R PMX Best Practices Checklist – v2

Andy Stein with R Subject Matter Expert Team

Reproducibility

- Number R scripts to reflect the order of the workflow (e.g. `Task01_ExploreData.R`, `Task02_MonolixRun.R`, `Task03_MonolixPlot.R`). Data preparation and modeling should be in separate scripts.
- Final dataset for modeling should be saved as a csv file.
- Use an umbrella script that calls all other scripts. Options include:
  - R script. Run with `R CMD BATCH "Task00_RunAll.R"` or
  - Umbrella Unix Script.
- Set `echo=TRUE` when calling scripts to produce a comprehensive output file. `source("Task01.R", echo=TRUE)`
- Set the random number generator seed `set.seed(123456)`
- Don't use `setwd()` in any Rscript. Write your code such that it executes when the working directory is the directory where the script is stored.
- At the top of all scripts, consider sourcing an `initialize_script.R file`. This file should:
  - Declare constants upfront (never hardcode numbers). Be sure to include units and a reference, e.g. `Kd = 1.5 #nM, from Investigator's Brochure version 1, page 27.`
  - Specify in variables the paths where data is located and where results are stored. Use these variables in subsequent scripts when loading data and saving results.
- Specifying global paths at initialization for traceability. Elsewhere, either use local paths: (`../monolix/run001.mlxtran`) or the parent a directory variable: `file.path(gps_directory, "monolix", "run001.mlxtran")`.
- At the end of all scripts, call `sessionInfo()` to document R version and all packages used.

Readability

- Ensure sufficient comments, especially units and references
- Create `Readme.txt` to summarize how to understand and execute your code.
- Follow a style guide. We like the tidyverse style guide, but any consistent naming convention is ok.

Graphs and Tables

- Plots should be labeled "draft" unless final. Use `xgxr::xgx_annotate_status()`
- Plots should list program that created them and graphics file output location. `xgxr::xgx_annotate_filenames()`
- Start of plot filename should match the start of the script name that created it (e.g. `Task02_ExploreData.R` would create `Task02_Spaghetti.pdf`).
- Use `ggplot2, xgxr, tidyr, dplyr,` so others can easily read/validate code.
- Use a white background with faint grid lines. These settings can be default with: `xgxr::xgx_theme_set()`
- When saving a figure or displaying it in Rmarkdown, set the width and height so that the font is easy to read.
- Use `ggPMX` for standard Monolix outputs.