



# Building a Deeper Bench: Writing ADaM Specifications

Shane Rosanbalm is a Senior Biostatistician at Rho. He started working in clinical trials while studying biostatistics at the University of North Carolina at Chapel Hill in a time before there were such things as hybrid cars, GPS, or texting.

Kristen Harrington is a Principal Programmer at Rho, who does not feel the need to put jokes in her bio; she defers to her co-author for that non-sense.

# Introduction

- ~~How do you teach employees to write ADaM specs?~~
- What goes into ***successfully*** training an employee to the point where they can ***independently*** write ***high-quality*** ADaM dataset specifications?

# Context

- Create protocol
- Create CRFs
- Create SAP/TLF mocks
- Create SDTM specifications
- Program SDTM
- ***Create ADaM specifications***
- Program ADaM
- Program TLFs
- Create define packages
- Fight with P21E

# Pre-requisites

- Programmed ADaM dataset for prior studies
  - No real ADaM expertise required to program
- Skim the ADaM model and IG documents
  - No more than 30 minutes per document
- Wrote SDTM specifications for the current study
  - A similar "building a deeper bench" process is in place for SDTM specification writing

# Training Overview

- In-person presentations (read, PowerPoint)
  - Big picture overview of ADaM model and IG
  - Informally annotating SAP and TLF mocks
  - Leveraging in-house gold standard metadata
  - Common conformance issues
  - Define and ADRG basics

# Training Overview

- Hands-on activities
  - Write specifications
  - Program datasets
  - Run P21E reports
- Support infrastructure
  - On-demand group chat
  - Weekly in-person check-in meetings
  - Timely **written** mentor feedback

# Skim This PDF

| Fragment |
|----------|
| GRy      |
| FL       |
| DT       |
| TM       |
| DTM      |

| Variable Name | Variable Label                          |
|---------------|-----------------------------------------|
| TRTSDT        | Date of First Exposure to Treatment     |
| TRTSTM        | Time of First Exposure to Treatment     |
| TRTSDTM       | Datetime of First Exposure to Treatment |
| TRTSDTF       | Date of First Exposure Input. Flag      |
| TRTSTMf       | Time of First Exposure Input. Flag      |

| Row | PARAM       | PARAMCD | AVISIT    | ABLFL | AVAL | BASE | CHG |
|-----|-------------|---------|-----------|-------|------|------|-----|
| 1   | Weight (kg) | WEIGHT  | Screening |       | 99   | 100  | .   |
| 2   | Weight (kg) | WEIGHT  | Run-In    |       | 101  | 100  | .   |
| 3   | Weight (kg) | WEIGHT  | Baseline  | Y     | 100  | 100  | 0   |
| 4   | Weight (kg) | WEIGHT  | Week 24   |       | 94   | 100  | -6  |
| 5   | Weight (kg) | WEIGHT  | Week 48   |       | 92   | 100  | -8  |
| 6   | Weight (kg) | WEIGHT  | Week 52   |       | 95   | 100  | -5  |

# Presentation Highlights

Old-school example of writing on paper.

## 8.2.1 Primary Endpoint

ADTTE

### 8.2.1.1 Overall Survival (OS)

PARAM (CD)

OS is defined as the time from the date of randomization to the date of death due to any cause. OS will be calculated in months as  $(\text{date of death} - \text{date of randomization} + 1) / 30.44$ .

$$AVAL = (ADT - STARTDT + 1) / 30.44$$

Subjects alive or lost to follow-up at the time of analysis will be censored at the date when they were last known to be alive. Subjects lacking data beyond randomization will have their OS censored at the date of randomization. Since the primary analysis will occur when the number of required events have been observed, the reasons for OS censoring at the time of analysis will be presented (Lost to follow-up/Withdrew consent, Ongoing, Other).

CNSR

CNSDTDSC

For the primary analysis of OS, subjects will not be censored at the time of subsequent non-protocol AML therapy including HSCT.

*New-school alternative: insert Comment in Word.*

# Presentation Highlights

## 6.1 Demographics and Baseline Characteristics

The following demographic and baseline characteristics will be summarized.

- Age (years)
- Age categories of (<60 years and  $\geq 60$  years; <65 years and  $\geq 65$  years,)
- Gender (male, and female)
- Race (White, Black or African American, Native Hawaiian or Other Pacific Islander, Asian, American Indian or Alaska Native, Other, and Not Reported)
- Ethnicity (Hispanic or Latino, Non-Hispanic or Latino, and Not Reported)
- Weight (kg)
- Height (cm)
- Body mass index ( $\text{kg}/\text{m}^2$ )
- Chemotherapy backbone (MEC or FAI)
- Eastern Cooperative Oncology Group (ECOG) performance status (0, 1, 2, and Other)
- Geographic region (North America, Europe, and Australia)

AGE

AGEGR1, AGEGR2

SEX

RACE

ETHNIC

WEIGHTBL

HEIGHTBL

BMIBL

CHEMOREG

ECOGBL

REGION

# Presentation Highlights

| CHG | CRIT1             | CRIT1FL |
|-----|-------------------|---------|
| 5   | >=20 bpm increase |         |
| 24  | >=20 bpm increase | Y       |
| 16  | >=20 bpm increase |         |

Verbose: populate  
CRIT1 for all records,  
populate CRIT1FL for  
the hits

When header N is  
the denominator

| CHG | CRIT1             | CRIT1FL |
|-----|-------------------|---------|
| 5   |                   |         |
| 24  | >=20 bpm increase | Y       |
| 16  |                   |         |

Minimalist: only  
populate the hits

| CHG | CRIT1             | CRIT1FL |
|-----|-------------------|---------|
| 5   | >=20 bpm increase | N       |
| 24  | >=20 bpm increase | Y       |
| 16  | >=20 bpm increase | N       |

Maximalist:  
populate the  
misses too

When visit n is  
the denominator

# Presentation Highlights

- Anytime post-baseline
  - Cannot use ONTRTFL as this has P21E checks with TRTSDT/TRTEDT
  - Use an analysis flag instead
    - ANL02FL
- Other examples
  - Leading up to Week 16
  - After Washout
  - During Follow-up

# Gold-standard Metadata

- Excel-based primer

| Name612  | Label612                 | ODMType | OriginVal                              | Notes                                                            |
|----------|--------------------------|---------|----------------------------------------|------------------------------------------------------------------|
| AVISIT   | Analysis Visit           | text    | Analysis visit.                        | Often involves visit windows. Add sap reference...               |
| AVISITN  | Analysis Visit (N)       | float   | Numeric version of analysis visit.     | For display sorting. If using VISITNUM, reset for ABLFL records. |
| ATPT     | Analysis Timepoint       | text    | Analysis timepoint.                    |                                                                  |
| ATPTN    | Analysis Timepoint (N)   | float   | Numeric version of analysis timepoint. |                                                                  |
| PARAM    | Parameter                | text    | Parameter.                             | Must include test, position, units, etc.                         |
| PARAMCD  | Parameter Code           | text    | Parameter code.                        | Check 1:1 with PARAM early.                                      |
| PARAMN   | Parameter (N)            | integer | Numeric version of parameter.          | For display sorting, but only if helpful.                        |
| PARCAT1  | Parameter Category 1     | text    | Parameter category.                    |                                                                  |
| PARCAT1N | Parameter Category 1 (N) | integer | Numeric version of parameter category. |                                                                  |
| AVAL     | Analysis Value           | float   | Analysis value.                        | <i>Either this...</i>                                            |
| AVALC    | Analysis Value (C)       | text    | Character analysis value.              | <i>...or this (but not both {within a PARAM}).</i>               |
| BASE     | Baseline Value           | float   | Baseline value.                        |                                                                  |
| BASEC    | Baseline Value (C)       | text    | Character baseline value.              |                                                                  |
| BASETYPE | Baseline Type            | text    | Baseline type.                         | If multiple baselines.                                           |
| CHG      | Change from Baseline     | float   | Change from baseline.                  |                                                                  |

# Do the Work

- Write specifications
- On-demand group chat
  - Trainee cohort and trainers
- Timely written mentor feedback
- Weekly in-person check-in meetings
- Program datasets
  - Mentor performs validation
- Run P21 checks

# Common Conformance Issues

!!!RUN P21 EARLY AND OFTEN!!!

- First draft of ADSL is validated?
  - Run P21
- In the middle of your first draft of ADLB?
  - Do your own programmatic checks on common issues
- First draft of all ADaM ready to send to sponsor?
  - Run P21
- First draft of define is created?
  - Run P21

# Common Issues

- Message: Inconsistent value for PARAM within a unique PARAMCD
- Source: ADLB.PARAMCD, ADLB.PARAM
- Why: Specs for PARAM say to concatenate LBTEST and LBSTRESU. Values for PARAMCD="ALB" include PARAM="Albumin (g/dL)" and PARAM="Albumin"

# Define and ADRG Basics

- Creating define.xml at Rho

## Rho Cheat Sheet to Create ADaM Define File

1) Before you can start specs, create a RhoME for your study, and create necessary study macros

- a. Run "Prog\Config\run\_createMetadataFiles"
- b. Always specify option "defineVersion=2"
- c. FDA no longer accepts define v1.0

2) Ensure all metadata is in RhoME. Make sure at least these columns are populated on all variables (these are the variables that are displayed in the define file):

- a. Name612: Variable Name
- b. Label612: Variable Label
- c. OriginType:
  - i. date (character date, ISO 8601, without time component)
  - ii. datetime (character datetime, ISO 8601, with date and time component)
  - iii. time (character time, ISO 8601, without date component)
  - iv. text (other character variables)
  - v. integer (numeric, no decimals)
  - vi. float (numeric, with decimals)

# ADRG Support

- Who is your audience?
  - FDA Reviewer
  - NOT a programmer or statistician or CDISC expert
  - Little to no knowledge of study
- Include anything that is...
  - Weird
  - Confusing
  - Important
  - Against CDISC Standards
  - Specific to this study/indication
  - Too long to fit in define

# ADRG Support

- Practical examples

## 3. Analysis Considerations Related to Multiple Analysis Datasets

### 3.1 Comparison of SDTM and ADaM Content

- Are data for screen failures, including data for run-in screening (for example, SDTM values of ARMCD='SCRNFAIL', or 'NOTASSGN') included in ADaM datasets?

Example 1:

Yes, the ARMCD value of 'SCRNFAIL' is used to identify subjects who failed randomization criteria. The ACTARMCD value of 'NOTTRT' is used to identify subjects who were randomized but were never treated.

Example 2:

No, screen failures (SDTM.DM.ARMCD='SCRNFAIL') are excluded from ADSL because no analyses were conducted on the population of all enrolled subjects. Analyses were only conducted on the safety population (all treated subjects) and the ITT population (all randomized subjects).

# ADRG Support

| Dataset(s) | Diagnostic Message and/or Check ID                                    | Severity | Count/<br>Issue Rate | BAD Explanations                                                                                                                                         | GOOD Explanations                                                                                                                                             |
|------------|-----------------------------------------------------------------------|----------|----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ADVS       | Secondary variable is present but its primary variable is not present | Error    | 6 (33.33%)           | It's fine<br><br>AVALCAT1 is in dataset, but AVALCAT1N is not.<br><br>AVALCAT1N has 9 characters                                                         | Character variable names that contain 8 characters must be shortened to add the suffix 'N' and create the secondary variable. For example, AVALCAT1/AVALCA1N. |
| ADSL       | RACE value not found in 'Race' extensible codelist                    | Warning  | 2071 (0.53%)         | See CRF pg 5<br><br>Values are same as SDTM<br><br>Values are directly from CRF, and do not match CDISC CT. See DERIVE.ADSL where INDEX(RACE,'OTHER')>0. | RACE values of 'OTHER' and 'MULTIPLE' are included as collected on the CRF, despite not conforming to the expected controlled terminology                     |
| ADMI       | Variable label mismatch between dataset and ADaM standard             | Error    | 1235 (100.00%)       | IG is wrong<br><br>Label was copied directly from SDTM MI dataset<br><br>Variable label in IG is more than 40 characters                                 | SDTM IG v3.2 provides a variable label that is > 40 characters for MISTRESC. Label is shortened to meet 40 character limit by changing 'Standard' to 'Std'    |

# Conclusion

- Context matters
- Pre-requisites exist
- A series of group trainings occur
- Work happens
- Mentoring, feedback, and questions
- Successful submission; drug approved!