# DIA Biostatistics Industry and Regulator Forum

April 6-8 | Virtual

#BioStats22

# Achieving Regulatory Approval Using R

**Tae Hyun (Ryan) Jung, Ph.D.**

**FDA/CDER/OB**

DIA

# Disclaimer

This presentation reflects the views of the presenter and should not be construed to represent Food and Drug Administration views or policies

# Statistical Software Clarifying Statement

FDA does not require use of any specific software for statistical analyses, and statistical software is not explicitly discussed in Title 21 of the Code of Federal Regulations [e.g., in 21CFR part 11]. However, the software package(s) used for statistical analyses should be fully documented in the submission, including version and build identification.

# Outline

- **Overview of Prograf RWE sNDA**
  - Clinical/Regulatory Background
  - Summary of Study
  - Scientific Registry of Transplant Recipients
  - Standard Analytic Files
- **Pre-/Post-Submission Discussion**
- **Using R-Markdown for Regulatory Submission**
  - Raw Data Format and Analysis Data Creation
  - Examples of R-Markdown Practices and Output
  - Lessens Learned from Regulatory Perspective

# Overview of Prograf RWE Application

## Clinical/Regulatory Background

- Prograf® (Tacrolimus) was indicated for the prophylaxis of organ rejection in adult and pediatric patients receiving allogeneic lung transplant in combination with other immunosuppressants
- Prograf was approved for prophylaxis of organ rejection in patients receiving liver transplants in 1994 (later for kidney & heart), based on RCT Evidence
- RCTs for lung was not submitted to FDA, but drug has been used widely in clinical care; sponsor (Astellas) submitted supplemental New Drug Application to FDA on Dec 15, 2020
- Approval for preventing rejection/death in lung transplant granted Jul 16, 2021

# Overview of Prograf RWE Application

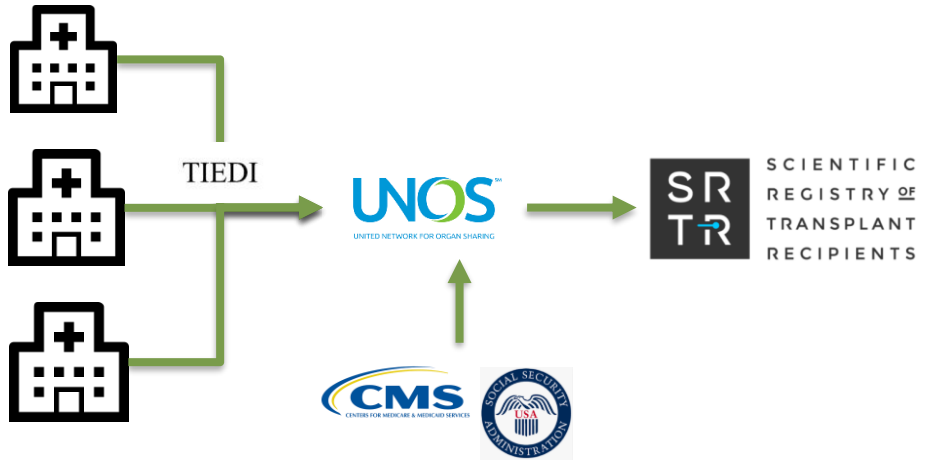
## Summary of Study

- **Study Design**: Non-interventional (observational) treatment arm, compared to historical controls
- **Primary Endpoint**: A composite endpoint of graft failure (GF) or death (due to any cause) within one year (365 days) after transplant
- **Data Source**: Scientific Registry of Transplant Recipients (SRTR) data on all lung transplants in US during 1999–2017
- **Study Population**: Adult and pediatric patients in tacrolimus immediate release (TAC IR) in combination with mycophenolate mofetil (MMF) or azathioprine (AZA)


*https://www.fda.gov/drugs/news-events-human-drugs/fda-approves-new-use-transplant-drug-based-real-world-evidence*

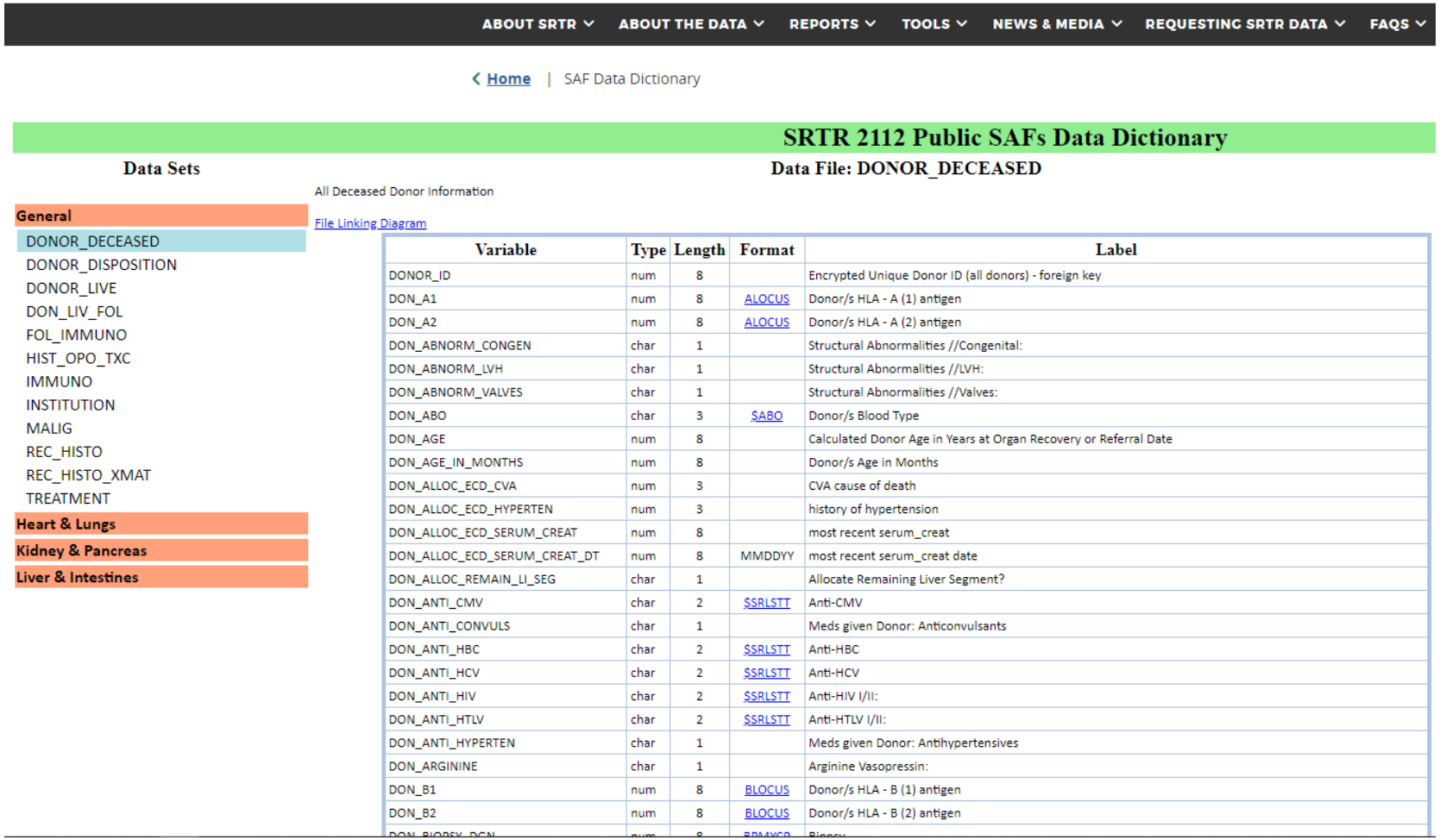
# Overview of Prograf RWE Application

- **Scientific Registry of Transplant Recipients (SRTR)**
  - The SRTR is a national transplant registry and made available under a data use agreement to external researchers and includes outcomes for all transplant recipients, candidates, recipients and donors in the United States from Oct 1987 onward
  - For Prograf RWE application, the sponsor submitted the SRTR standard analytic files (SAFs) that contain data on all lung transplant candidates, recipients, and donors in the United States between 1999 and 2017



# Overview of Prograf RWE Application

## Public Data Dictionary of SRTR Standard Analytic Files



ABOUT SRTR | ABOUT THE DATA | REPORTS | TOOLS | NEWS & MEDIA | REQUESTING SRTR DATA | FAQS

Home | SAF Data Dictionary

**SRTR 2112 Public SAFs Data Dictionary**

**Data Sets**

- General
  - DONOR_DECEASED
  - DONOR_DISPOSITION
  - DONOR_LIVE
  - DON_LIV_FOL
  - FOL_IMMUNO
  - HIST_OPO_TXC
  - IMMUNO
  - INSTITUTION
  - MALIG
  - REC_HISTO
  - REC_HISTO_XMAT
  - TREATMENT
- Heart & Lungs
- Kidney & Pancreas
- Liver & Intestines

**Data File: DONOR_DECEASED**

All Deceased Donor Information

File Linking Diagram

| Variable | Type | Length | Format | Label |
|---|---|---|---|---|
| DONOR_ID | num | 8 | | Encrypted Unique Donor ID (all donors) - foreign key |
| DON_A1 | num | 8 | ALOCUS | Donor/s HLA - A (1) antigen |
| DON_A2 | num | 8 | ALOCUS | Donor/s HLA - A (2) antigen |
| DON_ABNORM_CONGEN | char | 1 | | Structural Abnormalities //Congenital: |
| DON_ABNORM_LVH | char | 1 | | Structural Abnormalities //LVH: |
| DON_ABNORM_VALVES | char | 1 | | Structural Abnormalities //Valves: |
| DON_ABO | char | 3 | $ABO | Donor/s Blood Type |
| DON_AGE | num | 8 | | Calculated Donor Age in Years at Organ Recovery or Referral Date |
| DON_AGE_IN_MONTHS | num | 8 | | Donor/s Age in Months |
| DON_ALLOC_ECD_CVA | num | 3 | | CVA cause of death |
| DON_ALLOC_ECD_HYPERTEN | num | 3 | | history of hypertension |
| DON_ALLOC_ECD_SERUM_CREAT | num | 8 | | most recent serum_creat |
| DON_ALLOC_ECD_SERUM_CREAT_DT | num | 8 | MMDDYY | most recent serum_creat date |
| DON_ALLOC_REMAIN_LI_SEG | char | 1 | | Allocate Remaining Liver Segment? |
| DON_ANTI_CMV | char | 2 | $SRLSTT | Anti-CMV |
| DON_ANTI_CONVULS | char | 1 | | Meds given Donor: Anticonvulsants |
| DON_ANTI_HBC | char | 2 | $SRLSTT | Anti-HBC |
| DON_ANTI_HCV | char | 2 | $SRLSTT | Anti-HCV |
| DON_ANTI_HIV | char | 2 | $SRLSTT | Anti-HIV I/II: |
| DON_ANTI_HTLV | char | 2 | $SRLSTT | Anti-HTLV I/II: |
| DON_ANTI_HYPERTEN | char | 1 | | Meds given Donor: Antihypertensives |
| DON_ARGININE | char | 1 | | Arginine Vasopressin: |
| DON_B1 | num | 8 | BLOCUS | Donor/s HLA - B (1) antigen |
| DON_B2 | num | 8 | BLOCUS | Donor/s HLA - B (2) antigen |

**20680 subjects with single observation, 453 variables included in one final dataset**



https://www.srtr.org/requesting-srtr-data/saf-data-dictionary

# Pre-submission Discussions

| | Typical NDA Submission | Prograf RWE Submission |
|---|---|---|
| Study Report | Y | Y |
| Define Document | Y | Y |
| FDA ESG | Y | Y |
| Standardized Data | Y | N |
| SAS XPORT | V5 | V8 |
| Programming Language | SAS | R |

*ESG: Electronic Submission Gateway*

# Pre-submission Discussions

- Type C meeting on Feb 2020: Asked whether hybrid of SAS and R (employing packages such as “survival”, the “tidyverse” suite and “rmarkdown”) is acceptable
- Type B meeting on Aug 2020: Asked whether the programs submitted in .Rmd and html formats will be acceptable for review
  - Agreed with the program submission using R markdown. Asked sponsor to generate the final format into both html and pdf formats. Also, specify alternative packages if applicable.
  - Asked sponsor to submit a sample data (200 subjects) with relevant programs and document.
- Sample data/code submission and testing on Nov 2020: FDA assessed sample data with relevant R-markdown before sNDA submission on Dec 2020

# Post-submission Discussions

- Information Request on Jan 2021: Experienced difficulties installing company-owned R package; Not compatible with FDA (R 3.6.1 and R 4.0.3) and does not disclose a GitHub allowing for manual download
    - Submit the GitHub address
    - Clarify the required R version
    - Provide a reference manual (static pdf vignette)
- Information Request on Apr 2021: Experienced difficulties using sponsor suggested LaTeX editor
    - Submitted Rmd programs in a LaTeX-free environment

# Using R-Markdown for Regulatory Submission

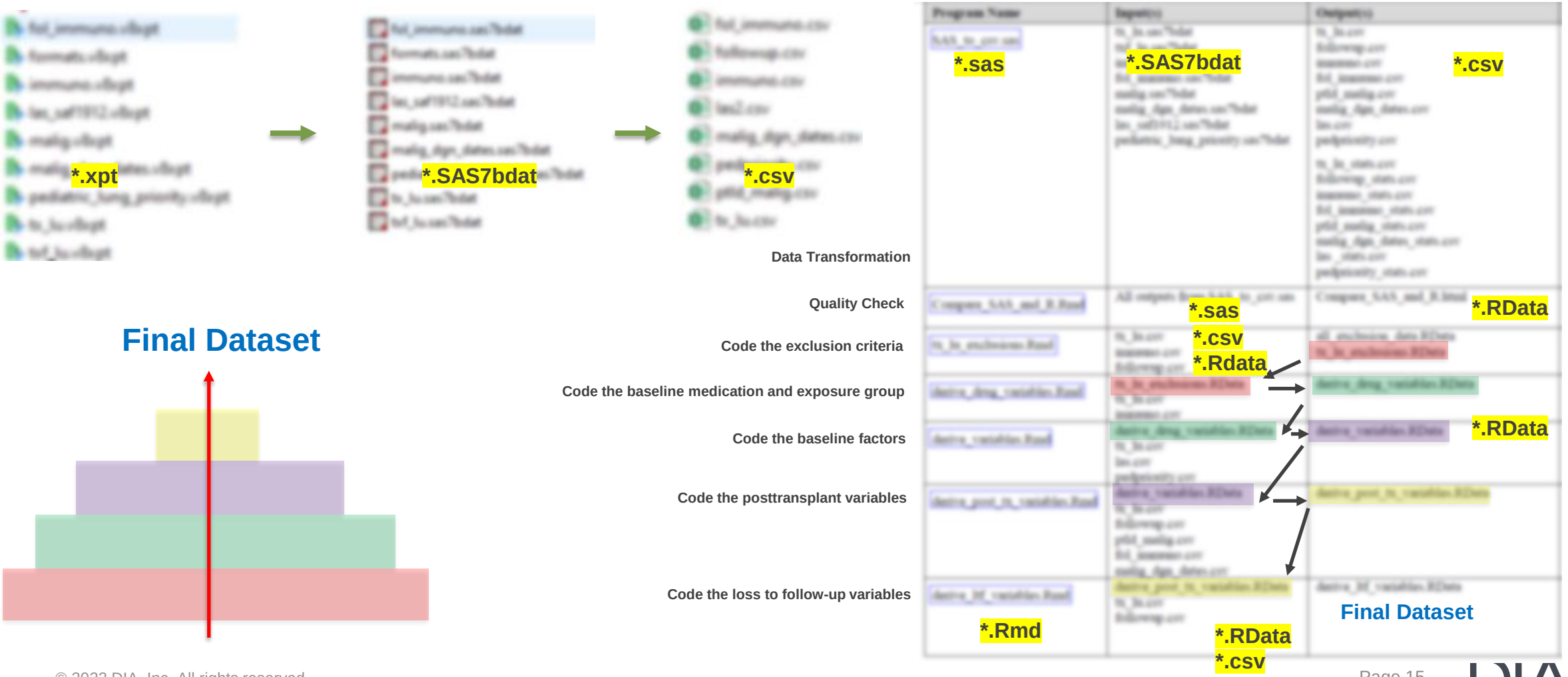
- **Raw Data Format of Standard Analytic Files (SAFs)**
  - SRTR provides data only in SAS format (“.xpt/sas7bdat” extension)
  - SAS 9.4 (SAS/STAT 15.1) was used to read SAF .xpt/sas7bdat files
  - Sponsor provided programming code and analysis data in R-Markdown (.Rmd)
  - Formatted SAS variables were converted into corresponding CSV variables that were composed of both the unformatted and formatted values
  - Imported the CSV files into R to create dataset of interest
  - R-Markdown provided both the R code as well as comments and intermittent results that create transparent documentation of the process

# Using R-Markdown for Regulatory Submission

- **Analysis Dataset Description**
  - Analysis datasets support all protocol- and statistical analysis plan-specified objectives
    - tx_data (Lung Transplant Analysis Data): A dataset with one record per patient containing all variables used in the tables and figures
    - Efficacy, safety, baseline characteristics
    - No PK/PD
  - No ADaM conformance checks were performed
  - Although the define file has the “look and feel” of define.xml v2.0, the underlying xml is not compliant with the standard

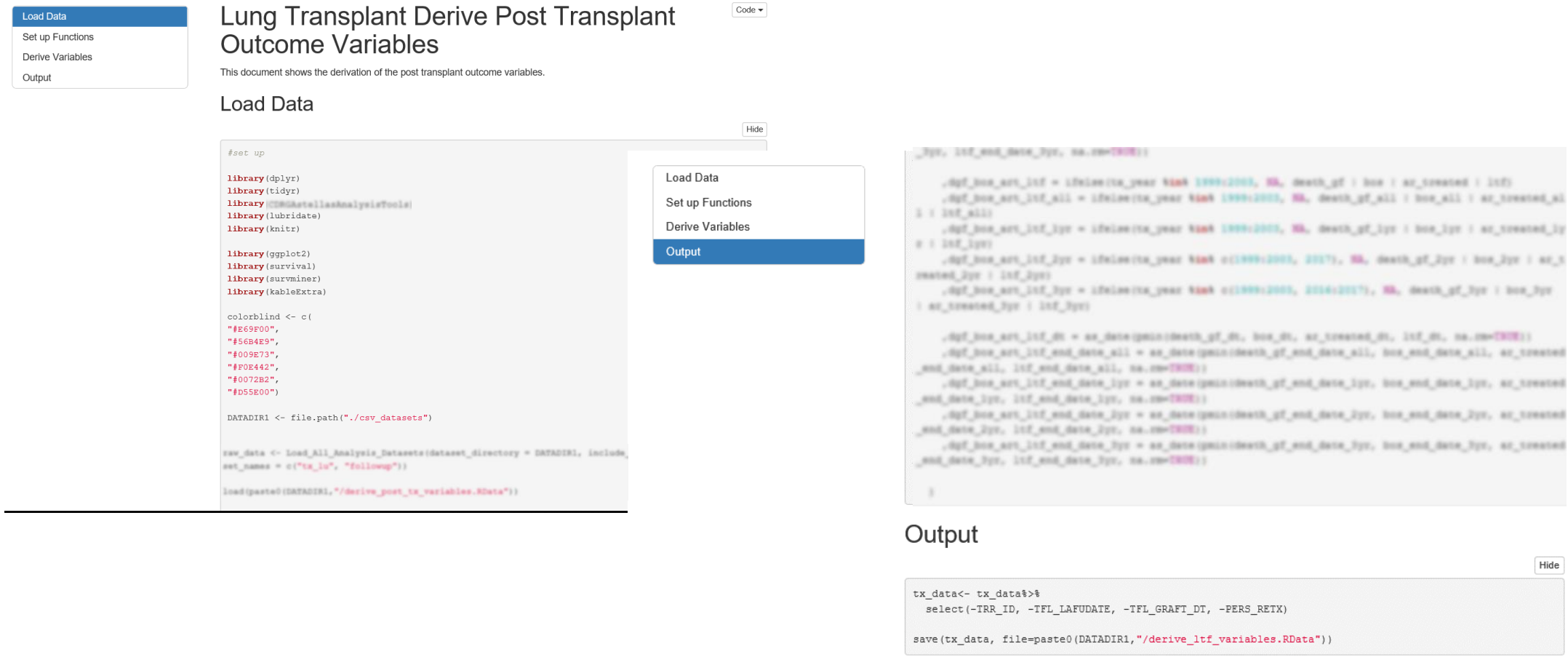
# Using R-Markdown for Regulatory Submission

## Steps for Data Management



# Using R-Markdown for Regulatory Submission

## Snapshots of final data management stage



# Using R-Markdown for Regulatory Submission

## HTML Output of R-Markdown



- Derive Drug Variables: See: derive_drug_variables.html
- Derive Other Variables: See: derive_variables.html

### Study Population

There were a total of 29772 lung transplants between 1999 and 2017; after exclusions 26080 lung transplant recipients remained in the study cohort.

### Non-Cumulative Exclusions

Study Population Exclusions by Age: Non-cumulative

| Grouping | Characteristic | Total | Adult | Pediatric |
|---|---|---|---|---|
| All | | 29772 | 28817 | 955 |
| Received any previous transplant, lung or otherwise | | 1240 (4.2%) | 1174 (4.1%) | 66 (6.9%) |
| Multi-organ transplant | | 134 (0.5%) | 120 (0.4%) | 14 (1.5%) |
| Living donor transplant | | 245 (0.8%) | 152 (0.5%) | 93 (9.7%) |
| Missing discharge date | | 248 (0.8%) | 236 (0.8%) | 12 (1.3%) |
| Discharged after 1 year | | 34 (0.1%) | 34 (0.1%) | 0 |
| Missing maintenance immunosuppression information at discharge | | 975 (3.3%) | 941 (3.3%) | 34 (3.6%) |
| Death, graft failure, or retransplant before discharge | | 1837 (6.2%) | 1767 (6.1%) | 70 (7.3%) |
| Incorrect Death Date | | 1 (0.0%) | 1 (0.0%) | 0 |
| Incorrect Retransplant Date | | 1 (0.0%) | 1 (0.0%) | 0 |
| Any Exclusion | | 3692 (12.4%) | 3462 (12.0%) | 230 (24.1%) |
| Included | | 26080 (87.6%) | 25355 (88.0%) | 725 (75.9%) |

*Note:*
A patient may be excluded for more than one reason. Non-cumulative data report counts and percent among all patients. The sum of all non-cumulative exclusions exceeds the total number of excluded patients.
* Program: './Create_cohort_files/NoLatex_Demographics_Report.Rmd'



### One Year Death or Graft Failure By Subgroup

#### Adult: Cumulative Incidence

Kaplan-Meier Estimate of Cumulative Incidence Percent (95% CI) of Death or Graft Failure at 1y Post-transplant by Subgroup: Adults

| Grouping | | At Risk | Events | % (95% CI) |
|---|---|---|---|---|
| Age at Transplant | 18-34 years | 2804 | 225 | 18.3% (0.0, 35.2) |
| Age at Transplant | 35-49 years | 3786 | 272 | 16.5% (1.5, 29.3) |
| Age at Transplant | 50-64 years | 13257 | 1214 | 11.7% (8.0, 15.3) |
| Age at Transplant | 65+ years | 5508 | 633 | 11.8% (11.0, 12.7) |
| Recipient Sex | Female | 10829 | 920 | 12.7% (7.2, 17.8) |
| Recipient Sex | Male | 14526 | 1424 | 13.2% (8.7, 17.4) |
| Lung Transplant Procedure | Double | 16190 | 1297 | 13.9% (8.2, 19.2) |
| Lung Transplant Procedure | Single | 9165 | 1047 | 12.0% (11.1, 12.8) |
| Transplant Era | 1999-2005 | 6284 | 725 | 16.8% (7.9, 24.8) |
| Transplant Era | 2006-2009 | 5024 | 519 | 11.7% (9.3, 14.0) |
| Transplant Era | 2010-2017 | 14047 | 1100 | 12.3% (6.3, 17.9) |
| Diagnosis Group | A - COPD | 9529 | 856 | 17.0% (6.8, 26.1) |
| Diagnosis Group | B - Pulmonary hypertension | 966 | 80 | 23.6% (0.0, 46.6) |
| Diagnosis Group | C - Cystic fibrosis | 3198 | 247 | 15.5% (0.0, 28.8) |
| Diagnosis Group | D - Pulmonary fibrosis | 11621 | 1154 | 10.3% (9.7, 10.8) |
| Diagnosis Group | Unknown | 41 | 7 | 17.1% (4.7, 27.8) |
| BMI | <18.5 | 2323 | 243 | 10.6% (9.4, 11.9) |
| BMI | 18.5-<25 | 10228 | 844 | 13.5% (6.5, 19.9) |
| BMI | 25-<30 | 8889 | 857 | 13.8% (7.9, 19.4) |
| BMI | 30+ | 3646 | 360 | 10.5% (9.4, 11.6) |
| BMI | Unknown | 259 | 40 | 15.3% (10.8, 19.5) |

# Using R-Markdown for Regulatory Submission

- **Lesson Learned from Regulatory Perspective**
  - RWD sources are disparate and non-uniform
  - For this submission, RWD structure was complex and had no prior knowledge or experience
  - R-Markdown provided transparent data management and analysis processes for review
  - Easy for reproducing and navigating results and supported interdisciplinary collaboration
  - Challenge existed when using R-Markdown throughout the review
    - Provide all necessary documents to aid review process
    - Submit all relevant analytic packages
    - Complex features are unnecessary and may not be compatible with FDA environment
    - Make it analyzable/executable: simple coding practices are acceptable

DIA

# Appendix: Submission Issues

- Data Submission with SAS XPORT (.xpt) version 8

| | Version 5† | Version 8 |
|---|---|---|
| Variable Name | 8 chars | 32 chars |
| Variable Label | 40 chars | 256 chars |
| Variable Length | 200 chars | 32,767 chars |

†FDA Study Data Technical Conformance Guide (version 4.5.1, July 2020)

- Data Submission through File Transfer Protocol (FTP)
- Sending a link to sponsor's database instead of using Electronic Submission Gateway Code Submission using R Markdown
- No SAS codes for data management and analysis

# Appendix: Submission Issues

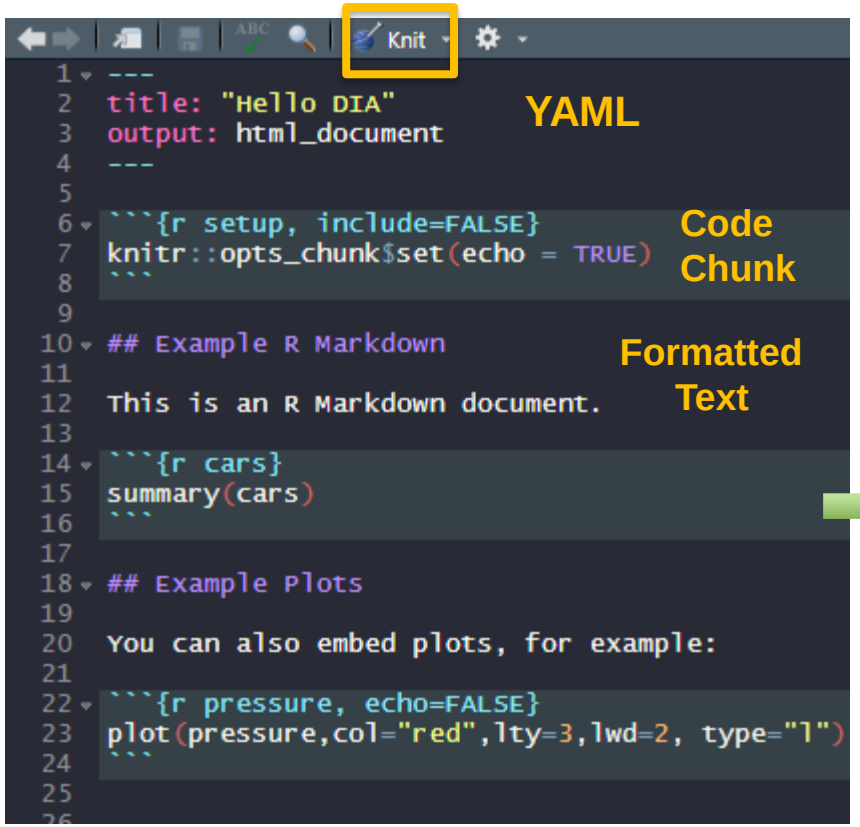
- **Data Submission with SAS XPORT (.xpt) version 8**
  - Comment: V8 is beneficial for lifting limits of characters. However, it will further increase file size, concerns on no native mechanism for audit trails, referencing source data. Although, it is advertised as nonproprietary, not sure whether it is vendor neutral.
  - Conclusion: Although FDA cannot require v8, accepting it is different. We decided to accept v8 by reviewing a sample data.
- **Data Submission through File Transfer Protocol (FTP)**
  - Comment: According to eCTD binding guidance† (section III M), the sponsor must use the FDA electronic submission gateway (ESG) for all submissions that are 10 gigabytes or smaller

    †Providing Regulatory Submissions in Electronic Format — Certain Human Pharmaceutical Product Applications and Related Submissions Using the eCTD Specifications (Feb 2020).
  - Conclusion: FDA does not accept regulatory submissions with File Transfer Protocol (FTP). The FDA ESG enables the secure submission of regulatory information for review and is our preferred method of transmission
- **Code Submission using R Markdown**
  - Conclusion: FDA agree with the program submission using R markdown. If your programs use R packages, specify alternative packages if applicable.

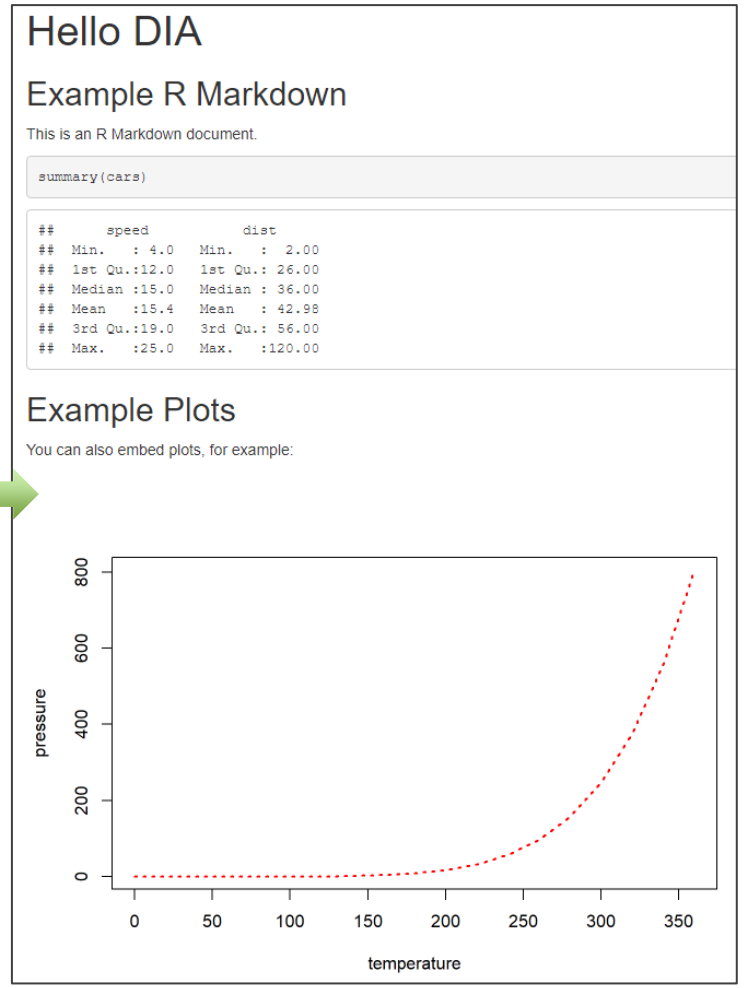
# Using R-Markdown for Regulatory Submission

## R Markdown

- Provides a unified framework for combining code, results, and comments
- Reproducible and shareable
- Comes pre-installed with RStudio




*. Rmd




*. html