Use of R-Scripts Can Help to Decrease Time and Improve Accuracy on Summary Tables for IND and Semi-Annual Reports

Brett Palmer BS, CCRC; Anlil Brikha, BS; Fang Lin, MS; Jill Woodman BS, CCRP

Robert H. Lurie Comprehensive Cancer Center of Northwestern University

Abstract

Annual IND and semi-annual reports (SARs) required manual generation of tables with data including All Adverse Events (AEs), Related AEs, and demographics. This process was time consuming and open to the possibility of errors. Using our home-grown clinical trial management system, the Quality Assurance (QA) team worked with the developers to implement R-Scripts for automatic table generation to minimize manual data manipulation. After the R-Scripts were implemented, the team tested the current QA team on the previous method of compiling tables by hand and using the new R-Script method. The process for each was timed and tables were collected and assessed for accuracy. Time data was analyzed, and tables were made with error rates, and graphs with time to complete each task. Both error rate and time decreased when using the R-Scripts for generation of summary tables.

Introduction

The Robert H. Lurie Comprehensive Cancer Center (RHLCCC) is responsible for submitting roughly 50 annual IND reports to the FDA and 180 SARs to the IRB. Creating these reports involves generating summary tables for All AEs, Related AEs, and demographics. On average, these three tables take 142.4 minutes to complete for one IND report. SARs require only the demographics table averaging 11.6 minutes per report. There is variation in time required to complete each table due to differing Excel abilities within the team. Further, the process is error prone because of the amount of manual work required. The RHLCCC spends roughly 153 hours a year on table creation for SARs and IND reports. Custom programmatic R-scripts were written to automatically calculate and complete each of the tables. R-scripts remove the manual work required, thus removing human error, and reducing the time spent on the task. These scripts also enable everyone to complete the tables regardless of Excel skills.

Methods and Materials

The R-Scripts were integrated into our home-grown clinical trial management system utilizing a sandboxed environment on a remote server. This method was chosen to ensure the security and protection of patient data. The testing was done on a curated set of demographics data, pulled from an old study and all patient identifiers were redacted. The AE data also utilized a curated set of AE data, for ten patients, and all patient identifiers were redacted. Four Quality Assurance Monitors (QAMs) (n=4) responsible for SAR and IND reporting were timed on a standardized version of creating each table before and after R-Script implementation. Their work was collected and compared with an answer key to check for accuracy. Errors in calculations and transcriptions were counted and tabulated into tables.

Table Type	Demographics	All AEs	Highest AEs
Number of Errors	13	6	4

Table 1. Summary of errors in tables generated by hand

Table Type	R-Script Demographics	R-Script All AEs	R-Script Highest AEs
Number of Errors	2	0	0

Table 2. Summary of errors in tables generated by R-Scripts

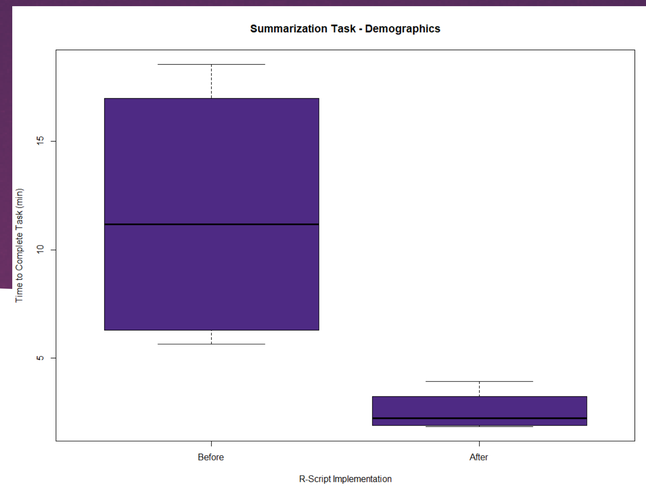

Figure 1. Summarization times for completion of Demographics Tables

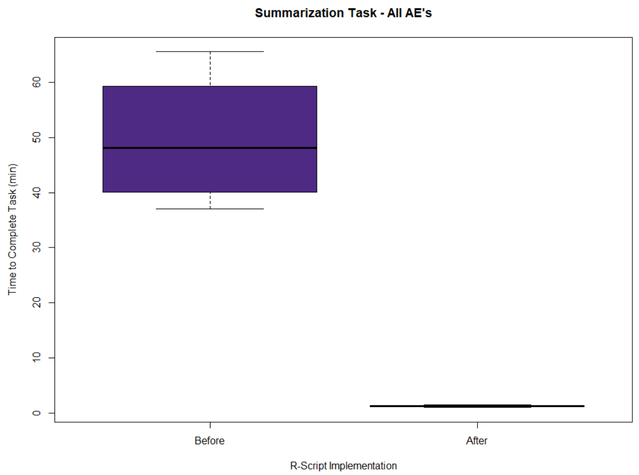

Figure 2. Summarization times for completion of All AEs Tables

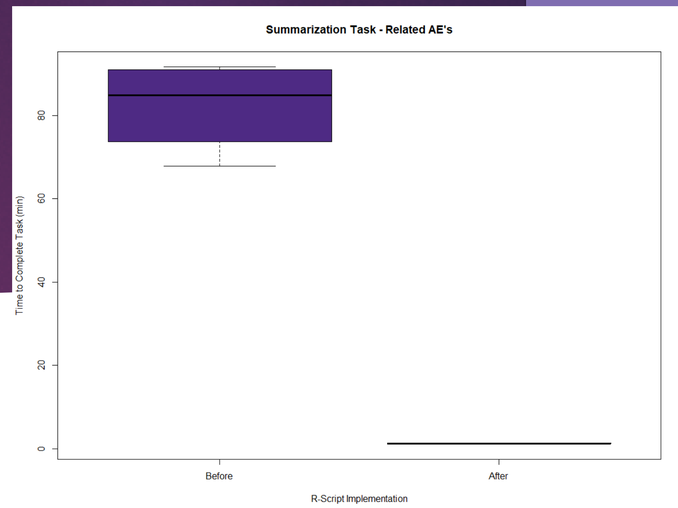

Figure 3. Summarization times for completion of Related AEs Tables

Results

There was a 91.3% reduction in incorrect calculations and typos after R-script implementation. After R-Script implementation there were two errors which were attributed to copying/pasting the table into the format used for our reports.

A paired-samples t-test was conducted on the three tables to compare time to complete before and after R-script implementation. There was a significant difference in the time to complete the table for all AEs before (M=49.7 minutes, SD=12.4) and after R-script implementation (M=1.3, SD=0.2); t(3)=7.8, p = 0.998. There was also a significant difference in the time to complete the table for related AEs before (M=81.1, SD=11.1) and after R-script implementation (M=1.2, SD=0.1); t(3)=14.7, p = 0.999. Finally, there was also a significant difference in the time to complete the table for demographics before (M=11.6, SD=6.3) and after R-script implementation (M=2.6, SD=1.0); t(3)=2.9, p = 0.970. There was a 92.2% reduction in time spent on tables for IND and SARs reports.

Discussion

Overall, R-script implementation has reduced the total annual time spent on all these tables from 153 to 12 hours. Saving 141 hours allows more time for effective trial monitoring, auditing, and other process improvement. Further, accuracy and consistency are of the utmost importance when reporting to the FDA/IRB. User error was noted and reduced in multiple places while implementing this process. In the future, further improvement to the demographics R-Script can eliminate the copy and paste errors uncovered when testing. To achieve this, the output will be updated to a format that matches our report templates. Ultimately, R-scripts can reduce time and error on any task that involves data summarization.

Contact

Quality Assurance Department
Robert H. Lurie Comprehensive Cancer Center
Northwestern University
croqualityassurance@northwestern.edu