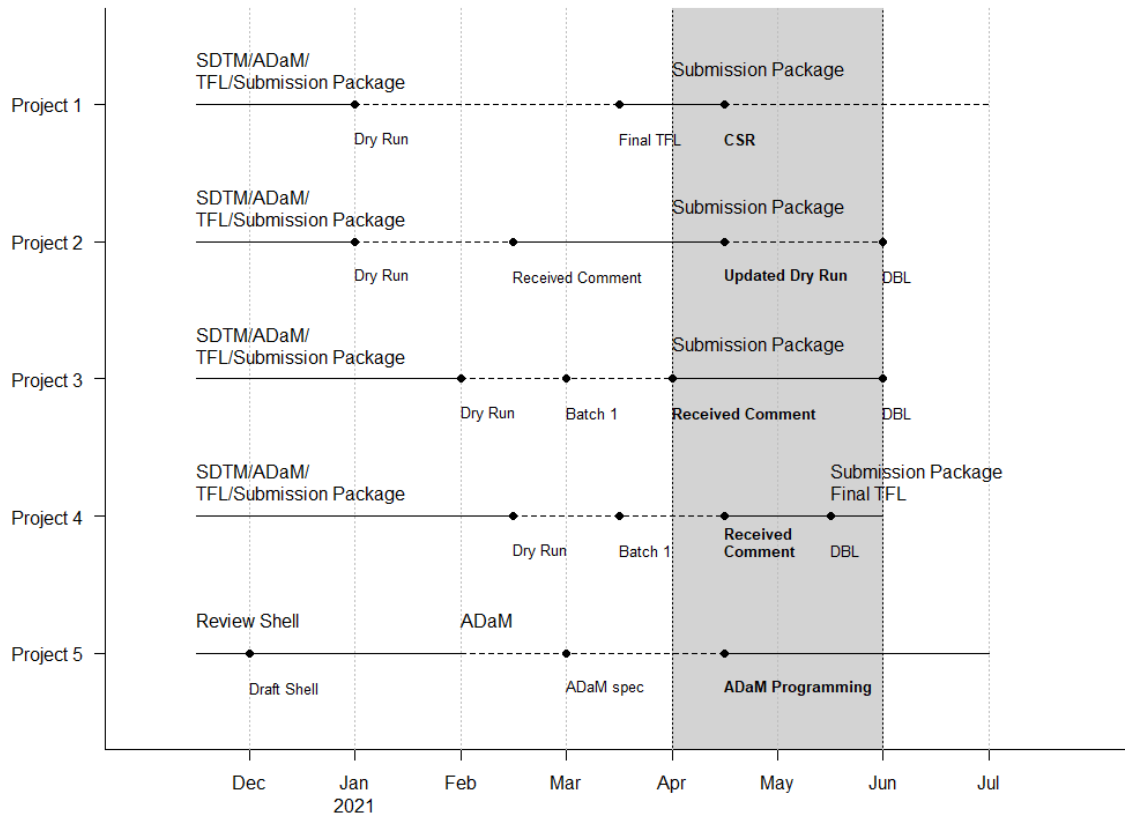


Figure 1. Static timeline execution period highlighted with gray

THE SHINY APPLICATION

To introduce the Shiny application, we use three subsections, including user interface (UI), server outputs and R packages. First, we illustrate how to add timeline information via UI. And then, we demonstrate server outputs, including editable timeline table and visualization of timeline. Furthermore, we share used R packages of the application and highlight the purpose of those packages in the application.

USER INTERFACE (UI)

There are two ways to upload timeline information to the application. One is creating one or multiple timeline information in EXCEL file and the other is adding each timeline information via UI. No matter which one is used, users need to provide the same timeline information. Figure 2 shows the structure of EXCEL file.

	A	B	C	D	E	F
1	project	task	start_dt	end_dt	note	timepoint
2	Project 1	Final TFL	2020-12-01	2021-04-30	DBL	2021-02-01
3	Project 1	Final TFL	2020-12-01	2021-04-30	Delivered to Medical Writer	2021-02-19
4	Project 1	Submission Package	2021-05-01	2021-07-31	Biostat Review	2021-06-23
5	Project 1	Submission Package	2021-05-01	2021-07-31	Biostat 2nd Review	2021-07-05
6	Project 2	Dry Run	2020-12-01	2021-02-05	Biostat Review	2021-02-05
7	Project 2	Dry Run	2020-12-01	2021-02-24	Delivered to Medical Reviewer	2021-02-24
8	Project 2	Final TFL	2021-10-01	2021-10-31	SAP/Shell finalized	2021-10-04
9	Project 2	Final TFL	2021-10-01	2021-10-31	DBL	2021-10-29
10	Project 3	Offline Listing	2020-12-01	2021-06-30	1st delivery to CDA	2021-01-20
11	Project 3	SDTM	2020-12-01	2021-06-30	SDTM mapping spec	2021-03-05
12	Project 3	SDTM	2020-12-01	2021-06-30	SDTM programming	2021-04-06

Figure 2. Example of timeline information in EXCEL file

The explanation of each column as below,

- project: It typically will be protocol number or company codename of protocol.
- task: For each project, it will contain one or more tasks, such as offline listing, SDTM/ADaM/TFL and submission package.
- start_dt: The start date for the duration of each task (date format: yyyy-mm-dd).
- end_dt: The end date for the duration of each task (date format: yyyy-mm-dd).
- note: The description of stage during each task, such as internal review, deliverable and milestone.
- timepoint: The date of stage (date format: yyyy-mm-dd).

There are fields corresponding to above timeline information in UI as Figure 3.

Choose xlsx file

No file selected

Project

Task

Start Date input

End Date input

Note

Timepoint

Excel File Name

Figure 3. UI

Furthermore, the UI also contains features of import/export timeline information EXCEL file. There is a need to pay special attention that existed EXCEL file should be uploaded before adding, once adding is executed, then the field of import EXCEL file will be hidden in UI.

SERVER OUTPUTS

Figure 4 shows the editable timeline table based on above example from Figure 2.

Project Timeline Table						
Project Timeline Figure						
Show 25 entries						
Search:						
	project	task	start_dt	end_dt	note	timepoint
1	Project 1	Final TFL	2020-12-01	2021-04-30	DBL	2021-02-01
2	Project 1	Final TFL	2020-12-01	2021-04-30	Delivered to Medical Writer	2021-02-19
3	Project 1	Submission Package	2021-05-01	2021-07-31	Biostat Review	2021-06-23
4	Project 1	Submission Package	2021-05-01	2021-07-31	Biostat 2nd Review	2021-07-05
5	Project 2	Dry Run	2020-12-01	2021-02-05	Biostat Review	2021-02-05
6	Project 2	Dry Run	2020-12-01	2021-02-24	Delivered to Medical Reviewer	2021-02-24
7	Project 2	Final TFL	2021-10-01	2021-10-31	SAP/Shell finalized	2021-10-04
8	Project 2	Final TFL	2021-10-01	2021-10-31	DBL	2021-10-29
9	Project 3	Offline Listing	2020-12-01	2021-06-30	1st delivery to CDA	2021-01-20
10	Project 3	SDTM	2020-12-01	2021-06-30	SDTM mapping spec	2021-03-05
11	Project 3	SDTM	2020-12-01	2021-06-30	SDTM programming	2021-04-06
Showing 1 to 11 of 11 entries						
Previous						Next

Figure 4. Editable timeline table

In addition, Figure 5 shows the visualization of timeline. There are some notices before reviewing the plot as below,

- The project will be displayed based on the order of column “project” in the timeline table. Furthermore, the individual line will be created for each task per project. Once there are more than one tasks in a project, the line with earlier start date of task duration will be drawn at the same y-axis level of project, the later one will be displayed a bit higher than the y-axis level of project. If there are multiple tasks with the same start date of task duration, alphabetical order will be considered.
- The timeline information from column of “note” and “timepoint” are applied to create hover text. Once users move mouse to the target date, the competitive tasks which date of stage are close will be displayed at the same time to remind users. Besides, users can choose the note(s) which want to be highlighted with color in UI once any timeline information is uploaded to the application.
- In order to highlight the current execution period, there will be a rectangle with color of light blue from one month before current month to two months later.

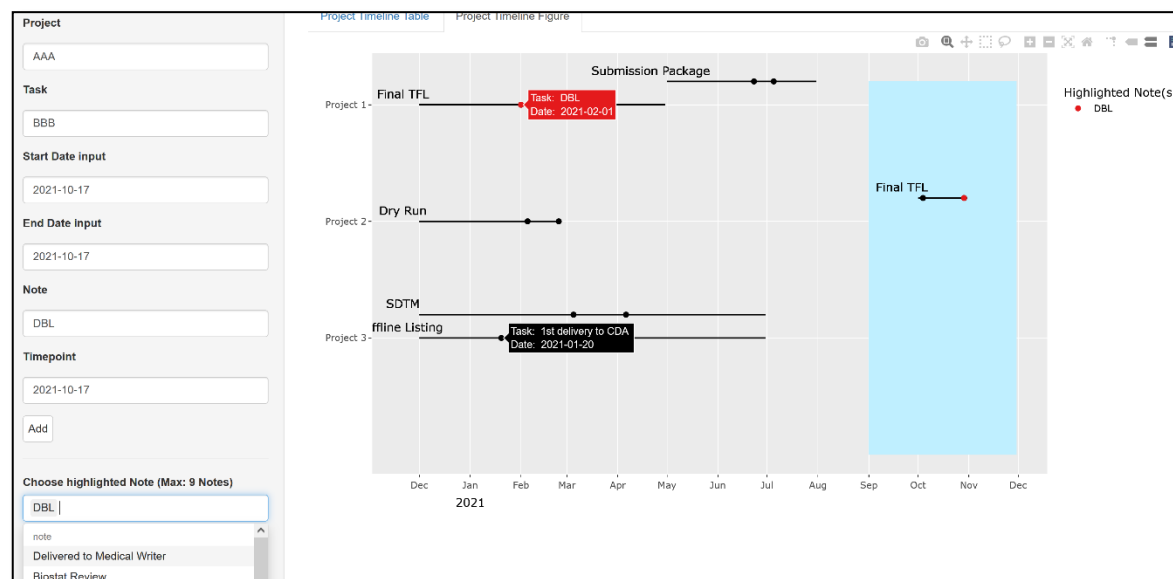


Figure 5. Visualization of timeline

R PACKAGES

In this section, we will use Table 1 to introduce those R packages which are used in the application and highlight the purpose of those packages in the application.

Table 1. Introduction to R packages in the application

Package	Purpose
shiny	It provides a flexible framework for creating interactive web applications with R
DT	Creating editable table of timeline
openxlsx	Import existed timeline EXCEL file
readxl	Export timeline table which is edited via the application as a timeline EXCEL file
dplyr	"dplyr is a grammar of data manipulation" by https://dplyr.tidyverse.org/
lubridate	Making it easier to work with dates and times from timeline information
ggplot2	Using it to draw the graph which can be simply connect with plotly package
plotly	Making interactive timeline graph from ggplot2 via function of ggplotly
RColorBrewer	Setting the list of color to highlight specific note from timeline information

CONCLUSION

Comparing Figure 1 and Figure 5, the limitation and inconvenience that is mentioned as the motivation of this article overcomes via the proposed application. To visit <https://kuenhunglin.shinyapps.io/TimelineOverview/> and start to arrange timeline. It will dramatically reduce users' time. Besides, programs of the application are shared on [GitHub](#). The example EXCEL file is also out there, make them help you to be familiar with the application. For the next step, the file of working hours is one of the most import sources for arranging resource should also be considered in the application. Hence, further development of the application is to add the tab of descriptive statistics for resource per month to connect with visualization timeline tab. To let arrange timeline more precise and efficient.

REFERENCES

- [1] S. Childress, J. Wildfire, R. Bailey and B. Sikora, "Clinical Timelines Visualized" *Proceedings of the PhUSE US Connect 2018*, paper DV01. Available at <https://www.lexjansen.com/phuse-us/2018/dv/DV01.pdf>.
- [2] N. Masel, S. Haesendonckx and A. Papadopoulou, "An Automation Proof of Concept of Periodic Reporting via R Shiny" *Proceedings of the PhUSE EU Connect 2019*, paper AD03. Available at <https://www.lexjansen.com/phuse/2019/ad/AD03.pdf>.
- [3] Graham M. Wheeler, Michael J. Sweeting and Adrian P. Mander, "AplusB: A Web Application for Investigating A + B Designs for Phase I Cancer Clinical Trials" *PloS One*. 2016; 11(7): e0159026. doi: [10.1371/journal.pone.0159026](https://doi.org/10.1371/journal.pone.0159026) PMID: [27403961](https://pubmed.ncbi.nlm.nih.gov/27403961/).

ACKNOWLEDGMENTS

The author would like to thank all participants in [TSDC](#) (Tamkang Statistics Data Club) that expand my horizons of visualization area and inspire me to keep my way on learning to be a data scientist.

CONTACT INFORMATION

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

Author Name: KuenHung Lin

Company: Athenex Pharmaceuticals Co., Ltd.

Address: 19F., No.460, Sec. 4, Xinyi Rd., Xinyi Dist., Taipei City 110, Taiwan (R.O.C.)

Email: kuenhunglin@athenex.com

Web: <https://github.com/liinhater>

Brand and product names are trademarks of their respective companies.