

Single Day Event

Efficient Regulatory Submissions with ARD: Leveraging R for Structured Results

Madhan Kumar N
ICON PLC



Disclaimer

The Opinions expressed in this presentation and on the following slides are solely those of the presenter and not necessarily those of ICON. ICON does not guarantee the accuracy or reliability of the information provided herein.



Agenda

- ❑ Workflows and Standards
- ❑ Analysis Results Standards (ARS)
- ❑ Analysis Results Workflow
- ❑ Principles and Benefits
- ❑ Introduction to {cards} package
- ❑ Vital Signs Example
- ❑ Limitations of ARS
- ❑ References



Workflows and Standards

STUDYID	DOMAIN	USUBJID	SUBJID	RFSTDTC	RFENDTC	RFSTDTC	RFENDTC	RFSTDTC	RFENDTC
1	CDISCPILOT01	DM	01-701-1015	1015	2014-01-02	2014-07-02	2014-01-02	2014-07-02	<NA>
2	CDISCPILOT01	DM	01-701-1023	1023	2012-06-05	2012-09-02	2012-06-05	2012-09-01	2013-02-18
3	CDISCPILOT01	DM	01-701-1028	1028	2013-07-19	2014-01-14	2013-07-19	2014-01-14	<NA>
4	CDISCPILOT01	DM	01-701-1033	1033	2014-03-18	2014-04-14	2014-03-18	2014-03-18	2014-09-15
5	CDISCPILOT01	DM	01-701-1034	1034	2014-07-01	2014-12-30	2014-07-01	2014-12-30	2014-12-30T09:50
6	CDISCPILOT01	DM	01-701-1047	1047	2013-02-12	2013-03-29	2013-02-12	2013-03-29	2013-07-28
7	CDISCPILOT01	DM	01-701-1057	1057	<NA>	<NA>	<NA>	<NA>	2013-12-27
8	CDISCPILOT01	DM	01-701-1097	1097	2014-01-01	2014-07-09	2014-01-01	2014-07-09	2014-07-09T09:39
9	CDISCPILOT01	DM	01-701-1111	1111	2012-09-07	2012-09-17	2012-09-07	2012-09-16	2013-02-22
10	CDISCPILOT01	DM	01-701-1115	1115	2013-11-30	2013-01-23	2013-11-30	2013-01-23	2013-05-20

SDTM

Standards for Study
Data

STUDYID	USUBJID	SUBJID	STED	AGE	AGEU	SEX	RACE	ETHNIC	ARMCD	ARM	
AE	AE	A	-				AE	AE	A	AE	
	CDISCPILOT01	01-701-1015	1015	701	63	YEARS	F	WHITE	HISPANIC OR LATINO	Plac	Placebo
2	CDISCPILOT01	01-701-1023	1023	701	64	YEARS	M	WHITE	HISPANIC OR LATINO	Plac	Placebo
3	CDISCPILOT01	01-701-1028	1028	701	71	YEARS	M	WHITE	NOT HISPANIC OR LATINO	Xan_Hi	Xanomeline High Dose
4	CDISCPILOT01	01-701-1033	1033	701	74	YEARS	M	WHITE	NOT HISPANIC OR LATINO	Xan_Lo	Xanomeline Low Dose
5	CDISCPILOT01	01-701-1034	1034	701	77	YEARS	F	WHITE	NOT HISPANIC OR LATINO	Xan_Hi	Xanomeline High Dose
6	CDISCPILOT01	01-701-1047	1047	701	80	YEARS	F	WHITE	NOT HISPANIC OR LATINO	Plac	Placebo
7	CDISCPILOT01	01-701-1057	1057	701	59	YEARS	F	WHITE	HISPANIC OR LATINO	Scritfal	Screen Failure
8	CDISCPILOT01	01-701-1097	1097	701	68	YEARS	M	WHITE	NOT HISPANIC OR LATINO	Xan_Lo	Xanomeline Low Dose
9	CDISCPILOT01	01-701-1111	1111	701	81	YEARS	F	WHITE	NOT HISPANIC OR LATINO	Xan_Lo	Xanomeline Low Dose
10	CDISCPILOT01	01-701-1115	1115	701	84	YEARS	M	WHITE	NOT HISPANIC OR LATINO	Xan_Lo	Xanomeline Low Dose

ADaM

Standards for Analysis
Data

Characteristic	Placebo N = 86	High Dose N = 84
Age		
Mean (SD)	75 (9)	74 (8)
Median (IQR)	76 (69, 82)	76 (71, 80)
Range	52 - 89	56 - 88
Baseline BMI (kg/m²)		
Mean (SD)	23.6 (3.7)	25.3 (4.2)
Median (IQR)	23.4 (21.2, 25.6)	24.8 (22.7, 27.9)
Range	15.1 - 33.3	13.7 - 34.5
Discontinuation, n (%)		
Completed	58 (67.4%)	27 (32.1%)
Adverse Event	8 (9.3%)	40 (47.6%)
Withdrew Consent	9 (10.5%)	8 (9.5%)
Sponsor Decision	2 (2.3%)	3 (3.6%)
Lack of Efficacy	3 (3.5%)	1 (1.2%)
I/E Not Met	1 (1.2%)	2 (2.4%)
Physician Decision	1 (1.2%)	2 (2.4%)
Death	2 (2.3%)	0 (0.0%)
Protocol Violation	1 (1.2%)	1 (1.2%)
Lost to Follow-up	1 (1.2%)	0 (0.0%)

No Standards for Results??

CDISC is developing a
standard called **Analysis
Results Standards (ARS)**



Analysis Results Standards (ARS)

The CDISC Analysis Results Standard model refers to the framework or structure used to store and organize the information related to an analysis result. This includes both the results data (ARD) generated by the analysis and the metadata (ARM-TS) that describes the characteristics of that results data. ARS is based on LinkML logical model which is an open-source schema development language and framework for generating machine-readable models.

Analysis Results Data
(ARD)

Standardized format for producing statistical analysis results in a consistent manner (e.g., treatment, variable, statistical method)

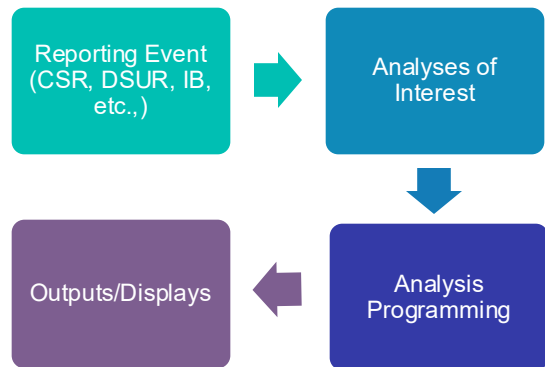
Analysis Results Metadata -
Technical Specifications
(ARM-TS)

Specifies the metadata needed to generate statistical analysis results and associated displays (e.g., title, population set, where-clause)

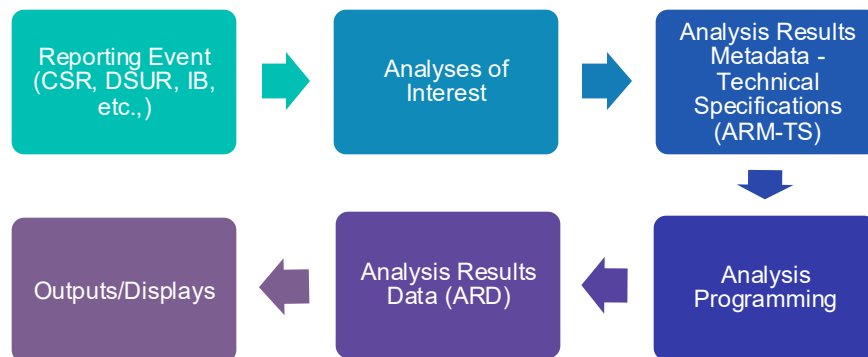


Analysis Results Workflow

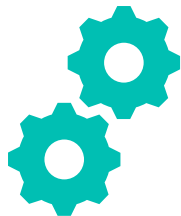
Present



Future



Principles and Benefits



Standardized, Machine-Readable Format

ARDs provide a consistent structure for results across sponsors. This makes it easier for regulators and teams to load, search, and validate results data. Automating TLF production from ARDs reduces manual effort.

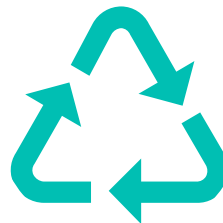
Automation



Faster Review & Reproducibility

The reviewers can easily reproduce the results with ARDs along with the metadata, without re-running the code, which makes the review faster.

Reproducibility



Reusability

A single ARD can produce multiple outputs (e.g., summary table, figure) by filtering and formatting its data.

Reusability



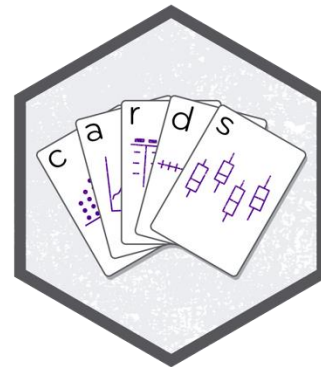
Efficiency

Enables traceability of the results, back to the ARD, and from there to ADaM and further back to SDTM and beyond through the use of existing CDISC standards and metadata.

Traceability

Introduction to {cards} package

- The {cards} package (**CDISC Analysis Result Data Sets**) adopts the CDISC ARS framework and creates ARD objects from data
- It computes common summary statistics (continuous, categorical, etc.) and returns a structured ARD “data.frame” that follows the ARS model.
- The {cardx} package is an extension to {cards} that supports specialized analyses (e.g., t-test, ANOVA, regression models, etc.,)
- The {cards} ARDs integrate seamlessly with packages such as {tfrm} and {gtsummary} for the formatted outputs creation.
- The ARDs can also be exported to XPT and JSON objects which makes it future ready.





Vital Signs Example

Packages Used - {pharmaverseadam}, {cards}, {tfrmt}, {gt}, {dplyr}, {stringr}.

Datasets Used - ADSL and ADVS datasets from the {pharmaverseadam} package.

```
advs <- pharmaverseadam::advs
adsl <- pharmaverseadam::adsl
```

Data Preparation - Filtering to include only the safety population and the systolic blood pressure records

```
# Filtering ADSL for Safety population
adsl_filter <- adsl |>
  filter(SAFFL == "Y")

# Filtering the dataset to include only Safety population and Analysis Flag 01
advs_filter <- advs |>
  filter(!is.na(SAFFL) & SAFFL=="Y" &
    ((!is.na(ANL01FL) & ANL01FL=="Y") | (!is.na(ABLFL) & ABLFL=="Y")) &
    PARAMCD %in% c("SYSBP") &
    is.na(DTYPE) &
    !is.na(AVISITN)) |>
  select(STUDYID, USUBJID, TRT01A, PARAM, PARAMCD, ADT, VSPOS, VISIT, VISITNUM, AVISIT, AVISITN, AVAL, BASE, DTYPE, ABLFL, CHG)

# Filtering only the Supine records for the Analysis
advs_pos <- advs_filter |>
  filter(str_trim(str_to_upper(VSPOS))=="SUPINE")
```



Vital Signs Example

Summary Statistics Calculation - Calculating the descriptive summary statistics using the `cards::ard_continuous` function

```
# Calculating the descriptive summary statistics using the ard_continuous function
adv_s_ard <- ard_summary(  
  data = advs_pos,  
  by = c(TRT01A, AVISITN),  
  variables = AVAL  
)
```

Here's how the ARD looks

	group1	group1_level	group2	group2_level	variable	context	stat_name	stat_label	stat	fmt_fun	warning	error
	/	All	/	All	AI	A	All	All	All	A	AI	
1	TRT01A	Placebo	AVISITN	0	AVAL	summary	N	N	85	0		
2	TRT01A	Placebo	AVISITN	0	AVAL	summary	mean	Mean	138.635294117647	1		
3	TRT01A	Placebo	AVISITN	0	AVAL	summary	sd	SD	16.7537110734874	1		
4	TRT01A	Placebo	AVISITN	0	AVAL	summary	median	Median	140	1		
5	TRT01A	Placebo	AVISITN	0	AVAL	summary	p25	Q1	129	1		
6	TRT01A	Placebo	AVISITN	0	AVAL	summary	p75	Q3	150	1		
7	TRT01A	Placebo	AVISITN	0	AVAL	summary	min	Min	90	1		
8	TRT01A	Placebo	AVISITN	0	AVAL	summary	max	Max	180	1		



Vital Signs Example

By default, the basic descriptive statistics got returned. If needed, we can also specify or pass any functions.

Here's how the refined ARD looks

```
ard_select <- advs_ard |>  
  rename_ard_columns() |>  
  unlist_ard_columns() |>  
  select(-c(fmt_fun, warning, error))
```

	TRT01A	AVISITN	AVAL	context	stat_name	stat_label	stat
	A	A		A	All	All	All
1	Placebo	0	AVAL	summary	N	N	85
2	Placebo	0	AVAL	summary	mean	Mean	138.635294117647
3	Placebo	0	AVAL	summary	sd	SD	16.7537110734874
4	Placebo	0	AVAL	summary	median	Median	140
5	Placebo	0	AVAL	summary	p25	Q1	129
6	Placebo	0	AVAL	summary	p75	Q3	150
7	Placebo	0	AVAL	summary	min	Min	90
8	Placebo	0	AVAL	summary	max	Max	180
9	Placebo	2	AVAL	summary	N	N	83
10	Placebo	2	AVAL	summary	mean	Mean	135



Vital Signs Example

Now, we have created the ARD needed for the analysis.

After creating the formatting using the {tfrmt} and {gt} packages, here's the final output of the summary table

Summary of Systolic Blood Pressure - Safety Population				
Analysis Visit	Statistic	Placebo (N=86)	Xanomeline High Dose (N=72)	Xanomeline Low Dose (N=96)
Baseline	n	85	72	96
	Mean	138.6	139.0	139.8
	SD	16.75	16.71	17.57
	Median	140.00	140.00	138.50
	Q1	129.00	130.00	128.50
	Q3	150.00	150.00	151.00
	Min	90.00	100.00	100.00
	Max	180.00	180.00	188.00
Week 2	n	83	72	75
	Mean	135.0	134.0	137.3
	SD	16.09	14.88	17.50
	Median	132.00	133.00	137.00
	Q1	122.00	122.00	122.00
	Q3	148.00	142.00	148.00
	Min	96.00	106.00	108.00
	Max	172.00	183.00	190.00



Vital Signs Example

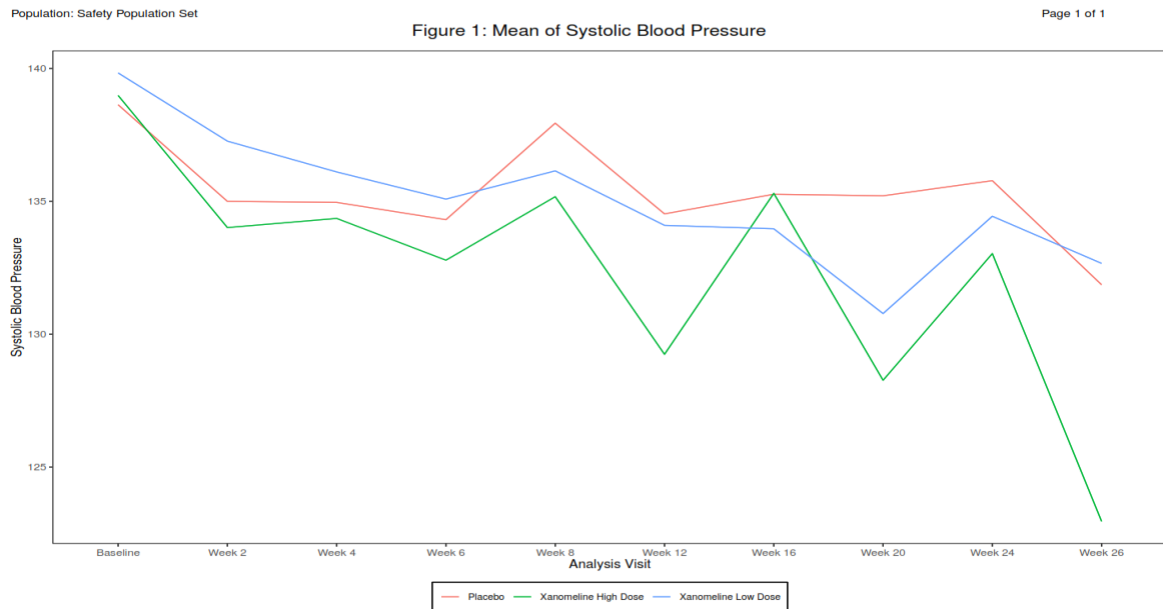
Now, let's create a mean plot of systolic blood pressure using the same ARD dataset with {ggplot2} package.

```
plot <- and_final |>
  filter(stat_name=="mean") |>
  arrange(AVISITN) |>
  mutate(AVISIT = fct_inorder(AVISIT)) |>
  ggplot(
    aes(x = AVISIT, y = stat, group = TRT01A, color = TRT01A)) +
  geom_line() +
  labs(
    x = "Analysis Visit",
    y = "Systolic Blood Pressure") +
  theme_elements
```



Vital Signs Example

Here's the final output of the mean plot figure.



The above figure was produced using ADVS dataset from the {pharmaverseadam} package

15-Aug-2025 06:51



Limitations of ARS

- ARS is a new standard (ARS v1.0 published April 2024) and is still being in the development stage.
- ARS uses LinkML logical model which is complex to understand and difficult to implement.
- CDISC is actively working to simplify the model to enhance wider adaptability.
- It will evolve based on the industry feedback.
- As it's still in the development stage, no standard template available for the Analysis Results Metadata - Technical Specifications (ARM - TS) for usage.

References

- CDISC. (2024). [*Analysis Results Standard User Guide v1.0*](#). CDISC Wiki
- Wallendszus, K., & Frenzel, H. (2022). [*CDISC Analysis Results Standard – Update and Progress*](#). PHUSE EU Connect Conference Proceedings, Paper DS09
- LeRoy, B., Marshall, R., & Busa, B. (2024). [*Getting Started with the New CDISC Analysis Results Standard*](#). PharmaSUG Conference Proceedings, Paper DS04
- Busa, B., LeRoy, B., & Marshall, R. (2023). [*All You Need to Know about the New CDISC Analysis Results Standards!*](#) PharmaSUG 2023 Conference Proceedings, Paper MM-327
- Sjoberg, D. D., & Steed, N. (2025). [*Harnessing Analysis Results Datasets \(ARD\) for Clinical Reporting in R: Our First ARD-Based Filing Experience*](#). PHUSE US Connect 2025 Conference Slides.

Thank You!!

Any Questions



**The Global Healthcare
Data Science Community**

Contact Channels

📞 UK +44 1843 609600
✉ office@phuse.global
🌐 phuse.global

Social Media

🐦 [@phusetwitta](https://twitter.com/phusetwitta)
📘 [/phusebook](https://facebook.com/phusebook)
▶ [/phusetube](https://youtube.com/phusetube)
🌐 [/company/phuse](https://linkedin.com/company/phuse)

Backup Slides



{cards}: Other Summary functions

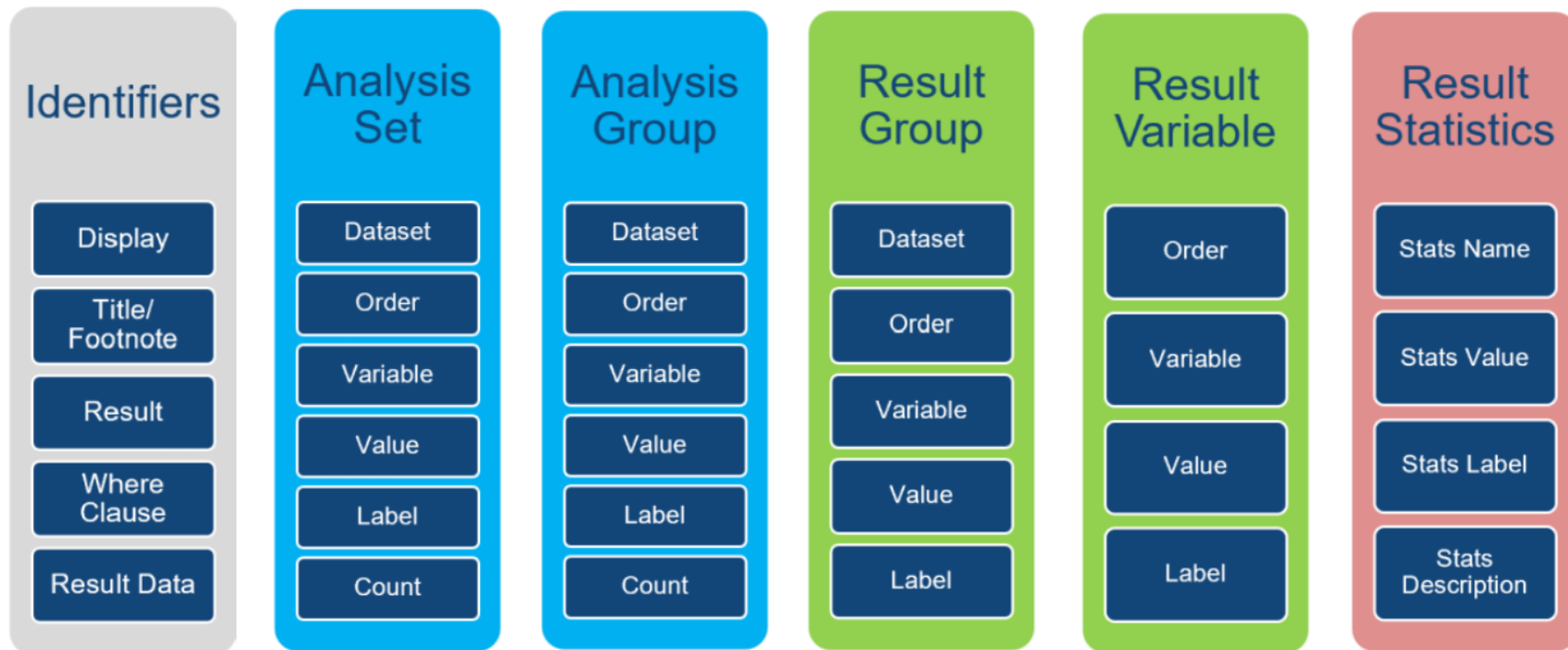
- `ard_categorical()` renamed to `ard_tabulate()*`: useful for calculating the n counts for the categorical variables
- `ard_hierarchical()`: useful for nested tabulations, e.g. AE terms within SOC
- `ard_dichotomous()` renamed to `ard_tabulate_value()*`: similar to `ard_categorical()`, and tabulates a single value of the variable

*Some functions are renamed in the newer version 0.7.0 which published on 27AUG2025.





Key Elements of the ARD



Key Elements of the ARM-TS

Output	StudyID
	Task
	Group
	Order
	DisplayID
	Filename
	Version
	FileType
	StyleID

Display	DisplayID
	ReferenceDisplayID
	Version
	Name
	Title
	DisplaySectionID
	AnalysisSetID
	AnalysisGroupID
	DisplayTemplateID
	Document

DisplaySection	DisplaySectionID
	ReferenceDisplayID
	Section
	SectionSubID
	Order
	Text
	Reference

AnalysisSet	AnalysisSetID
	Dataset
	Order
	Variable
	Comparator
	Value
	Label
	CompoundExpression

AnalysisGroup	AnalysisGroupID
	Dataset
	Order
	Variable
	Comparator
	Value
	Label
	CompoundExpression

Template	DisplayTemplateID
	TemplateContext
	ElementName
	ElementValue

Result	DisplayID
	AnalysisResultID
	Version
	ResultDescription
	DisplayPattern
	Reason
	Purpose
	Dataset
	AnalysisVariable
	AnalysisGroupID
	WhereClauseID
	GroupingByVar
	GroupingByOrdFmt
	Documentation
	ProgrammingCodeContext
	ProgrammingCode
	CodeReference

Style	StyleID
	StyleContext
	ElementName
	ElementValue

WhereClause	WhereClauseID
	Dataset
	Order
	Variable
	Comparator
	Value
	Label
	Compound Expression