

ADaM Data in High Definition

Varia Cartledge, Alexion Pharmaceuticals, Boston, USA

Overview

- Why is a good define important?
- Who should be involved in define creation?
- What are the basic style guidelines for clear define content
- How did my team approach Analysis Results Metadata?
- When will this presentation end?

Why is a good define important?

- A define:
 - Is a mandatory submission deliverable
 - Meets conformance guidelines and can be validated
- A *good* define:
 - Gives reviewers insight into your data
 - Allows reviewers to find the supporting evidence they need to understand the story your data tells
 - Provides clear and specific traceability, boosting confidence in your analysis

Who should be involved in define creation?

- Often define is treated as an after-the-fact deliverable and prepared by a more junior programmer based on metadata extracted from datasets and information from specs
- Lead study programmers and statisticians are ultimately responsible for the content of definitions and algorithms
- To avoid a step where junior programmers attempt to edit the spec contents to provide clarity, lead programmers and statisticians should be trained on an in-house style guide

What are the basic style guidelines for clear define content?

- Completeness
 - Include all information needed to show the traceability of a variable
 - If it's not possible to clearly describe how a variable was created in the define, use additional documentation such as the ADRG
 - Don't use cross-references to information in other variable or dataset definitions

'="Y" if record is the analysis record for baseline and the scheduled relapse VISITs. Refer Tab ANLxxFL_SV for visit info.

What are the basic style guidelines for clear define content?

- Consistency
 - Be consistent in the use of wording and punctuation
 - For example, pick a style for quotation marks and use it throughout the define
 - Determine the style of pseudocode you want and use it throughout
 - Avoid specific SAS language such as “in”, “floor”, “.” in place of “missing”
 - Use a consistent approach to dataset and variable name references
 - Include all information needed for traceability, e.g. including protocol numbers if data are being carried from another study

```
if ASTDT lt TR01SDT then ASTDT -  
TR01SDT;  
else if ASTDT >= TR01SDT then ASTDT -  
TR01SDT +1;
```


Basic style guidelines, continued

- Keep it clean
 - Avoid special characters which may not process correctly during any automation
 - Do not use extended ascii or non-ascii characters such as slanted quotes, double quotes, carriage returns, extra spaces, bullets
 - Do not use Greek characters or mathematical symbols such as the degree symbol
 - Don't use formatting such as bolding text, highlighting text, italics, or strikethrough
 - Strikethrough can disappear when being imported from a spreadsheet to an automation tool, leaving the incorrect information in the definition!
 - Don't use reference paths to internal directories (e.g. references to lookup tables stored in an internal documentation directory)



How did my team approach Analysis Results Metadata?

- The purpose of ARM is to give reviewers a clear and concise way to understand the traceability and algorithms used to create key endpoint results
- To that end, better to be selective in what to present rather than attempt to be comprehensive
 - PMDA Technical Conformance Guide (Jan 2019) “main results of efficacy, safety and clinical study results that provide the rationales for setting of the dosage and administration”
 - This is good guidance for now, it would be smart to build a company standard as to what this encompasses
 - However, ultimately result selection will always be study specific
- The difference between “results” and “displays” is important to communicate – not necessarily intuitive
 - “When analysis results metadata is provided for a specific result within (a) display, there is no requirement to describe all individual results within a display. For example, analysis results metadata may be provided for complex statistics (e.g., p-values) but not necessarily for descriptive statistics like mean or median.” (Analysis Results Metadata Specification, Version 1.0, CDISC ADaM Metadata Sub-Team, January 27, 2015)

How did my team approach Analysis Results Metadata?

- Lead programmer lobbied for inclusion of ARM because of submission to PMDA and opportunity to early adopt
- Kickoff meeting to familiarize team with ARM concepts – important to show examples and clarify how our results metadata would be presented
- Essential to engage statisticians to identify what results are the most critical to telling the story of the data

Standard	ADaM-IG 1.0
Study Name	CDISC-Sample
Study Description	CDISC-Sample Data Definition
Protocol Name	CDISC-Sample
Metadata Name	Study CDISC-Sample, Data Definitions
Metadata Description	Study CDISC-Sample, Data Definitions

Analysis Results Metadata (Summary) for Study CDISC-Sample

Table 14-3.01 Primary Endpoint Analysis: ADAS-Cog - Summary at Week 24 - LOCF (Efficacy Population) Dose response analysis for ADAS-Cog changes from baseline Pairwise comparisons to placebo for ADAS-Cog changes from baseline
Table 14-5.02 Incidence of Treatment Emergent Serious Adverse Events by Treatment Group Incidence of Treatment Emergent Serious Adverse Events by Treatment Group

Analysis Results Metadata (Detail) for Study CDISC-Sample

Table 14-3.01

Display	Table 14-3.01 Primary Endpoint Analysis: ADAS-Cog - Summary at Week 24 - LOCF (Efficacy Population)
Analysis Result	Dose response analysis for ADAS-Cog changes from baseline
Analysis Parameter(s)	PARAMCD = "ACTOT" (Adas-Cog(11) Subscore)
Analysis Variable(s)	CHG (Change from Baseline)
Analysis Reason	SPECIFIED IN SAP
Analysis Purpose	PRIMARY OUTCOME MEASURE
Data References (incl. Selection Criteria)	ADQSADAS [PARAMCD = "ACTOT" and AVISIT = "Week 24" and EFFFL = "Y" and ANL01FL = "Y"]
Documentation	Linear model analysis of CHG for dose response; using randomized dose (0 for placebo; 54 for low dose; 81 for high dose) and site group in model. Used PROC GLM in SAS to produce p-value (from Type III SS for treatment dose). SAP Section 10.1.1
Programming Statements	[SAS version 9.2] proc glm data = ADQSADAS; where EFFFL='Y' and ANL01FL='Y' and AVISIT='Week 24' and PARAMCD="ACTOT"; class SITEGR1; model CHG = TRTPN SITEGR1; run;

How did my team approach Analysis Results Metadata?

- Created a template for ARM from Pinnacle21 Enterprise define tool for use during kickoff to record the selected displays and results
 - Decided on a simple convention of using the Display name plus “Rxx” for the Result ID
 - Found it useful to add a leading zero to the result ID numbering for correct sorting within the define tool
- This template was then used to build the remaining fields for ARM by define programmer, study programmers, and statisticians
 - Statisticians selected which text from SAP was appropriate for Documentation input
 - Study programmers identified SAS code to include – we preferred to present it in the define as a specific piece of a program rather than link to whole program
 - Other fields such as Reason and Purpose were simple for the define programmer to complete

Display	ID	Description
Table 14.1.1.2.1	Table 14.1.1.2.1 R01	Baseline EDSS
Table 14.1.1.2.1	Table 14.1.1.2.1 R02	Baseline HAI S
Table 14.1.1.2.1	Table 14.1.1.2.1 R03	Baseline mRS
Table 14.1.3.2.1	Table 14.1.3.2.1 R01	Number of rel
Table 14.1.3.2.1	Table 14.1.3.2.1 R02	Historical Ann

How did my team approach Analysis Results Metadata?

- Template could be loaded directly into Pinnacle21 Enterprise when complete
 - Except for the manual Selection Criteria (where clause) builder:

The screenshot shows the 'Selection Criteria Builder' interface. It features four input fields: 'Dataset' with the value 'ADEFF', 'Variable' with the value 'PARAMCD', 'Comparator' with the value 'EQ', and 'Value' with the value 'ARRHS24'. An 'and' button is positioned to the right of the 'Value' field. Below these fields, the 'Selection Criteria' section displays the generated query: 'ADEFF[PARAMCD EQ 'ARRHS24']'. At the bottom right, there are 'Submit' and 'Cancel' buttons.

Dataset	Variable	Comparator	Value
ADEFF	PARAMCD	EQ	'ARRHS24'

and

Selection Criteria
ADEFF[PARAMCD EQ 'ARRHS24']

Submit Cancel

How did my team approach Analysis Results Metadata?

- The template generated from Pinnacle21 Enterprise was a reasonably effective communication tool
- Statisticians and study programmers only edited the Analysis Display and Analysis Results tabs; other tabs including Analysis Display Criteria and Comments tab were managed by define programmer
- A generated draft of ARM as they would appear in the final define was the most useful review tool
- An additional pdf document was created in the datasets folder containing SAP extracts to avoid linking between folders in module 5

When will this presentation be over? (Conclusion)

- Define.xml for ADaM is a powerful tool for creating data submissions that will be easy to navigate and understand
- Creating the Analysis Results Metadata section adds value for reviewers
- The construction of a well-formed define requires a level of technical expertise that often involves a dedicated define programmer
- However, the work of developing the key content of the define belongs in the hands of the study programmers and study statisticians
- Ensuring that programmers and statisticians across organizations have a fundamental understanding of the content and purpose of the define is key to:
 - Building the value of the define
 - Creating workflows that focus on what is meaningful in define creation

When will this presentation be over?

