



CDISC ARS Template Code: Gearing up with automation in TFL Programming

Malan Bosman, Lead Automation Engineer

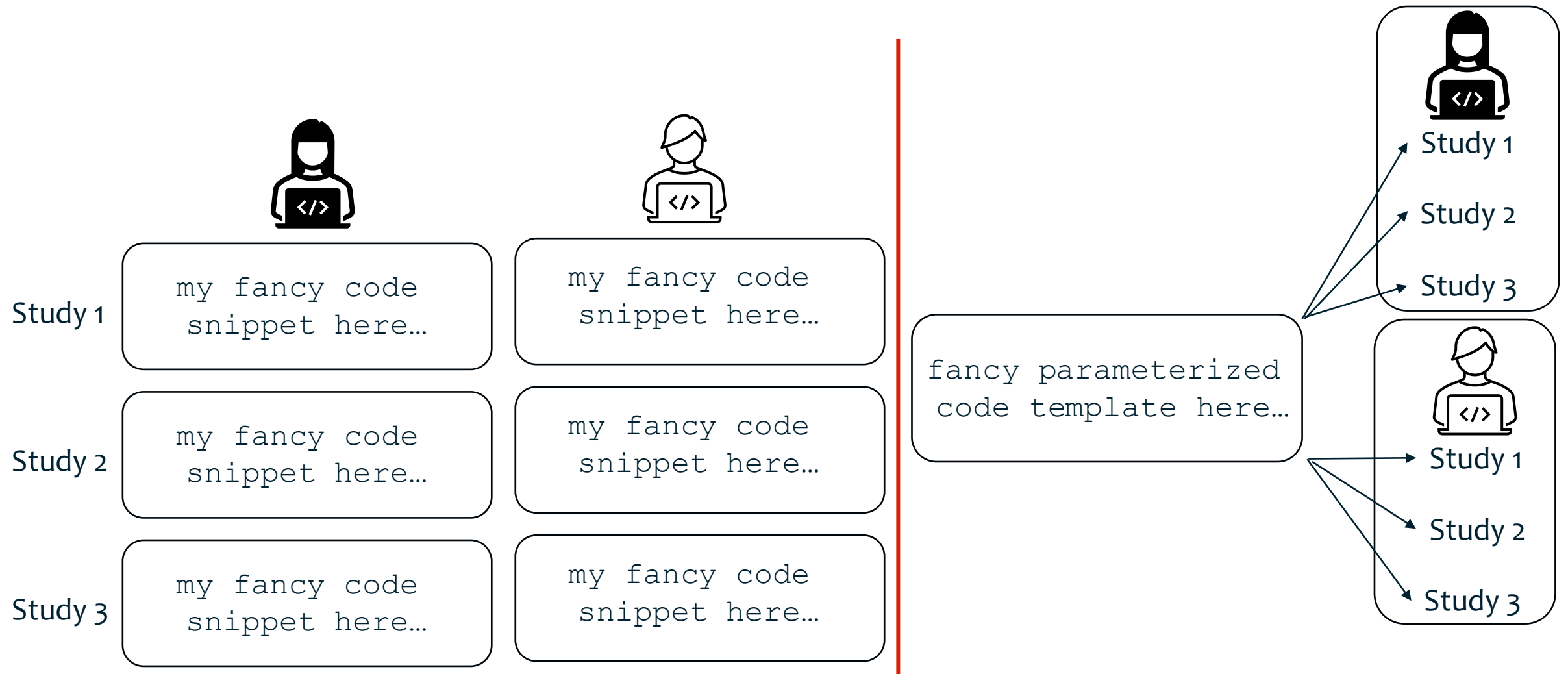
Anna Yaggi, Product Manager

Clymb Clinical

Outline

- The opportunity: re-use code!
- CDISC ARS
- AnalysisMethodCodeTemplate
- CARDS and CARDX Packages
- Implementation and Examples
- Key Takeaways + Next Steps

The opportunity: re-use code!



CDISC Analysis Results Standard - a Logical Data Model



Analysis Results Standard (ARS) v1.0



Large trials generate many analysis results in the form of tables, figures, and written reports, yet these results are rarely output in a form that is machine-readable. Previously, there has been no standard way of describing and organizing these results, making it difficult to automate their generation, make them reproducible, trace their origin, or enable them to be reused in other outputs.

To address these inefficiencies, CDISC has developed the [Analysis Results Standard \(ARS\)](#), which aim to facilitate automation, reproducibility, reusability, and traceability of analysis results data.

Features of ARS v1.0

- A Logical Data Model that describes analysis results and associated metadata.
- A User Guide to illustrate and exercise the model with common safety displays.

<https://cdisc-org.github.io/analysis-results-standard/>

Analysis Results Standard (ARS)

Schema Diagram

Classes

Slots

Enumerations

Types

Subsets

Analysis Results Standard (ARS)

Logical model to support both the prospective specification of analyses and the fully contextualized representation of the results of the analyses.

URI: <https://www.cdisc.org/ars/1-0> Name: ars_idm

Schema Diagram

Classes

Classes provide templates for organizing data. Data objects instantiate slots (aka fields, attributes) that are applicable to it. See [LinkML docs](#)

Class	Description
NamedObject	An object with a name
ReportingEvent	A set of analyses and o requirements...
ListOfContents	A structured list of ana

cdisc

Analysis Results Standard User Guide

Version 1.0 (Final)

Prepared by the
Analysis Results Standard Team

Notes to Readers

- This is the final Version 1.0 of the Analysis Results Standard User Guide.
- This document is based on ADaM v2.1 and Analysis Results Metadata (ARM) v1.0 for Define-XML v2.0

Revision History

Date	Version
2024-04-19	Final

© 2024 Clinical Data Interchange Standards Consortium, Inc. All rights reserved.

<https://wiki.cdisc.org/display/ARSP/Analysis+Results+Standard+User+Guide+v1.0>

Model Components

Analysis Set

Data Subset

Analysis Grouping

Analysis Method

Analysis

Results

Summary of Demographics

Study - CDISC 360

Page x of y

Table 14.1.1
Summary of Demographics
Safety Population

Characteristics	Placebo (N=XX)	Xanomeline Low Dose (N=XX)	Xanomeline High Dose (N=XX)
Age (years)			
n	XX	XX	XX
Mean (SD)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
Median	XX.X	XX.X	XX.X
Q1, Q3	XX.X, XX.X	XX.X, XX.X	XX.X, XX.X
Min, Max	XX, XX	XX, XX	XX, XX
Age Group, n (%)			
< 65 years	XX (XX.X)	XX (XX.X)	XX (XX.X)
≥ 65 years	XX (XX.X)	XX (XX.X)	XX (XX.X)
Gender, n (%)			
Male	XX (XX.X)	XX (XX.X)	XX (XX.X)
Female	XX (XX.X)	XX (XX.X)	XX (XX.X)
Ethnicity, n (%)			
Hispanic or Latino	XX (XX.X)	XX (XX.X)	XX (XX.X)
Not Hispanic or Latino	XX (XX.X)	XX (XX.X)	XX (XX.X)

Output

Source dataset: adsl, Generated on: DDMONYYYY:HH:MM
Program: <pid>.sas, Output: <pid><oid>.rtf, Generated on: DDMONYYYY:HH:MM

Output

Summary of TEAE by SOC and PT

Study - CDISC 360

Page x of y

Table 14.3.1.1
Summary of TEAE by System Organ Class and Preferred Term
Safety Population

System Organ Class	Placebo	Xanomeline Low Dose	Xanomeline High Dose
Preferred Term [a], n (%)	(N=XX)	(N=XX)	(N=XX)
Number of subjects with at least one event	XX (XX.X)	XX (XX.X)	XX (XX.X)
<SOC 1>	XX (XX.X)	XX (XX.X)	XX (XX.X)
<Preferred Term 1>	XX (XX.X)	XX (XX.X)	XX (XX.X)
...	XX (XX.X)	XX (XX.X)	XX (XX.X)
<Preferred Term n>	XX (XX.X)	XX (XX.X)	XX (XX.X)
<SOC 2>	XX (XX.X)	XX (XX.X)	XX (XX.X)
<Preferred Term 1>	XX (XX.X)	XX (XX.X)	XX (XX.X)
...	XX (XX.X)	XX (XX.X)	XX (XX.X)
<Preferred Term n>	XX (XX.X)	XX (XX.X)	XX (XX.X)

Notes: TEAE=Treatment-Emergent Adverse Events.
Subjects are counted once within each system organ class and preferred term.
[a] All investigators adverse events were coded using MedDRA version xx.x.

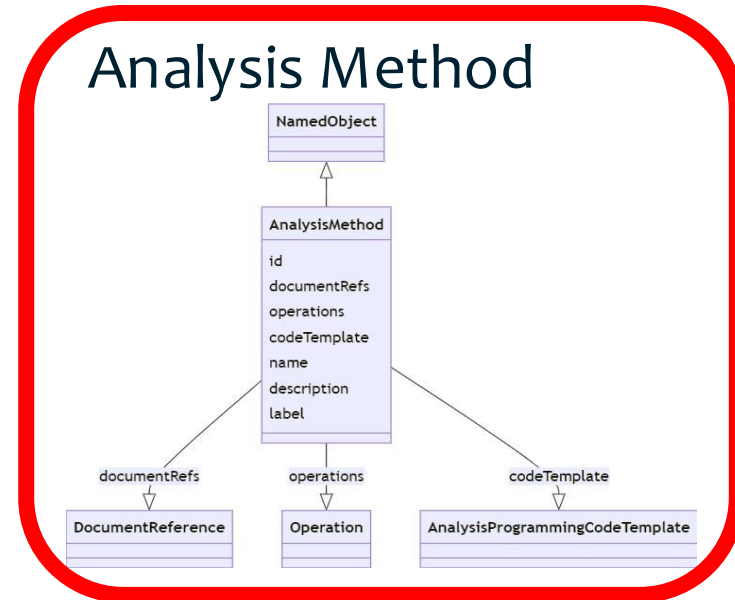
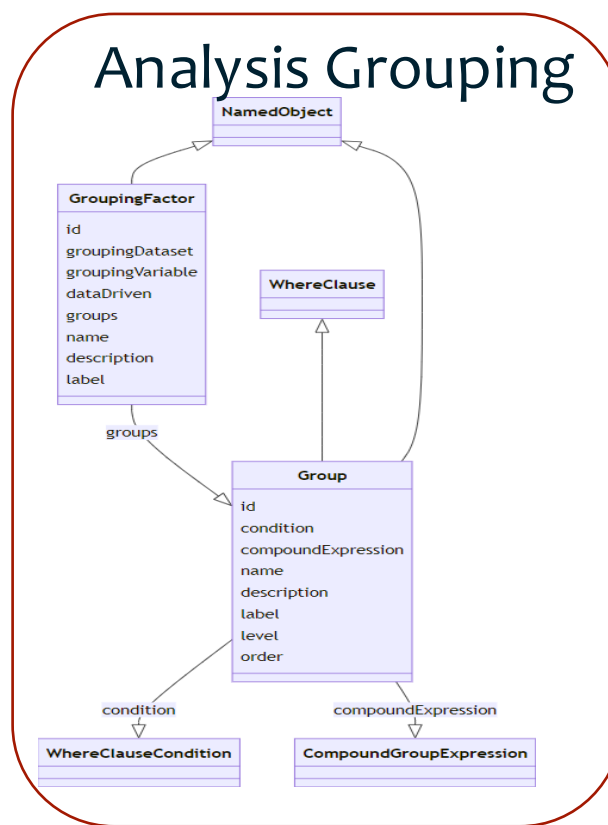
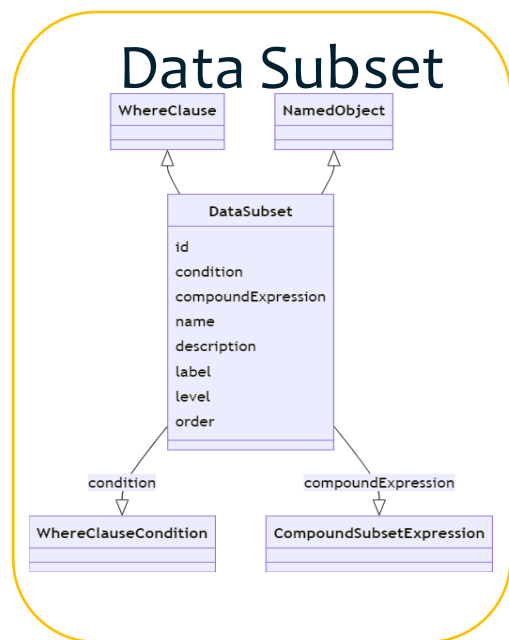
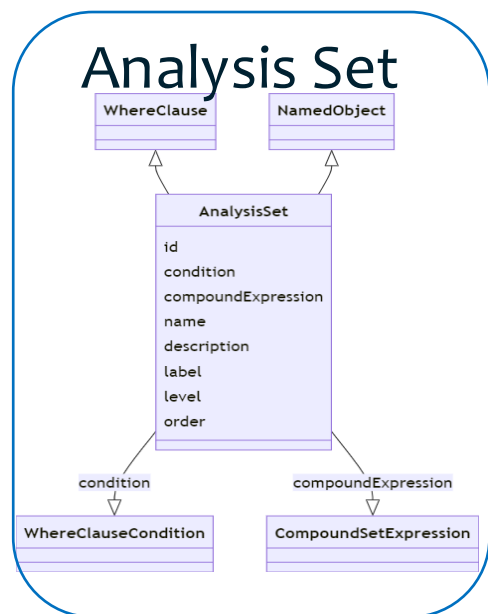
Source dataset: adae, Generated on: DDMONYYYY:HH:MM
Program: <pid>.sas, Output: <pid><oid>.rtf, Generated on: DDMONYYYY:HH:MM

Notes: TEAE=Treatment-Emergent Adverse Events.
Subjects are counted once within each system organ class and preferred term.
[a] All investigators adverse events were coded using MedDRA version xx.x.

Source dataset: adae, Generated on: DDMONYYYY:HH:MM
Program: <pid>.sas, Output: <pid><oid>.rtf, Generated on: DDMONYYYY:HH:MM

CDISC ARS Metadata

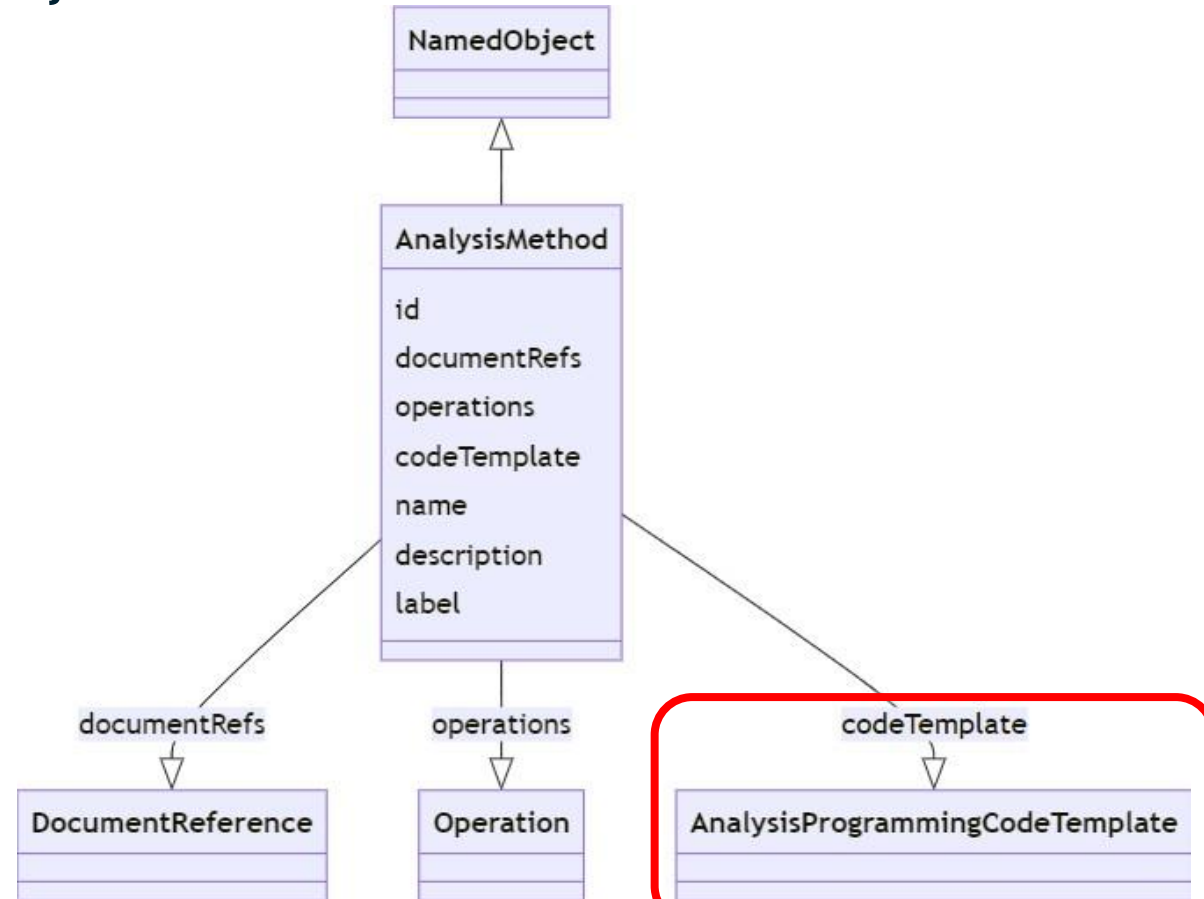
One **Output** has many **Analyses**. One **Analysis** has the following:



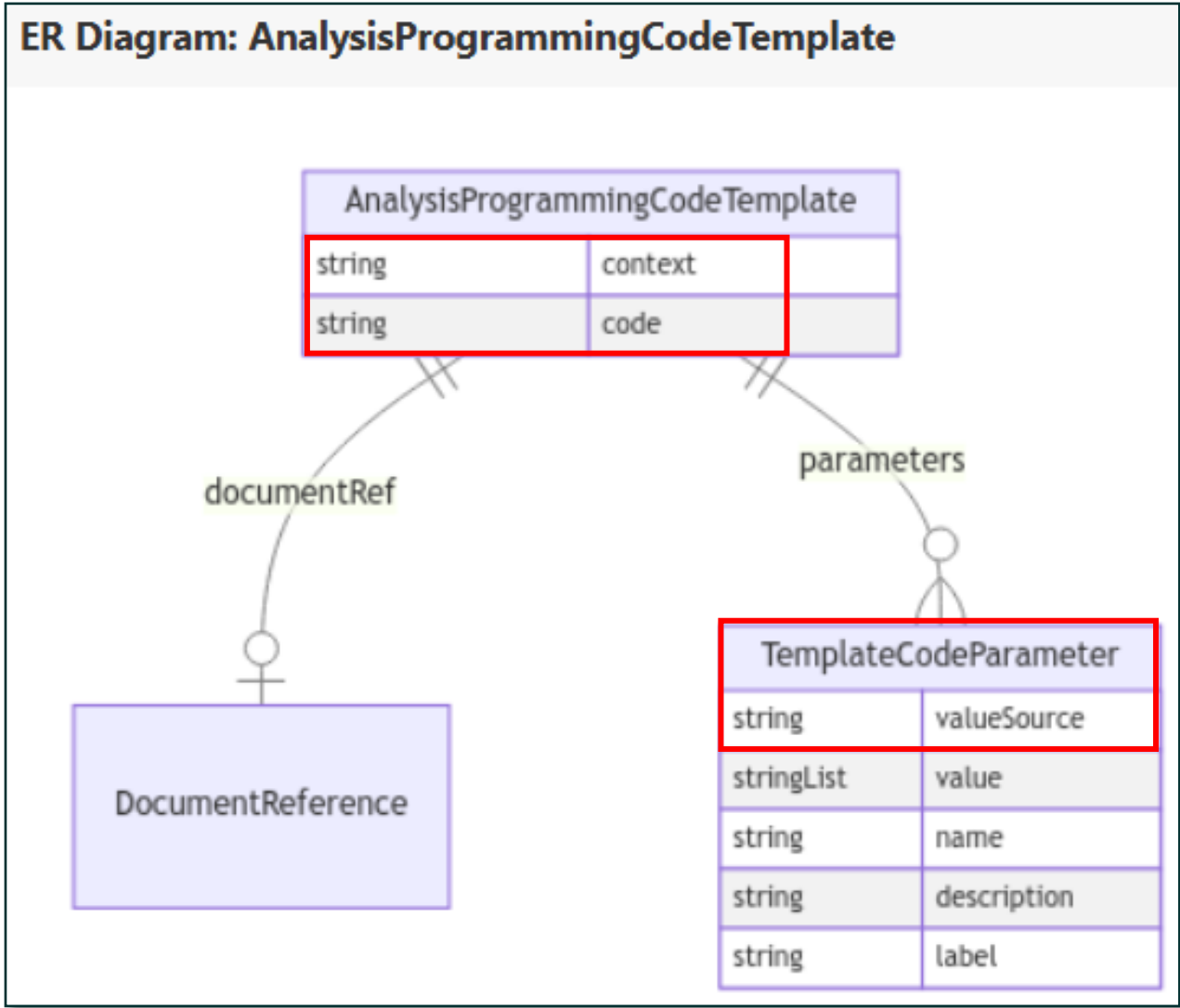
One **Analysis Method** consists of one or more **Operations**

CDISC ARS Metadata

Analysis Method



AnalysisMethodCodeTemplate

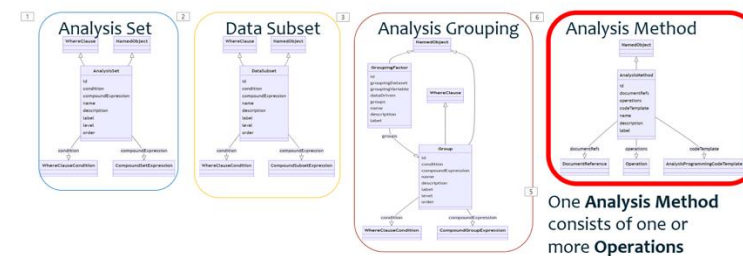


In one sentence...

Results for each analysis...

Characteristics	Xanomeline Med Dose (N=XX) n (%)	Xanomeline High Dose (N=XX) n (%)	Placebo (N=XX) n (%)
Age			
n	XX	XX	XX
Mean (SD)	XX.X (XX.XX)	XX.X (XX.XX)	XX.X (XX.XX)
Median	XX.X	XX.X	XX.X

Are described in the ARS Metadata...



Which includes AnalysisMethods

id	name	operation_id	operation_name	operation_description	operation_label	operation_order	operation_resultPattern
Mth02_ContVar_Summ_ByGrp	Summary by group of a continuous variable	Mth02_ContVar_Summ_ByGrp_1_n	Count of non-missing values	n	1 XX		
Mth02_ContVar_Summ_ByGrp	Summary by group of a continuous variable	Mth02_ContVar_Summ_ByGrp_2_Mean	Mean	Mean	2 XXX		
Mth02_ContVar_Summ_ByGrp	Summary by group of a continuous variable	Mth02_ContVar_Summ_ByGrp_3_SD	Standard deviation	SD	3 (XX.XX)		
Mth02_ContVar_Summ_ByGrp	Summary by group of a continuous variable	Mth02_ContVar_Summ_ByGrp_4_Median	Median	Median	4 XXX		
Mth02_ContVar_Summ_ByGrp	Summary by group of a continuous variable	Mth02_ContVar_Summ_ByGrp_5_Q1	First quartile	Q1	5 XXX		
Mth02_ContVar_Summ_ByGrp	Summary by group of a continuous variable	Mth02_ContVar_Summ_ByGrp_6_Q3	Third quartile	Q3	6 XXX		
Mth02_ContVar_Summ_ByGrp	Summary by group of a continuous variable	Mth02_ContVar_Summ_ByGrp_7_Min	Minimum	Min	7 XX		
Mth02_ContVar_Summ_ByGrp	Summary by group of a continuous variable	Mth02_ContVar_Summ_ByGrp_8_Max	Maximum	Max	8 XX		

Linked to Analysis Method Code

method_id	context	specifiedAs	templateCode
Mth02_ContVar_Summ_ByGrp R		Code	<pre>df3_analysisidhere <- cards::ard_continuous(data = df2_analysisidhere, by = c(byvarshere), variables = anavarhere) > dplyr::mutate(operationid = dplyr::case_when(stat_name == 'N' ~ 'Mth_03_01_n', stat_name == 'mean' ~ 'Mth_03_02_Mean', stat_name == 'sd' ~ 'Mth_03_03_SD', stat_name == 'median' ~ 'Mth_03_04_Median', stat_name == 'p25' ~ 'Mth_03_05_Q1', stat_name == 'p75' ~ 'Mth_03_06_Q3', stat_name == 'min' ~ 'Mth_03_07_Min', stat_name == 'max' ~ 'Mth_03_08_Max'])</pre>

To create runnable code at run-time

```
ard_continuous(
  data = df2_An_03,
  by = c('TRT01AN'),
  variables = AGE
) |>
```

Containing parameters referencing ARS Metadata elements

method_id	parameter_name	parameter_description	parameter_valueSource	parameter_value
Mth02_ContVar_Summ_ByGrp	byvarshere	string for "by" parameter for the ard_continuous function	by_listc	
Mth02_ContVar_Summ_ByGrp	anavarhere	Analysis variable to be summarised	ana_var	

Key Features for Template Based Automation

1. Consistent Interface
2. Standardized ARD Output
3. Built-In Metadata
4. Flexible Parameterization

▼ Programming Code

Context

R (siera) ×☐ From existing document reference

Code

denom_dataset = df2_denomaidhere

in_data = df2_analysisidhere |>
 dplyr::distinct(distinctlisthere) |>
 dplyr::mutate(dummy = 'dummyvar')

Parameter 1 

Parameter Name	Parameter Label	Parameter Value	Parameter Value Source
distinctlisthere	Enter Parameter Label	Enter Parameter Value	distinct_list

Parameter Description

list of variables by which to apply the "distinct" function from dplyr

Extracting ARS Metadata from TFL Designer

▼ Programming Code

Context

SAS Version 9.4 ☐ From existing document reference

Code

%CatVar_Count_ByGrp(anid=&anid,gfids=&gfids);

Parameter Name	Parameter Label	Parameter Value	Parameter Value Source
anid	Analysis Id	Enter Parameter Value	#/id
gfids	Grouping Factor Identifie	Enter Parameter Value	#/orderedGroupings/[*]/

```

methods:
- id: Mth01_CatVar_Count_ByGrp
  name: Count by group for a categorical variable
  operations:
  - id: Mth01_CatVar_Count_ByGrp_1_n
    name: Count of subjects
    label: n
    order: 1
    resultPattern: (N=XXX)
  codeTemplate:
    context: SAS Version 9.4
    code: "%CatVar_Count_ByGrp(anid=&anid,gfids=&gfids);"
    parameters:
    - name: anid
      description: Analysis Id
      valueSource: '#/id'
    - name: gfids
      description: Grouping Factor Identifiers
      valueSource: '#/orderedGroupings/[*]/groupingId'

```

Metadata-based Parameterization

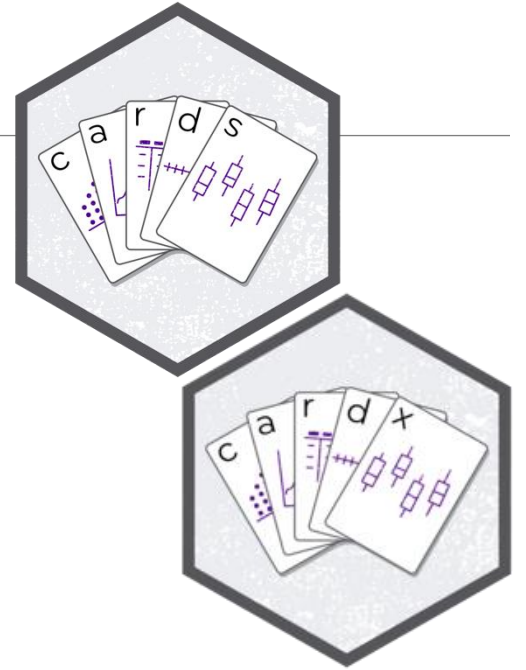
- Template programming code parameter values may be specified as a reference to a metadata element
- This means analysis-specific parameters values can be retrieved at runtime
- Generation of analysis result values can be automated based on the selection of analysis method

CARDS and CARDX Packages

- **Purpose-built for ARD generation:**
- *cards* and *cardx* are part of the pharmaverse ecosystem and designed specifically to generate Analysis Results Datasets (ARDs).
- They wrap common statistical procedures and return results in a standardized ARD format, making them ideal for ARS-based automation.
- **cards package (core ARD functionality):**
 - **ard_summary()** – Summary statistics for continuous variables
 - **ard_tabulate()** – Frequency and count summaries for categorical variables
 - **ard_hierarchical()** – Subject-level summaries for nested or hierarchical data
 - **ard_stack()** – Combines multiple ARD outputs into a single dataset



CARDS and CARDX Packages



■ cardx package (extended methods):

- Adds additional statistical techniques, hypothesis tests, and specialized analyses
- Expands ARD coverage beyond the core summaries in *cards*

■ Combined value:

- Together, *cards* and *cardx* provide a comprehensive, ARS-aligned toolkit for standardized ARD generation

Practical Implementation

Example: Summary statistics

```
Analysis_ARD <- cards::ard_summary(data = filtered_data,  
                                   by = c(byvariables_here),  
                                   variables = analysisvariable_here )
```

```
Analysis_ARD <- cards::ard_summary(data = filtered_ADSL,  
                                   by = c(TRT01A),  
                                   variables = AGE)
```

Example: Chi-square p-value

```
df3_analysisidhere <-  
  cardx::ard_stats_chisq_test(by = groupvar1here,  
                              data = df2_analysisidhere,  
                              variables = groupvar2here) |>  
  dplyr::filter(stat_name == "p.value") |>  
  dplyr::mutate(operationid = dplyr::case_when(  
    stat_name == 'p.value' ~ opid1here))
```

Example: Fisher's Exact Test

```
full = df_pop |>
  dplyr::filter(armfilterhere)

success = df2_analysisidhere |>
  dplyr::mutate(FL = 1)

ana = full |>
  dplyr::left_join(success |>
    dplyr::select(anavarhere,
                  FL),
    by = "anavarhere") |>
  dplyr::mutate(FL = ifelse(is.na(FL), 0, FL))

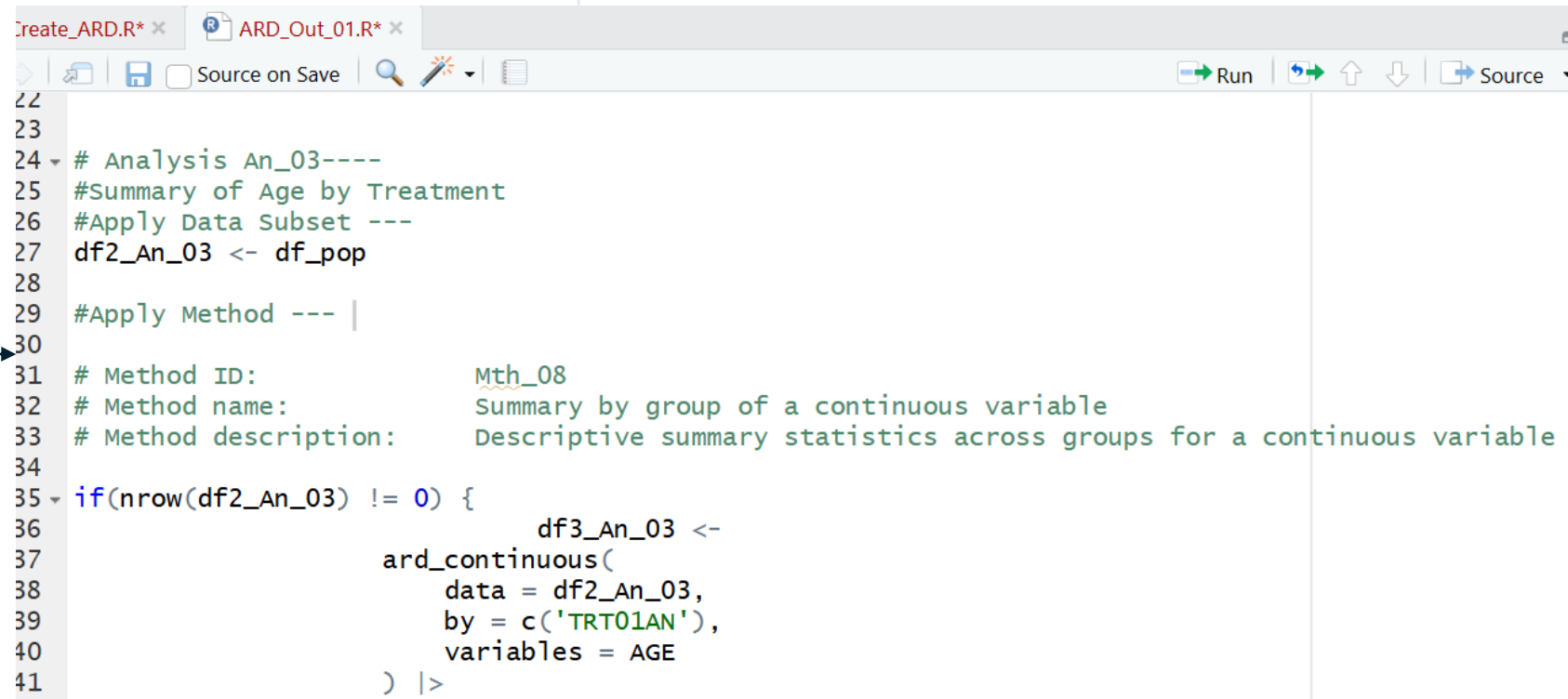
df3_analysisidhere = ana |>
  cardx::ard_stats_fisher_test(by = groupvar1here,
                              variables = FL) |>
  dplyr::filter(stat_name %in% c('estimate',
                                'conf.low',
                                'conf.high')) |>
  dplyr::mutate(operationid = dplyr::case_when(
    stat_name == 'estimate' ~ 'opid1here',
    stat_name == 'conf.low' ~ 'opid2here',
    stat_name == 'conf.high' ~ 'opid3here'))
```

Example automation engine: `siera`

```
library(siera)
```

```
ars_path_json <- "C:/Users/mbosm/OneDrive - Clymb Clinical/Documents/TFL_ARS/02MAR2026/demo1.json"
adampath <- "C:/Users/mbosm/OneDrive - Clymb Clinical/Documents/TFL_ARS/02MAR2026/csv_adam"
outputpath <- "C:/Users/mbosm/OneDrive - Clymb Clinical/Documents/TFL_ARS/02MAR2026/ard_programs"
```

```
readARS(  
  ARS_path = ars_path_json,  
  adam_path = adampath,  
  output_path = outputpath  
)
```



```
Create_ARD.R* x ARD_Out_01.R* x  
Source on Save  
22  
23  
24 # Analysis An_03----  
25 #Summary of Age by Treatment  
26 #Apply Data Subset ---  
27 df2_An_03 <- df_pop  
28  
29 #Apply Method --- |  
30  
31 # Method ID: Mth_08  
32 # Method name: Summary by group of a continuous variable  
33 # Method description: Descriptive summary statistics across groups for a continuous variable  
34  
35 if(nrow(df2_An_03) != 0) {  
36     df3_An_03 <-  
37     ard_continuous(  
38         data = df2_An_03,  
39         by = c('TRT01AN'),  
40         variables = AGE  
41     ) |>
```

Benefits

Efficiency Gains

Improved Quality and Consistency

Enhanced Maintainability

Improved Traceability

Facilitates Validation

Considerations

Metadata Quality and Completeness

Technical Infrastructure

Change Management

Edge Cases and Flexibility

Transparency and Review

Future Directions

1. Enhanced Template Libraries
2. AI-Assisted Metadata Creation
3. Expansion to Other Languages
4. ARD-driven processes





Malan Bosman

Lead Data Automation Engineer

mbosman@clymbclinical.com

Anna Yaggi

Product Owner

ayaggi@clymbclinical.com



EMAIL

info@clymbclinical.com



WEBSITE

www.clymbclinical.com



PHONE

781-692-3613