Enhancing Clinical Trials FDA Submission Documentation through the Power of Metadata

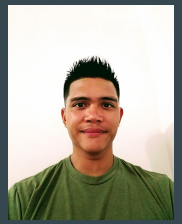
A Proposal for New Standards in Open-Source Submissions

FDA Quarterly Meeting : 16OCT2024

Presented by the PHUSE Open-Source Metadata Documentation Working Group (OSDocuMetaWG)

PHUSE Open-Source Metadata Documentation Working Group (OSDocuMetaWG)

Lovemore Gakava (Co-Lead Novo Nordisk)



Joel Laxamana (Co-Lead - Roche/ Genentech)

Steven Haesendonckx (Johnson and Johnson)

Nicholas Masel (Johnson & Johnson)

Room for others to contribute... :)

Problem Statement & Objective

- The objective of this project is to develop a new template or enhance an existing one, such as the Study Data Standardization Plan (SDSP) or Analysis Data Reviewer's Guide (ADRG).

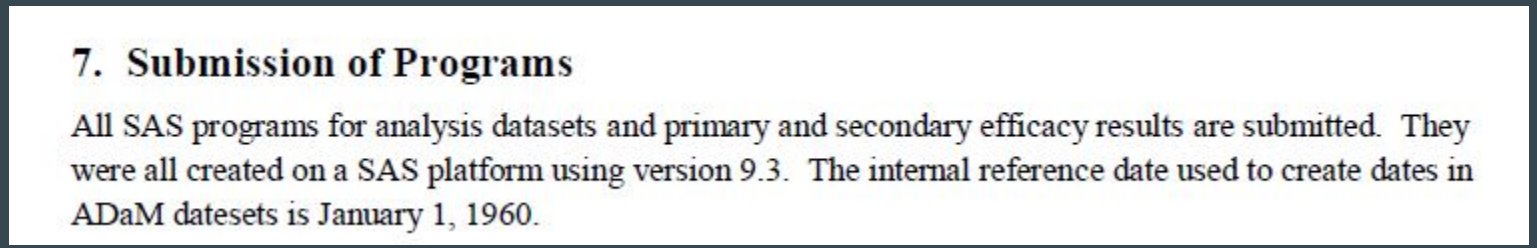
- The aim is to ensure consistent documentation of metadata pertaining to the versions of statistical packages and procedures, aligning with health authority expectations.

- The standardized template will streamline the submission of clinical study metadata to health authorities as part of the regulatory review process.

Documentation of versions of statistical packages and procedures

Proprietary Statistical Programming Languages

1. The version metadata of the statistical procedures/functions has been documented within the Statistical Analysis Plan with a general statement of software and version used (other companies may use other documentation)
2. Detailed information has been added to the Analysis Data Reviewer's Guide (ADRG). The ADRG has a section where the programming software version and functions are specified



7. Submission of Programs

All SAS programs for analysis datasets and primary and secondary efficacy results are submitted. They were all created on a SAS platform using version 9.3. The internal reference date used to create dates in ADaM datesets is January 1, 1960.

Documentation of versions of statistical packages and procedures

Open-Source software

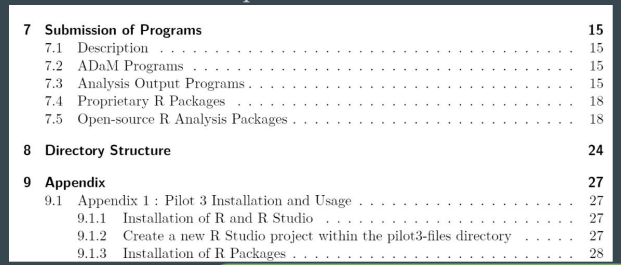
Challenges

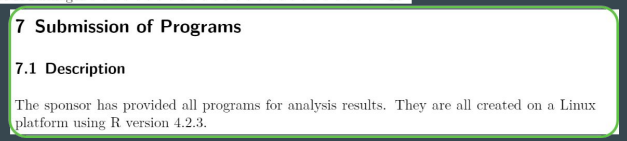
For the regulatory authority to reproduce the analysis, additional information is needed such as:

- A summary of analysis packages, their versions and download locations
- The software version and download location
- Additional instructions for reproducibility of the computing environment.

Current R-based submissions

- ADRG Example 1: R Submission Pilot 3



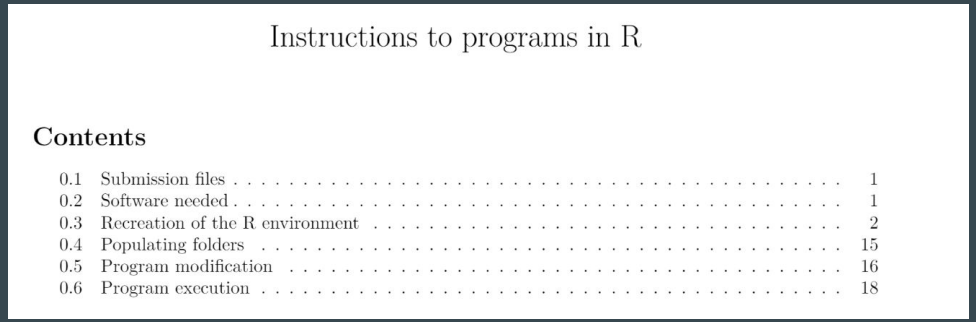


7 Submission of Programs

7.1 Description

The sponsor has provided all programs for analysis results. They are all created on a Linux platform using R version 4.2.3.

ADRG Example 2: Novo Nordisk R submission



Instructions to programs in R

Contents

Instructions for Reproducing the Analysis Environment for HA review : Pilot 3 ADRG

- Currently, in the Pilot 3 ADRG, we have a section, which covers detailed step-by-step instructions on how to set-up the analysis environment and re-execute the analysis scripts.
- Our proposal is to allow the addition of a section like this in the ADRG.
- An example demo can be found here :

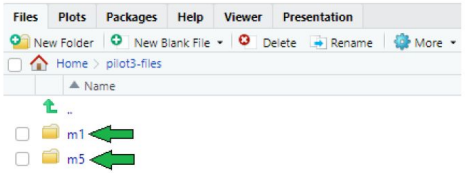
[R Consortium's R-based Test Submission Package for FDA Evaluation - Joel Laxamana]

9 Appendix

9.1 Appendix 1 : Pilot 3 Installation and Usage

To install and execute the R programs, follow all of the procedures below. Ensure that you note the location of where you downloaded the Pilot 3 eCTD submission files. For demonstration purposes, the procedures below assume the transfer has been saved to this location: C:\pilot3.

In addition, create a new directory to hold the unpacked Pilot 3 ADaM and tlf programs and files. For demonstration purposes, the procedures below assume the new directory is this location: C:\pilot3-files, where the unpacked files are shown as the m1 and m5 directories.



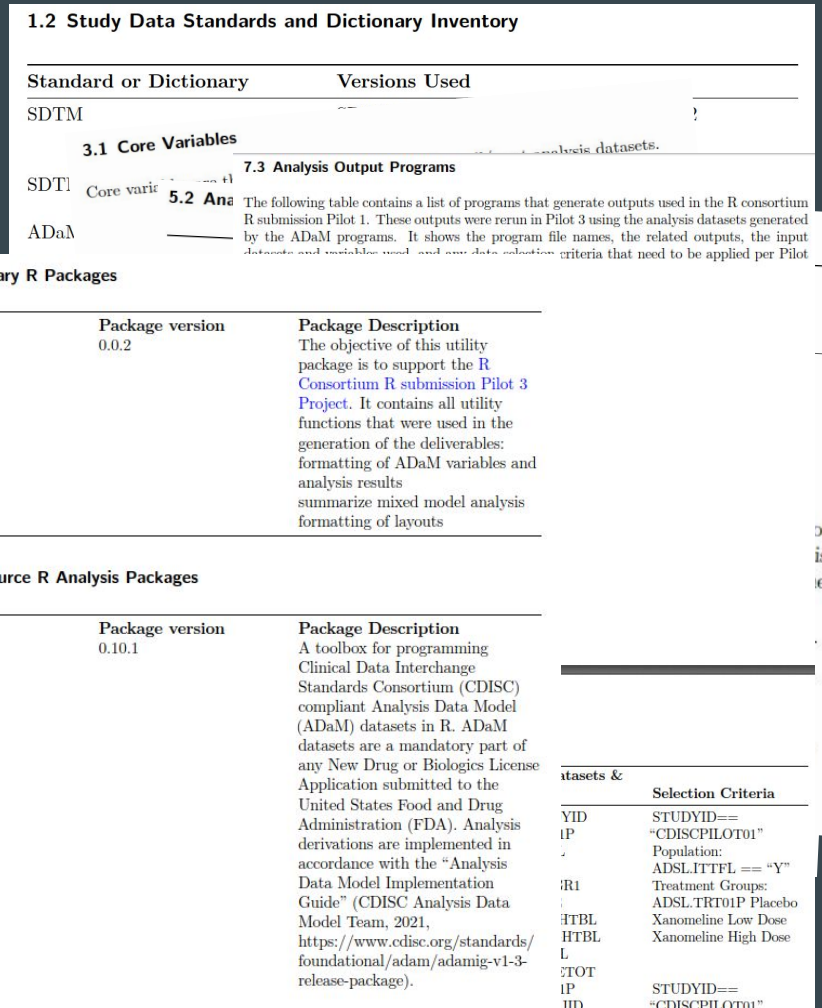
9.1.1 Installation of R and R Studio

Download and install R 4.2.3 for Windows from https://cran.r-project.org/bin/windows/ba

Other Metadata use Opportunities

Tables in the Current ADRG for Automation - from the Pilot 3 ADRG

- 1.2 Study Data Standards and Dictionary Inventory
  - Potential source: TS SDTM metadata
- 3.1 Core Variables
  - Source: define.xml specifications
- 4.2 Data Dependencies
  - Code scraping for ADaM programs.
- 5.2 Analysis Datasets
  - Metadata extraction from .xpts
- 6.2 Issues Summary
  - Report extraction from Pinnacle 21
- 7.2 ADaM Programs
  - Metadata extraction from .xpts
- 7.3 Analysis Output Programs
  - Read in folder that contains selected outputs
- 7.4 & 7.5 R Packages
  - Code utilizing SessionInfo() and {renv}



1.2 Study Data Standards and Dictionary Inventory

Standard or Dictionary	Versions Used
SDTM	

3.1 Core Variables

SDTM Core variables

5.2 Ana

ADaM

7.3 Analysis Output Programs

The following table contains a list of programs that generate outputs used in the R consortium R submission Pilot 1. These outputs were rerun in Pilot 3 using the analysis datasets generated by the ADaM programs. It shows the program file names, the related outputs, the input datasets and variables used, and any data selection criteria that need to be applied per Pilot

7.4 Proprietary R Packages

R Package	Package version	Package Description
pilot3utils	0.0.2	The objective of this utility package is to support the R Consortium R submission Pilot 3 Project. It contains all utility functions that were used in the generation of the deliverables: formatting of ADaM variables and analysis results summarize mixed model analysis formatting of layouts

7.5 Open-source R Analysis Packages

R Package	Package version	Package Description
admiral	0.10.1	A toolbox for programming Clinical Data Interchange Standards Consortium (CDISC) compliant Analysis Data Model (ADaM) datasets in R. ADaM datasets are a mandatory part of any New Drug or Biologics License Application submitted to the United States Food and Drug Administration (FDA). Analysis derivations are implemented in accordance with the "Analysis Data Model Implementation Guide" (CDISC Analysis Data Model Team, 2021, https://www.cdisc.org/standards/foundational/adam/adamig-v1-3-release-package).

atasets &

	Selection Criteria
YID	STUDYID==
1P	"CDISCPILOT01"
	Population:
	ADSL.ITTFL == "Y"
R1	Treatment Groups:
	ADSL.TRT01P Placebo
HTBL	Xanomeline Low Dose
HTBL	Xanomeline High Dose
L	
TOT	
1P	STUDYID==
	"CDISCPILOT01"

Other eCTD Structure and File Format Opportunities

Enhancing FDA Submissions with HTML and QMD Files
Introduction

- Current State: Submissions include ADRG as .pdf files
- Proposal: Include .html and .qmd files in submissions

Benefits of HTML and QMD Files
1. Reproducibility

- Dynamic content and up-to-date rendering
- Easier to replicate and verify analyses

2. Enhanced Documentation

- Rich features like hyperlinks, tables, and multimedia
- Improved user experience and readability

3. Versatility

- HTML Files: Interactive, easily navigable, and portable
- QMD Files: Can be rendered to both HTML and PDF formats for flexibility

4. Immediate Advantages

- Modern, dynamic presentations of data and metadata
- Facilitate better understanding and review processes
- Ability to quickly convert back to traditional PDF if necessary

5. Conclusion

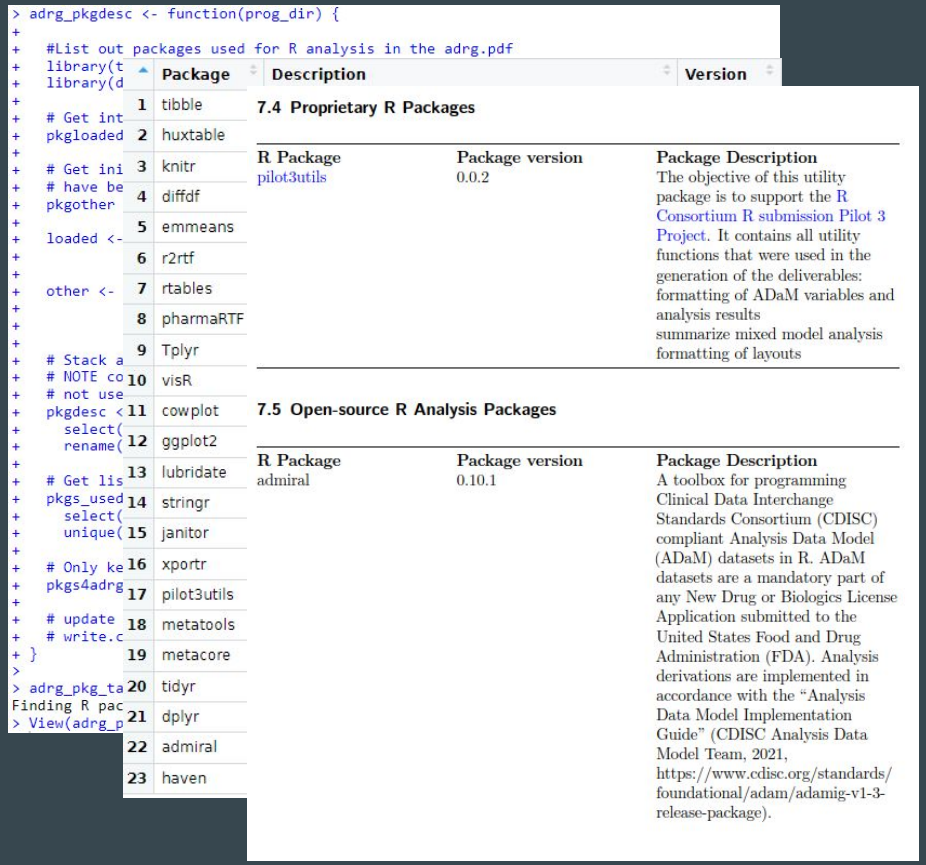
- Proposal for FDA and Health Agencies to accept .html and .qmd files
- Enhances data reproducibility, documentation, and overall submission quality

Open-Source Package Development

Possibility of an Open-Source Package

- **For example** : A R function that can utilize metadata from the project R Environment to help semi-automate a list of R packages to include in the ADRG.
  - *Code utilizing SessionInfo() and {renv}*
- **Benefits for stakeholders**
  - Provides necessary information needed for reproducibility
- **Standardization**
  - Allows future Open-Source submissions to follow the same approach in a consistent and efficient manner
- **Looking ahead**
  - From the list of “Other Metadata use Opportunities”, more functions like this can be developed and stored in an Open-source package to help collect further metadata to fill in the ADRG template. This can modernize the way we develop the ADRG today.

On the right you can see the function which collects the list of R packages used in a Clinical Trial Analysis, the actual list of R packages, Description and Version, then how this information is brought into the ADRG.



```
> adrg_pkgdesc <- function(prog_dir) {
+
+   #List out packages used for R analysis in the adrg.pdf
+   library(t
+   library(d
+
+   # Get int
+   pkgloaded
+
+   # Get ini
+   # have be
+   pkgother
+
+   loaded <-
+
+   other <-
+
+
+   # Stack a
+   # NOTE co
+   # not use
+   pkgdesc <
+     select(
+     rename(
+
+   # Get lis
+   pkgs_used
+     select(
+     unique(
+
+   # Only ke
+   pkgs4adrg
+
+   # update
+   # write.c
+ }
>
> adrg_pkg_ta
Finding R pac
> View(adrg_p
```

	Package	Description	Version
1	tibble		
2	huxtable		
3	knitr		
4	diffdf		
5	emmeans		
6	r2rtf		
7	rtables		
8	pharmaRTF		
9	Tplyr		
10	visR		
11	cowplot		
12	ggplot2		
13	lubridate		
14	stringr		
15	janitor		
16	xportr		
17	pilot3utils		
18	metatools		
19	metacore		
20	tidyr		
21	dplyr		
22	admiral		
23	haven		

7.4 Proprietary R Packages

R Package	Package version	Package Description
pilot3utils	0.0.2	The objective of this utility package is to support the R Consortium R submission Pilot 3 Project. It contains all utility functions that were used in the generation of the deliverables: formatting of ADaM variables and analysis results summarize mixed model analysis formatting of layouts

7.5 Open-source R Analysis Packages

R Package	Package version	Package Description
admiral	0.10.1	A toolbox for programming Clinical Data Interchange Standards Consortium (CDISC) compliant Analysis Data Model (ADaM) datasets in R. ADaM datasets are a mandatory part of any New Drug or Biologics License Application submitted to the United States Food and Drug Administration (FDA). Analysis derivations are implemented in accordance with the “Analysis Data Model Implementation Guide” (CDISC Analysis Data Model Team, 2021, https://www.cdisc.org/standards/foundational/adam/adamig-v1-3-release-package).

Possible solutions

- Open Source submissions
  - Add the information to the SAP and ADRG
  - Add the information to the ADRG - similar to the example submissions
  - Create a supplementary document to ADRG with the open source information
  - Other?
- Hybrid Submissions
  - Add the information to the SAP and ADRG?

Next Steps and Feedback

- **Review and Feedback :**
  - Current notes and brainstorming : OSDocuMeta Minutes
- **Call for Input :**
  - Encourage thoughts and suggestions offline (add a question here).

The End : THANK YOU FOR YOUR TIME!

Q&A