



# Use of R Script to Create Trial Summary (ts.xpt) Domains for Nonclinical SEND Studies

Bob Friedman, Xybion; Anthony Fata, PDS; William Varady, Boehringer Ingelheim; William Houser, BMS; Kevin Snyder, FDA

## Abstract

The "Study Data Technical Conformance Guide" for electronic submission of study data requires a Trial Summary (ts.xpt) dataset file even for legacy studies. As legacy studies may not be part of systems capable of producing SEND compliant datasets, the question arose of how to create trial summary files in isolation. The Nonclinical Script Assessment Project (part of the FDA/PhUSE "Nonclinical Topics" working group) was created to explore computer scripts to help with SEND data analysis. As part of this effort, an "R" script was created and shared that generates the ts.xpt to meet this requirement. The script takes as input an Excel (xlsx) file and outputs one or more ts.xpt datasets with flexibility on the parameter values to include. In addition, this project is also collaborating on other scripts to create and analyze SDTM/SEND datasets.

## Breakdown of R Script

Requirements: a Windows computer with Java installed, the R-Stats program (open source), downloadable R libraries (open source), and Microsoft Excel.

Optionally: SAS viewer (free to use), a SEND validator tool, and an integrated development environment for R (for example R Studio, also open source).

The non-clinical topics Github repository (see Figure 4) has the script, library method replacement, and example input and output files.

This script reads an XLSX file with multiple different study's TS information and creates separate ts.xpt files for each of those studies.

This script does not validate or check against any FDA rules. Please use a separate validator before submission.

## Other Nonclinical Contributions

- **Public SEND data packages**
  - Instem
  - Nimble Clinical Research (Includes a Study Data Reviewer's Guide)
  - PDS
- **Several R scripts (Including: "CreatingXPT" that creates TS domains.)**
- **Several SAS scripts for working with SEND datasets including tools to:**
  - Convert Excel files to \*.xpt files: SEND3\_0-from-xls
  - Convert \*.xpt files to Excel files: SEND3.0-to-xls
  - Create \*.pdf files from \*.xpt files: SAS3.0-to-print
  - Convert \*.xpt to simple \*.xml files: SEND3.0-to-xml



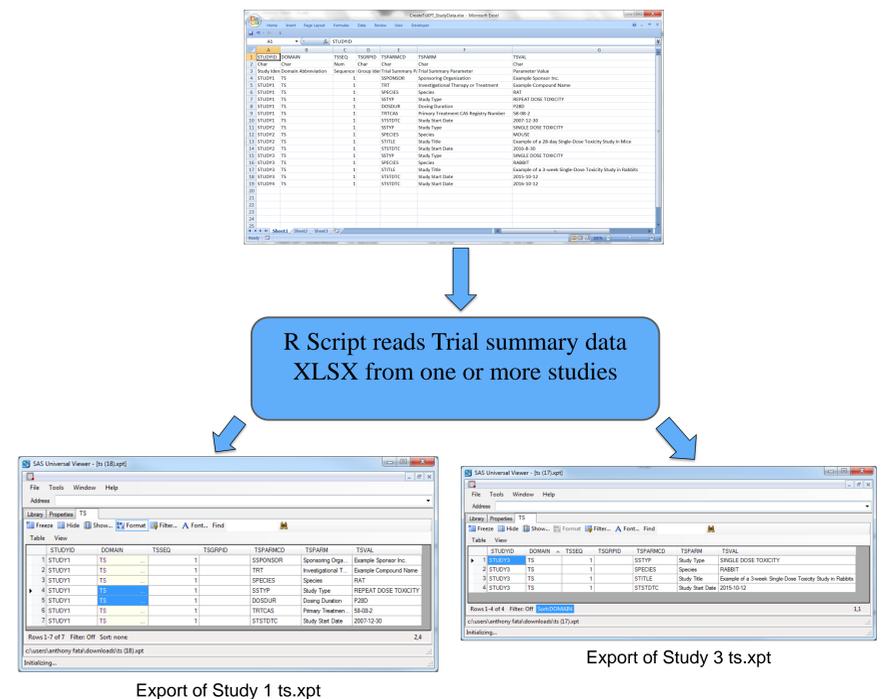
Figure 2: QR Code to access the Nonclinical Script Assessment Project PhUSE Wiki page



Figure 3: QR Code to access the Nonclinical Script Assessment Project Github Repository

Note: The opinions expressed in this poster are those of the authors and do not necessarily represent the opinions of their respective companies

## Flow Chart of Creation Process



```

33 # Select file to read
34 mainDir <- "c:/temp/r testing/xpt output"
35 # Set output files
36 setwd(mainDir)
37 myFile <- file.choose()
38 # Read in XLSX file
39 df <- readxl::read_excel(myFile,
40                           sheet=1,
41                           startRow = 1,
42                           endCol = 7)
43 # make copy skipping first two rows (field type and field label)
44 df2 = df[-1,-1]
45 df2 = df2[-1,-1]
46 # for those that are num, transform to numeric
47 df2$transform(df2, TSSSQ = as.numeric(TSSSQ))
48 # set labels for each field
49 Hmisc::label(df2)=df2[,1]
50 # For each set of rows belonging to a study create TS.XPT file
51 studyList <- unique(df2$STUDYID)
52 for(aStudy in studyList){
53   # Create subdirectory
54   # Set output files
55   setwd(mainDir)
56   if (file.exists(aStudy)){
57     setwd(file.path(mainDir, aStudy))
58   } else {
59     dir.create(file.path(mainDir, aStudy))
60     setwd(file.path(mainDir, aStudy))
61   }
62   # filter to this study
63   studyData <- subset(df2, STUDYID==aStudy, keepNA=FALSE)
64   # Set length for character fields
65   SASformat(studyData$DOMAIN) <- "S2."
66   # place this dataset into a list with a name
67   aList = list(studyData)
68   # name it
69   names(aList)[1] <- "TS"
70   # and label it
71   attr(aList, "label") <- "TRIAL SUMMARY"
72   # write out dataframe
73   write_xport2(
74     list=aList,
75     file = "ts.xpt",
76     verbose=FALSE,
77     sasVer="7.00",
78     osType="R 3.0.1",
79   ) # end of study loop
80 setwd(mainDir)

```

Figure 1: Screenshot of a portion of the R script

## Lessons Learned

The SEND Implementation PhUSE Wiki FAQ covers most questions regarding the criteria set forth by the FDA ([http://www.phusewiki.org/wiki/index.php?title=SEND\\_Implementation\\_Wiki\\_-\\_FAQ](http://www.phusewiki.org/wiki/index.php?title=SEND_Implementation_Wiki_-_FAQ))

All dates must follow the ISO 8601 format (ex: YYYY-MM-DD), otherwise the submission into the Gateway will trigger a technical rejection.

FDA is investigating expanding the criteria for legacy studies beyond just STSDTC to include: SSPONSOR, TRT, SPECIES, SSTYP, DOSDUR, TRTCAS

In order to create ts.xpt file with correct dataset name (TS), dataset label (Trial Summary) and correct lengths for each column, a replacement of a portion of the SASxport R library was created.

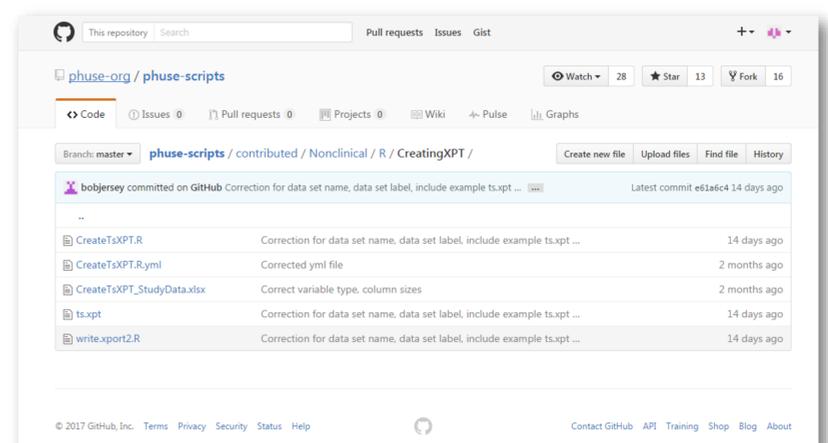


Figure 4: Screenshot of Script location within the PhUSE Github Repository. Can be found at the following link: <https://github.com/phuse-org/phuse-scripts/tree/master/contributed/Nonclinical/R/CreatingXPT>

Acknowledgements to the Group: Kevin Cahill, EPL; Bob Dorsam, FDA; Chris Eley, Pfizer; Elaine Thompson, FDA; Stephen MacMannis, Pfizer; Fred Mura, PDS