

Welcome!

Cardinal

A Collaborative Leap Towards Harmonization of Clinical Reporting Standards

Abi Yogasekaram, Roche
Emily de la Rua, Roche

November 3, 2025



Instructors

Abinaya Yogasekaram



linkedin.com/in/ayogasek



github.com/ayogasekaram



Emily de la Rúa



linkedin.com/in/emilydelarua



github.com/edelarua



Workshop Outline



1. What is Cardinal?
 1. History & Motivation
 2. Our Journey
 3. Navigating the Cardinal website
2. Learnings, Outlook, & Call for Collaboration
3. Technical Overview
 1. {gtsummary}
 2. {crane}
 3. ARDs
4. Workshop Exercises

Cardinal

Formerly {falcon}



Workshop Scope



Understand and navigate the Cardinal template catalog



Brief technical overview of the `{gtsummary}` and `{crane}` R packages



Exercises to create TLGs

Workshop Expectations

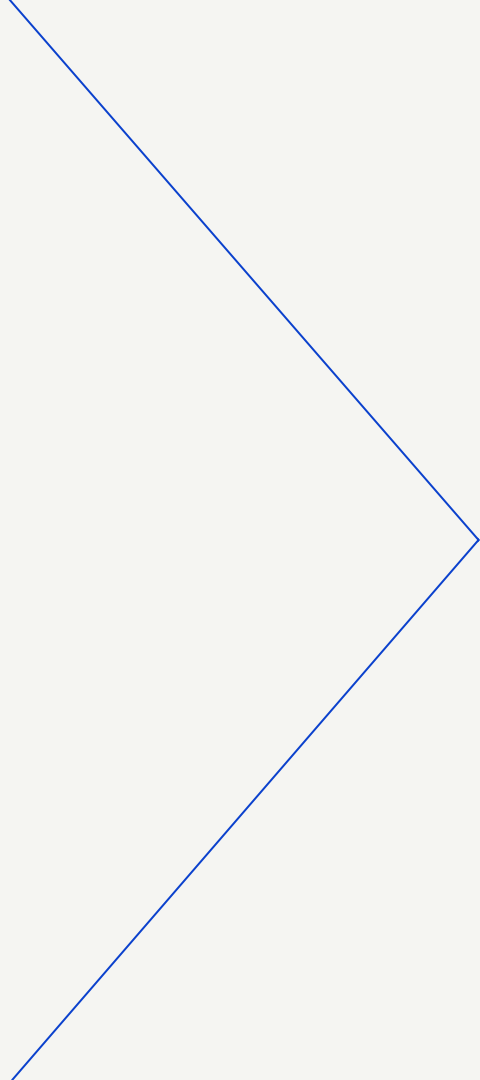


Please ask questions at any time!



Turn your cameras on if you're comfortable! We'd love to see you 😊

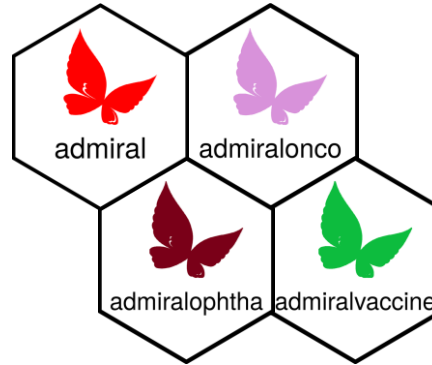
History & Motivation



Pharma Industry Has Very Well Established Data Standards

SDTM & ADaM have brought great benefits to clinical trial conduct & analyses

*Universally agreed upon standards not only enable **easier data sharing & re-use**, but also foster **industry collaboration***



cdisc

How About TLGs?

We all create demographic tables, yet in a thousand different ways

Table XX Baseline Demographics and Characteristics Safety Set			
Age (Years)			
n	XX (XX.X)	XX	
Mean (SD)	XX.X		
Median	XX.X		
Min, Max	XX.X		
Age Group, n (%)			
<<Age breakdown 1 per protocol>>	XX (XX.X)		
<<Age breakdown 1 per protocol>>	XX (XX.X)		
.....	XX (XX.X)		
Sex, n (%)			
Male	XX (XX.X)		
Female	XX (XX.X)		
Race, n (%)			
White	XX (XX.X)		
Black or African American	XX (XX.X)		
Asian	XX (XX.X)		
American Indian or Alaska Native	XX (XX.X)		
Native Hawaiian or Other Pacific Islander	XX (XX.X)		
Other	XX (XX.X)		
Unknown / Not Reported	XX (XX.X)		
Ethnicity, n (%)			
Hispanic or Latino	XX (XX.X)		
Not Hispanic or Latino	XX (XX.X)		
Unknown / Not Reported	XX (XX.X)		
Weight (kg)			
n	XX.X (XX.X)	XX.X, XX.X	
Mean (SD)	XX.X		
Median	XX.X, XX.X		
Min, Max	XX.X, XX.X		

	A: Drug X (N=134)	B: Placebo (N=134)	C: Comb (N=134)
Age (yr)			
n	134	134	132
Mean (SD)	33.8 (6.6)	35.4 (7.9)	35.4 (7.9)
Median	33.0	35.0	35.0
Min - Max	21.0 - 50.0	21.0 - 62.0	20.0 - 69.0
Age Group			
n	134	134	132
18-40	113 (84.3%)	103 (76.9%)	106 (80.3%)
41-64	21 (15.7%)	31 (23.1%)	26 (19.7%)
>=65	0	0	0
Sex			
n	134	134	132
Female	79 (59%)	82 (61.2%)	76 (57.6%)
Male	55 (41%)	52 (38.8%)	56 (42.4%)
Ethnicity			
n	134	134	132
NOT REPORTED	6 (4.5%)	10 (7.5%)	11 (8.3%)
HISPANIC OR LATINO	15 (11.2%)	18 (13.4%)	15 (11.4%)
NOT HISPANIC OR LATINO	104 (77.6%)	103 (76.9%)	101 (76.3%)
UNKNOWN	9 (6.7%)	3 (2.2%)	5 (3.8%)
Race			
n	134	134	132
ASIAN	68 (50.7%)	67 (50%)	73 (55.3%)
BLACK OR AFRICAN AMERICAN	31 (23.1%)	28 (20.9%)	32 (24.3%)
WHITE	27 (20.1%)	26 (19.4%)	21 (15.9%)
AMERICAN INDIAN OR ALASKA NATIVE	8 (6%)	11 (8.2%)	6 (4.6%)
MULTIPLE	0	1 (0.7%)	0
NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER	0	1 (0.7%)	0
OTHER	0	0	0
UNKNOWN	0	0	0
Continuous Level Biomarker 1			
n	134	134	132
Mean (SD)	6.0 (3.6)	5.7 (3.3)	5.6 (3.2)
Median	5.4	4.8	4.8
Min - Max	0.4 - 17.7	0.6 - 14.2	0.2 - 14.2

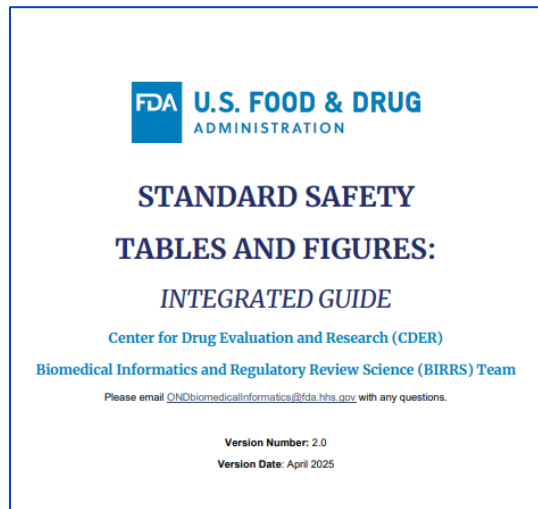
2.1.4.2.1 Example: Demographic data				
	Drug A	Drug B	Total	
Number of subjects (N, %)	142	100.0 219	100.0 361	100.0
Sex (N, %)	142	100.0 219	100.0 361	100.0
Male	111	78.2 173	79.0 284	78.7
Female	31	21.8 46	21.0 77	21.3
Race (N, %)	142	100.0 219	100.0 361	100.0
White	0	0.0 1	0.5 1	0.3
Black or African American	8	5.6 11	5.0 19	5.3
Asian	8	77.5 169	77.2 279	77.3
American Indian or Alaska Native	24	16.9 38	17.4 62	17.2
Native Hawaiian or Other Pacific Islander	142	100.0 219	100.0 361	100.0
Other	110	77.5 169	77.2 279	77.3
Unknown / Not Reported	10	5.6 12	5.5 20	5.5
Ethnicity (N, %)	142	100.0 219	100.0 361	100.0
Hispanic or Latino	110	77.5 169	77.2 279	77.3
Not Hispanic or Latino	32	16.9 38	17.4 62	17.2
Unknown / Not Reported	0	0.0 1	0.5 1	0.3

Demographic data, data of baseline and medication details				
Demographics and patient characteristics at baseline - Randomized population				
	A: Drug X (N=133)	B: Placebo (N=141)	C: Combination (N=126)	All (N=400)
Age (years)				
Number	133	141	126	400
Mean (SD)	35.4 (7.5)	34.9 (7.4)	34.3 (7.4)	34.9 (7.4)
Median	36.0	34.0	33.0	34.0
Q1; Q3	29.0; 40.0	30.0; 39.0	29.0; 38.0	29.0; 39.0
Min; Max	21; 58	20; 62	23; 69	20; 69
Age group (n (%))				
Number	133	141	126	400
From 18 - 64 years	133 (100)	141 (100)	126 (99.2)	399 (99.8)
From 65 - 84 years	0	0	1 (0.8)	1 (0.2)
Sex (n (%))				
Number	133	141	126	400
Male	56 (42.1)	66 (46.8)	47 (37.3)	169 (42.2)
Female	77 (57.9)	75 (53.2)	79 (62.7)	231 (57.8)
Race (n (%))				
Number	133	141	126	400
White	70 (52.6)	86 (61.0)	62 (49.2)	218 (54.5)
Black or African American	28 (21.1)	28 (19.9)	24 (19.0)	80 (20.0)
Asian	26 (19.5)	22 (15.6)	32 (25.4)	80 (20.0)
American Indian or Alaska Native	7 (5.3)	5 (3.5)	8 (6.3)	20 (5.0)
Native Hawaiian or Other Pacific Islander	1 (0.8)	0	0	1 (0.2)
Multiple	1 (0.8)	0	0	1 (0.2)
Unknown / Not Reported	0	0	0	0

BMI: Body mass index				
PGM-DEV/COMPOUND NAME-STUDY NAME-ANALYSIS NAME-REPORT/PGM-dm demo r 1 R-OUT-REPORT/OUTPUT/dm demo r 1 x.rtf (UNOCT2018.53)				

An Opportunity Arose

FDA proposed an integrated guide for standard safety tables & figures



Boehringer
Ingelheim

sanofi

moderna

Common Toolkit:

Open-source R packages for TLG creation are available



Shared Resource:

Developers come from different companies



One Layout:

A much easier entry point for collaboration



Instead of potentially implementing this guide individually, why don't we do it together?

Pivot from {falcon} to Cardinal

Early 2024

- CDISC publishes Analysis Results Datasets (ARDs)
 - Structured way to store analytic results
- Limited benefit from accommodating 3 different table engines

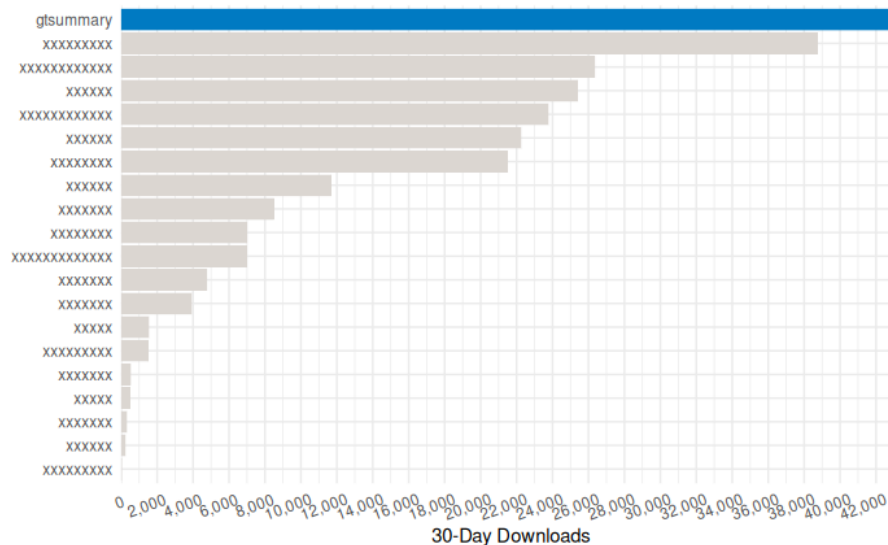


*Rather than developing {falcon} to accommodate different table engines – use a **single package***

{gtsummary}



- Recently refactored to have an ARD backbone
 - Widely adopted = more resources available
 - Channel of communication with the author
 - Readable code
 - Complex tables were easier to create using simpler tables
-
- The stats
 - 1,600,000 installations from CRAN
 - 1,100 GitHub stars
 - 1,000 citations in peer-reviewed articles
 - 350 contributors
 - 50 code contributors
 - Won the 2021 American Statistical Association (ASA) Innovation in Programming Award
 - Won the 2024 Posit Pharma Table Contest



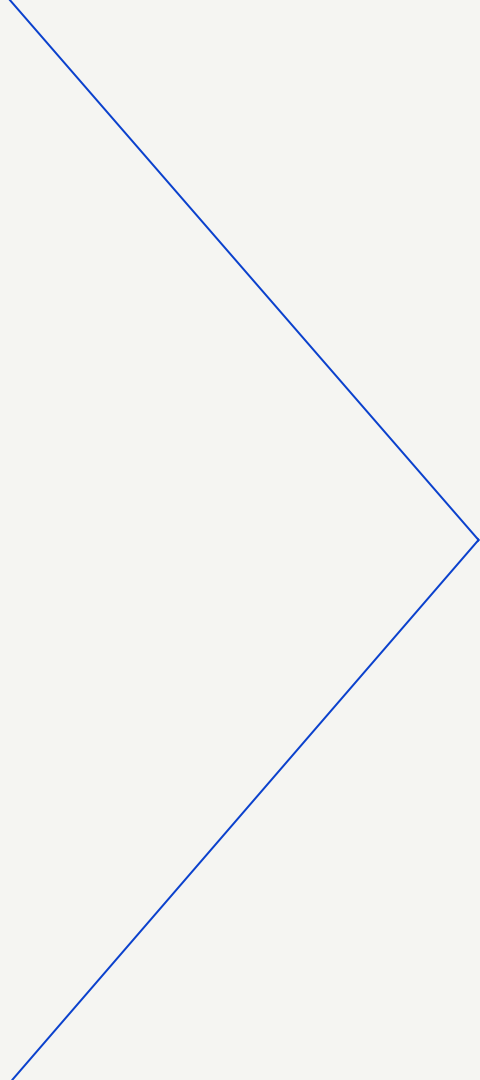
Cardinal

First industry collaborative effort for TLG creation



An industry collaborative effort with the aspiration of open-sourcing a catalog of harmonized TLGs for clinical study reporting and simplifying the process of output review, re-use, and meta-analyses

Current Progress



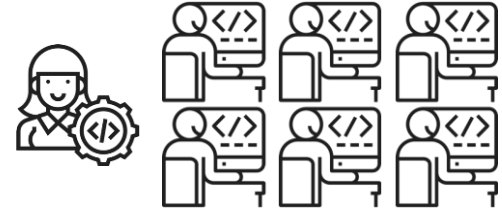
Project Coordination

How does a cross-company team work?



Product Owners

- Template prioritization
- Refine requirements
- Project roadmap



Developers

- Agile development
- Bi-weekly standup meeting
- GitHub project board to track progress

Our Journey

What have we achieved so far?



Kickoff

Companies agreed to co-develop TLG templates, using FDA's guide as a starting point

2022 Q4



Team Setup

Github repository & slack channel was created. Product owners & developers were onboarded to the project

2023 Q1



Active Development

A website was launched to share development progress, template code, upcoming events, etc. ~27 templates are now publicly available

Q2-Q3



Increasing Awareness

{falcon} workshop was conducted in this year's R/Pharma conference. CDISC and more pharma companies have reached out to inquire about future collaborations

Q4



Catalog Pivot

Instead of reinventing the wheel, we pivoted to developing a catalog of templates (FDA+) using the {gtsummary} package

2024



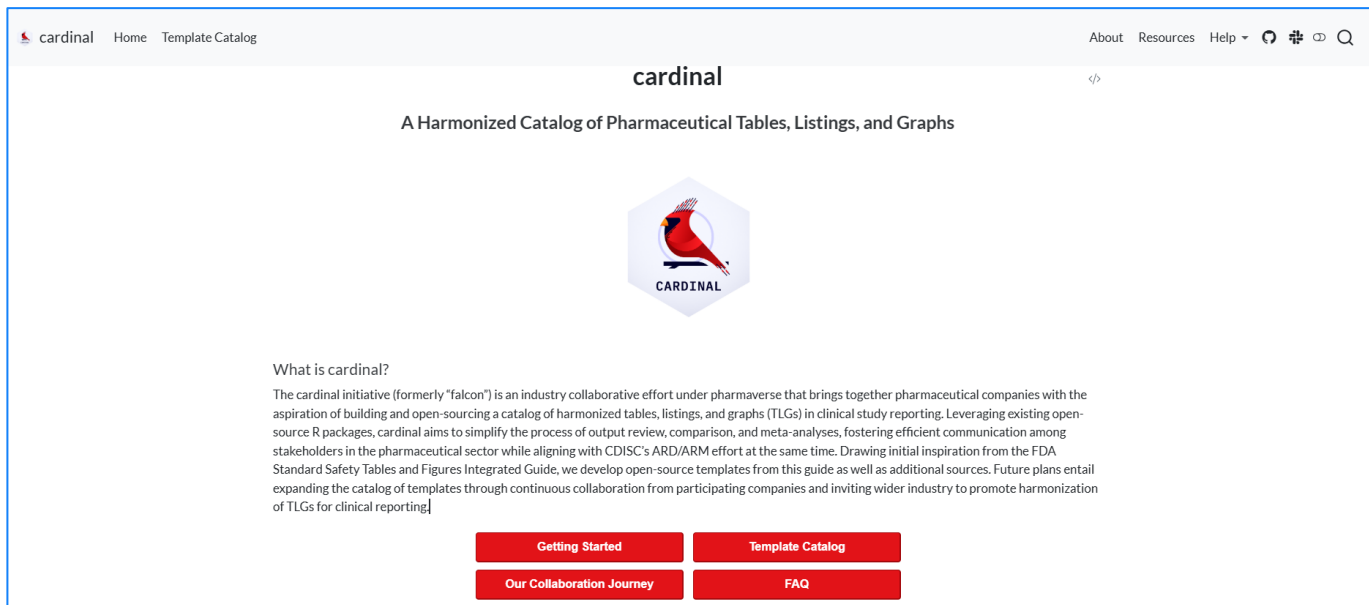
Future Plans

- Alignment with CDISC's ARD/ARM efforts
- Expanding scope of work to common clinical trial analyses, and further promoting a harmonized TLG standard

2025+

A Deeper Look

Explore Cardinal in detail



<https://pharmaverse.github.io/cardinal/>

Template Catalog

Template Catalog			Categories
1F	Order By		All (27)
			FDA (26)
			Roche (1)
			adverse events (18)
			change from baseline (1)
			deaths (1)
			demographics (1)
			disposition (2)
			exposure (1)
			parallel-group (1)
			safety (26)
			table (27)
			vital signs (3)
TLG Description	Source	Categories	
All Individual Subject Deaths, Safety Population, Pooled Analysis (or Trial X)	FDA Table 09	table, FDA, safety, adverse events	
Deaths, Safety Population, Pooled Analysis (or Trial X)	FDA Table 08	table, FDA, safety, deaths	
Demographics and Baseline Clinical Characteristics, Safety Population, Pooled Analysis (or Trial X)	FDA Table 02	table, FDA, safety, demographics	
Duration of Treatment Exposure, Safety Population, Pooled Analysis (or Trial X)	FDA Table 06	table, FDA, safety, exposure	
Laboratory Test Results and Change from Baseline by Visit	Roche LBT01	table, Roche, parallel-group, change from baseline	
Overview of Adverse Events by Demographic Subgroup, Safety Population, Pooled Analysis (or Trial X)	FDA Table 51	table, FDA, safety, adverse events	
Overview of Adverse Events, Safety Population, Pooled Analysis (or Trial X)	FDA Table 07	table, FDA, safety, adverse events	
Overview of Serious Adverse Events by Demographic Subgroup, Safety Population, Pooled Analysis (or Trial X)	FDA Table 50	table, FDA, safety, adverse events	

<https://pharmaverse.github.io/cardinal/>

Data Setup

Table Preview Setup Build Table Build ARD

▼ Code

```
1 # Load libraries & data -----
2 library(dplyr)
3 library(gtsummary)
4
5 adsl <- random.cdisc.data::cads1 |>
6 mutate(
7   AGEGR1 = factor(
8     case_when(
9       AGE >= 17 & AGE < 65 ~ ">=17 to <65",
10      AGE >= 65 & AGE < 75 ~ "65-74",
11      AGE >= 75 ~ ">=75"
12    ),
13    levels = c(">=17 to <65", "65-74", ">=75")
14  )
15 )
16
17 # Pre-processing -----
18 # Filter for the safety population, x
19 data <- adsl |>
20 filter(SAFFL == "Y")
```

<https://pharmaverse.github.io/cardinal/>

Build Table

Table Preview Setup **Build Table** Build ARD

▼ Code

```
1 tbl <- data |>
2   tbl_summary(
3     by = "TRT01A",
4     include = c("SEX", "AGE", "AGEGR1", "ETHNIC", "RACE", "BMRKR1", "BMRKR2"),
5     type = all_continuous() ~ "continuous2", # arranges statistics into multiple lines
6     statistic = list(
7       all_continuous() ~ c(
8         "{mean} ({sd})",
9         "{median} ({min}, {max})"
10      ),
11       all_categorical() ~ "{n} ({p}%"
12     ),
13     label = list(AGEGR1 = "Age Group, Years")
14   ) |>
15   add_overall(last = TRUE, col_label = "***Total Population**  \nN = {N}") |>
16   # remove default footnote
17   remove_footnote_header(columns = everything())
18
19 tbl
```

<https://pharmaverse.github.io/cardinal/>

Extract ARD

Table Preview Setup Build Table Build ARD

▼ Code

```
1 ard <- gather_ard(tbl)
2 ard
```

\$tbl_summary

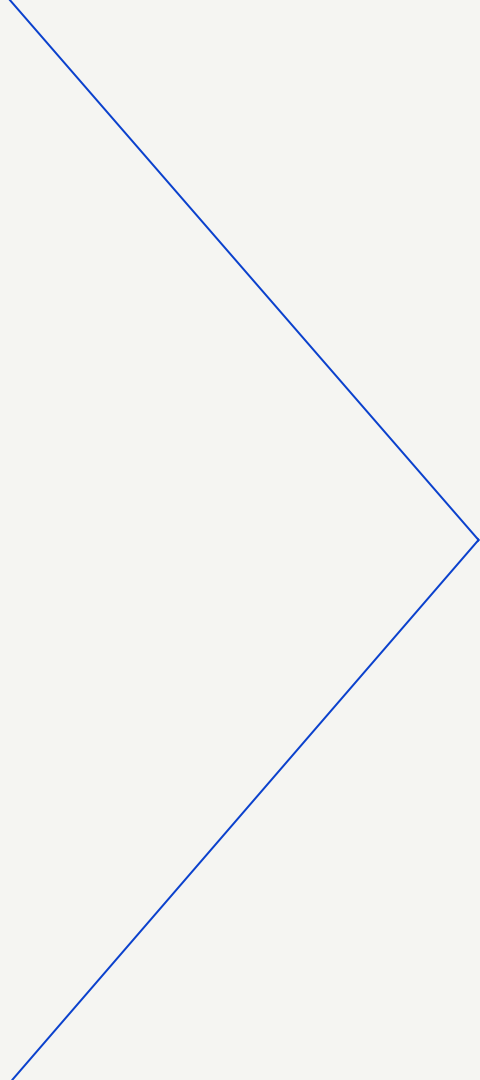
{cards} data frame: 347 x 12

	group1	group1_level	variable	variable_level	context	stat_name	stat_label	stat	fmt_fun	warning	error	gts_column
1	TRT01A	A: Drug X	SEX	F	tabulate	n	n	79	<fn>			stat_1
2	TRT01A	A: Drug X	SEX	F	tabulate	N	N	134	<fn>			stat_1
3	TRT01A	A: Drug X	SEX	F	tabulate	p	%	0.59	<fn>			stat_1
4	TRT01A	A: Drug X	SEX	M	tabulate	n	n	55	<fn>			stat_1
5	TRT01A	A: Drug X	SEX	M	tabulate	N	N	134	<fn>			stat_1
6	TRT01A	A: Drug X	SEX	M	tabulate	p	%	0.41	<fn>			stat_1
7	TRT01A	A: Drug X	RACE	ASIAN	tabulate	n	n	68	<fn>			stat_1
8	TRT01A	A: Drug X	RACE	ASIAN	tabulate	N	N	134	<fn>			stat_1
9	TRT01A	A: Drug X	RACE	ASIAN	tabulate	p	%	0.507	<fn>			stat_1
10	TRT01A	A: Drug X	RACE	BLACK OR...	tabulate	n	n	31	<fn>			stat_1

i 337 more rows

<https://pharmaverse.github.io/cardinal/>

Learnings, Outlook, & Call for Collaboration



Key Learnings

Reflections on our collaboration so far



*Collaboration entry point is significantly lower when an **industry-wide standard** is established*



*Developers are motivated to work on open-source project, which opens **new career opportunities***



*Building open-source solutions together across pharma companies is **less resource intensive** and **more efficient***

Future Outlook

How to fully realize the potential of Cardinal?



Engage more companies and collaborate closely with CDISC & health authorities



An industry harmonized TLG standard for clinical reporting would replace all internal standards, and the implementation is freely accessible for all

Call for Collaboration

The best time to join the journey was a year ago. The second best time is now.



<https://pharmaverse.org/>



<https://bit.ly/48KVL2R>



<https://pharmaverse.github.io/cardinal/>

Acknowledgements

Abinaya Yogasekaram - Roche

Alex Assuied - Sanofi

Daniel Sjoberg - Genentech

Emily de la Rua - Roche

Freeman Wang - Sanofi

Harsha Kalikivayi - Sanofi

Huan Lu - Sanofi

Jaime Pires - Roche

Jessica Knizia - Boehringer-Ingelheim

Juergen Boehl - Boehringer-Ingelheim

Kavitha Allala - Boehringer-Ingelheim

Korbinian Matthias - Boehringer-Ingelheim

Lian Lin - Moderna

Padmaja Chiruvolu - Amgen

Pawel Rucki - Roche

Vincent Shen - Roche

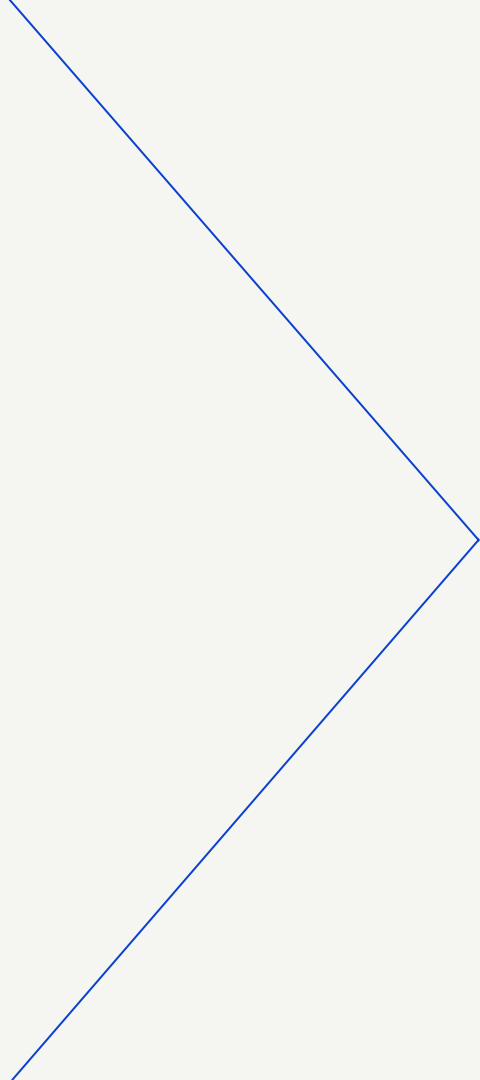
Yichen Wang - Moderna

Yoshito Koujin - Boehringer-Ingelheim

Yuye Wang - Moderna



Technical Overview



tbl_summary()

{pharmaverseadam} Data for Examples

Reduced sizes of `adsl` and `adae`.

```
library(tidyverse)

adsl <- pharmaverseadam::adsl |>
  filter(SAFFL == "Y") |>
  mutate(ARM2 = word(ARM), FEMALE = SEX == "F") |>
  labelled::set_variable_labels(FEMALE = "Female")

adae <- pharmaverseadam::adae |>
  filter(
    USUBJID %in% adsl$USUBJID,
    AESOC %in% c("CARDIAC DISORDERS", "EYE DISORDERS"),
    AEDECOD %in% c("ATRIAL FLUTTER", "MYOCARDIAL INFARCTION", "EYE ALLERGY", "EYE SWELLING")
  ) |>
  mutate(ARM2 = word(ARM))
```

Basic tbl_summary()

```
library(gtsummary)

ads1 |>
  tbl_summary(
    include = c(AGE, ETHNIC, FEMALE)
  )
```

Characteristic	N = 254 ¹
Age	77 (70, 81)
Ethnicity	
HISPANIC OR LATINO	12 (4.7%)
NOT HISPANIC OR LATINO	242 (95%)
Female	143 (56%)
¹ Median (Q1, Q3); n (%)	

- Four types of summaries: `continuous`, `continuous2`, `categorical`, and `dichotomous`
- Statistics are `median (IQR)` for continuous, `n (%)` for categorical/dichotomous
- Variables coded `0/1`, `TRUE/FALSE`, `Yes/No` treated as dichotomous by default
- Label attributes are printed automatically

Customize tbl_summary() output

```
ads1 |>
  tbl_summary(
    include = c(AGE, ETHNIC, FEMALE),
    by = ARM2,
  )
```

Characteristic	Placebo N = 86 ¹	Xanomeline N = 168 ¹
Age	76 (69, 82)	77 (71, 81)
Ethnicity		
HISPANIC OR LATINO	3 (3.5%)	9 (5.4%)
NOT HISPANIC OR LATINO	83 (97%)	159 (95%)
Female	53 (62%)	90 (54%)
¹ Median (Q1, Q3); n (%)		

- **by**: specify a column variable for cross-tabulation



Customize tbl_summary() output

```
ads1 |>
  tbl_summary(
    include = c(AGE, ETHNIC, FEMALE),
    by = ARM2,
    type = AGE ~ "continuous2",
  )
```

Characteristic	Placebo N = 86 ¹	Xanomeline N = 168 ¹
Age		
Median (Q1, Q3)	76 (69, 82)	77 (71, 81)
Ethnicity		
HISPANIC OR LATINO	3 (3.5%)	9 (5.4%)
NOT HISPANIC OR LATINO	83 (97%)	159 (95%)
Female	53 (62%)	90 (54%)
¹ n (%)		

- **by**: specify a column variable for cross-tabulation
- **type**: specify the summary type



Customize tbl_summary() output

```
ads1 |>
  tbl_summary(
    include = c(AGE, ETHNIC, FEMALE),
    by = ARM2,
    type = AGE ~ "continuous2",
    statistic =
      list(
        AGE ~ c("{mean} ({sd})",
                 "{min}, {max}"),
        FEMALE ~ "{n} / {N} ({p}%)"
      ),
  )
```



Characteristic	Placebo N = 86 ¹	Xanomeline N = 168 ¹
Age		
Mean (SD)	75 (9)	75 (8)
Min, Max	52, 89	51, 88
Ethnicity		
HISPANIC OR LATINO	3 (3.5%)	9 (5.4%)
NOT HISPANIC OR LATINO	83 (97%)	159 (95%)
Female	53 / 86 (62%)	90 / 168 (54%)
¹ n (%); n / N (%)		

- **by**: specify a column variable for cross-tabulation
- **type**: specify the summary type
- **statistic**: customize the reported statistics

Customize tbl_summary() output

```
ads1 |>
  tbl_summary(
    include = c(AGE, ETHNIC, FEMALE),
    by = ARM2,
    type = AGE ~ "continuous2",
    statistic =
      list(
        AGE ~ c("{mean} ({sd})",
                  "{min}, {max}"),
        FEMALE ~ "{n} / {N} ({p}%)"
      ),
    label =
      AGE ~ "Age, years",
  )
```



Characteristic	Placebo N = 86 ¹	Xanomeline N = 168 ¹
Age, years		
Mean (SD)	75 (9)	75 (8)
Min, Max	52, 89	51, 88
Ethnicity		
HISPANIC OR LATINO	3 (3.5%)	9 (5.4%)
NOT HISPANIC OR LATINO	83 (97%)	159 (95%)
Female	53 / 86 (62%)	90 / 168 (54%)
¹ n (%); n / N (%)		

- **by**: specify a column variable for cross-tabulation
- **type**: specify the summary type
- **statistic**: customize the reported statistics
- **label**: change or customize variable labels

Customize tbl_summary() output

```
ads1 |>
  tbl_summary(
    include = c(AGE, ETHNIC, FEMALE),
    by = ARM2,
    type = AGE ~ "continuous2",
    statistic =
      list(
        AGE ~ c("{mean} ({sd})",
                 "{min}, {max}"),
        FEMALE ~ "{n} / {N} ({p}%)"
      ),
    label =
      AGE ~ "Age, years",
    digits = AGE ~ list(sd = 1) # report SD(age)
  )
```

Characteristic	Placebo N = 86 ¹	Xanomeline N = 168 ¹
Age, years		
Mean (SD)	75 (8.6)	75 (8.1)
Min, Max	52, 89	51, 88
Ethnicity		
HISPANIC OR LATINO	3 (3.5%)	9 (5.4%)
NOT HISPANIC OR LATINO	83 (97%)	159 (95%)
Female	53 / 86 (62%)	90 / 168 (54%)
¹ n (%); n / N (%)		

- **by**: specify a column variable for cross-tabulation
- **type**: specify the summary type
- **statistic**: customize the reported statistics
- **label**: change or customize variable labels
- **digits**: specify the number of decimal places for rounding

{gtsummary} + formulas

This syntax is also used in {cards}, {cardx}, {crane}, and {gt}.

*select
variables*

*give
instructions*

```
sm_trial %>%  
tbl_summary(  
  label      = age ~ "Patient Age",  
  type       = c(age, marker) ~ "continuous",  
  digits     = starts_with("age") ~ 0,  
  statistic  = all_continuous() ~ "{mean} ({sd})"  
)
```

Use **lists** to pass ≥ 2 sets of instruction:

```
label = list(age ~ "Patient Age", marker ~ "Marker Level")
```

Named list are OK too! `label = list(age = "Patient Age")`

{gtsummary} selectors

- Use the following helpers to **select groups of variables**:
`all_continuous()`, `all_categorical()`
- Use `all_stat_cols()` to select the **summary statistic columns**

Add-on functions in {gtsummary}

`tbl_summary()` objects can also be updated using related functions.

- `add_*()` add **additional column** of statistics or information, e.g. p-values, q-values, overall statistics, treatment differences, N obs., and more
- `modify_*()` **modify** table headers, spanning headers, footnotes, and more

Update tbl_summary() with add_*

```
1 ads1 |>
2   tbl_summary(
3     by = ARM2,
4     include = c(AGE, ETHNIC, FEMALE)
5   ) |>
6   add_overall(last = TRUE)
```

Characteristic	Placebo N = 86 ¹	Xanomeline N = 168 ¹	Overall N = 254 ¹
Age	76 (69, 82)	77 (71, 81)	77 (70, 81)
Ethnicity			
HISPANIC OR LATINO	3 (3.5%)	9 (5.4%)	12 (4.7%)
NOT HISPANIC OR LATINO	83 (97%)	159 (95%)	242 (95%)
Female	53 (62%)	90 (54%)	143 (56%)
¹ Median (Q1, Q3); n (%)			

- `add_overall()`: adds a column of overall statistics

Update tbl_summary() with add_*

- `add_n()`: adds a column non-missing counts
- `add_p()`: adds a column of p-values
- `add_difference()`: mean and rate differences between two groups.
Can also be adjusted differences

Update tbl_summary() with modify_*

```
1 tbl <-  
2   adsl |>  
3   tbl_summary(by = ARM2, include = c("AGE"  
4   modify_header(  
5     stat_1 ~ "**Group A**",  
6     stat_2 ~ "**Group B**"  
7   ) |>  
8   modify_spanning_header(  
9     all_stat_cols() ~ "**Drug**") |>  
10  modify_footnote(  
11    all_stat_cols() ~  
12      paste("median (IQR) for continuous;"  
13            "n (%) for categorical")  
14  )  
15  tbl
```

Characteristic	Drug	
	Group A ¹	Group B ¹
Age	76 (69, 82)	77 (71, 81)
Ethnicity		
HISPANIC OR LATINO	3 (3.5%)	9 (5.4%)
NOT HISPANIC OR LATINO	83 (97%)	159 (95%)
Female	53 (62%)	90 (54%)
¹ median (IQR) for continuous; n (%) for categorical		

- Use `show_header_names()` to see the internal header names available for use in `modify_header()`

Column names

```
show_header_names(tbl)
```



<i>Column Name</i>	<i>Header</i>	<i>Level*</i>	<i>N*</i>	<i>n*</i>	<i>p*</i>
label	***Characteristic**		254 <int>		
stat_1	***Group A**	Placebo <chr>	254 <int>	86 <int>	0.339 <dbl>
stat_2	***Group B**	Xanomeline <chr>	254 <int>	168 <int>	0.661 <dbl>

* These values may be dynamically placed into headers (and other locations).

i Review the ``modify_header()`` help for examples.

`all_stat_cols()` selects columns "stat_1" and "stat_2"

Update tbl_summary() with add_*

```
1 ads1 |>
2   tbl_summary(
3     by = ARM2,
4     include = c(AGE, ETHNIC, FEMALE)
5   ) |>
6   add_stat(...)
```



- Customize statistics presented with `add_stat()`
- Added statistics can be placed on the label or the level rows
- Added statistics may be a single column or multiple

Add-on functions in {gtsummary}

And many more!

See the documentation at

<http://www.danielsjoberg.com/gtsummary/reference/index.html>

And a detailed `tbl_summary()` vignette at

http://www.danielsjoberg.com/gtsummary/articles/tbl_summary.html

tbl_hierarchical()

Adverse Event Reporting (and friends)

Use `tbl_hierarchical()` and `tbl_hierarchical_count()` for reporting of AEs, Con Meds, and more.

```
tbl_ae <- adae |>
  tbl_hierarchical(
    by = "ARM2",
    variables = c("AESOC", "AEDECOD"),
    id = "USUBJID",
    denominator = adsl
  )
```

Primary System Organ Class Dictionary-Derived Term	Placebo N = 86 ¹	Xanomeline N = 168 ¹
CARDIAC DISORDERS	4 (4.7%)	8 (4.8%)
ATRIAL FLUTTER	0 (0%)	2 (1.2%)
MYOCARDIAL INFARCTION	4 (4.7%)	6 (3.6%)
EYE DISORDERS	1 (1.2%)	0 (0%)
EYE ALLERGY	1 (1.2%)	0 (0%)
EYE SWELLING	1 (1.2%)	0 (0%)
¹ n (%)		



tbl_merge()/tbl_stack()

tbl_merge() for side-by-side tables

```
tbl_n <-  
  tbl_summary(adsl, include = ETHNIC, statistic = ETHNIC ~ "{n}") |>  
  modify_header(all_stat_cols() ~ "***N**") |> # update column header  
  remove_footnote_header() # remove footnote  
tbl_age <-  
  tbl_continuous(adsl, include = ETHNIC, variable = AGE, by = ARM2) |>  
  modify_header(all_stat_cols() ~ "**{level}**") # update header  
  
# combine the tables side by side  
list(tbl_n, tbl_age) |>  
  tbl_merge(tab_spanner = FALSE) # suppress default header
```

Characteristic	N	Placebo ¹	Xanomeline ¹
Ethnicity			
HISPANIC OR LATINO	12	64 (63, 86)	63 (56, 78)
NOT HISPANIC OR LATINO	242	76 (70, 82)	77 (71, 81)
¹ Age: Median (Q1, Q3)			



tbl_stack() to combine vertically

```
tbl_drug_a <- filter(adsl, ARM2 == "Placebo") |>
  tbl_summary(include = ETHNIC)
tbl_drug_b <- filter(adsl, ARM2 == "Xanomeline") |>
  tbl_summary(include = ETHNIC)

# stack the two tables
list(tbl_drug_a, tbl_drug_b) |>
  tbl_stack(group_header = c("Placebo", "Xanomeline"), quiet = TRUE) |> # optionally include header
  modify_header(all_stat_cols() ~ "**Summary Statistics**")
```

Characteristic	Summary Statistics ¹
Placebo	
Ethnicity	
HISPANIC OR LATINO	3 (3.5%)
NOT HISPANIC OR LATINO	83 (97%)
Xanomeline	
Ethnicity	
HISPANIC OR LATINO	9 (5.4%)
NOT HISPANIC OR LATINO	159 (95%)
¹ n (%)	



Cobbling Tables Together

- Many tables we create in the pharma space come from a catalog of standard tables.
- The {gtsummary} package makes it simple to break complex tables into their simple parts and cobble them together in the end.
- Moreover, the internal structure of a gtsummary table is **super simple**:
 - A data frame
 - Instructions to print that data frame to make it cute.

Characteristic	Control <i>tbl_merge()</i>		Case		<i>add_stat()</i>		
	Not Exposed	Exposed	Not Exposed	Exposed	Odds Ratio	CMH Odds Ratio	p-value
T Stage	<i>tbl_cross()</i>						
Crude	46	42	52	60	1.26 (0.72, 2.21)	1.23 (0.69, 2.18)	0.6
T1	16	13	12	12	1.23 (0.42, 3.64)		
T2	14	13	11	16	1.57 (0.53, 4.60)		
T3	9	12	13	9	0.52 (0.15, 1.74)		
T4	7	4	16	23	2.52 (0.63, 10.0)		
Grade	<i>tbl_stack()</i>						
Crude	46	42	52	60	1.26 (0.72, 2.21)	1.26 (0.71, 2.22)	0.5
I	19	16	16	17	1.26 (0.49, 3.27)		
II	16	16	16	20	1.25 (0.48, 3.25)		
III	11	10	20	23	1.27 (0.44, 3.60)		

Where are the ARDs?

- ARDs are the backbone for all calculations in gtsummary
- Every gtsummary table saves the ARDs from each calculation

```
tbl <- tbl_summary(ads1, by = "ARM", include = "AGE")  
gather_ard(tbl)
```

```
$tbl_summary  
{cards} data frame: 38 x 12  
  group1 group1_level variable variable_level stat_name stat_label stat  
1     ARM      Placebo     AGE              median      Median    76  
2     ARM      Placebo     AGE              p25           Q1      69  
3     ARM      Placebo     AGE              p75           Q3      82  
4     ARM      Xanomeli... AGE              median      Median    76  
5     ARM      Xanomeli... AGE              p25           Q1     70.5  
6     ARM      Xanomeli... AGE              p75           Q3      80  
7     ARM      Xanomeli... AGE              median      Median   77.5  
8     ARM      Xanomeli... AGE              p25           Q1      71  
9     ARM      Xanomeli... AGE              p75           Q3      82  
10    <NA>           AGE              label      Variable... Age  
i 28 more rows  
i Use `print(n = ...)` to see more rows  
i 5 more variables: context, fmt_fun, warning, error, gts_column
```

ARDs

Tables

A hexagonal icon with a blue background and orange border. It contains a table with the text 'tfrmt' overlaid.

	Treatment	Placebo	Total
Age			
n	XX	XX	XX
Mean	XX (X.XX)	XX (X.XX)	XX (X.XX)
SD	XX	XX	XX
Min	XX	XX	XX
Max	XX	XX	XX
Age Group			
<65 yrs	XX (X.XX)	XX (X.XX)	XX (X.XX)
65-80 yrs	XX (X.XX)	XX (X.XX)	XX (X.XX)
>80 yrs	XX (X.XX)	XX (X.XX)	XX (X.XX)



ARD + QC

ARDs are **wonderful** for QCing {gtsummary} tables. 🐱

- ARDs include the formatted and un-formatted numbers that appear in the table.
- Extract the ARD from the {gtsummary} table.
- Build fresh ARD from source data, and compare it to the ARD from the table.

ARD + QC: Build and Compare ARDs

```
ard_demog <- ads1 |>
  cards::ard_stack(
    cards::ard_summary(
      variables = "AGE",
      statistic =
        AGE ~ cards::continuous_summary_fns(c("median", "p25", "p75"))
    ),
    .by = "ARM2",
  )
```

The next step is to simply compare the two ARDs to confirm results. As this is done programmatically, it is quick to repeat as data continues to accrue.

ARD-first tables

ARD-first Tables

Similar to functions that accept a data frame, the package exports functions with nearly identical APIs that accept an ARD.

```
tbl_summary()
```



```
tbl_hierarchical()
```

```
tbl_continuous()
```

```
tbl_wide_summary()
```

```
tbl_ard_summary()
```



```
tbl_ard_hierarchical()
```

```
tbl_ard_continuous()
```

```
tbl_ard_wide_summary()
```

ARD-first Tables

We can use the skills we learned earlier today to create ARDs for gtsummary tables.

```
library(cards)

ard <- ard_stack(
  data = adsl,
  ard_summary(variables = AGE),
  ard_tabulate(variables = ETHNIC),
  ard_tabulate_value(variables = FEMALE),
  # add these for best-looking tables
  .attributes = TRUE,
  .missing = TRUE
)
ard
```

```
{cards} data frame: 38 x 9
  variable variable_level context stat_name stat_label  stat
1      AGE                summary      N          N      254
2      AGE                summary    mean        Mean  75.087
3      AGE                summary      sd          SD   8.246
4      AGE                summary    median        Median    77
5      AGE                summary    p25          Q1     70
6      AGE                summary    p75          Q3     81
7      AGE                summary    min          Min     51
8      AGE                summary    max          Max     89
9  ETHNIC      HISPANIC... tabulate      n          n     12
10 ETHNIC      HISPANIC... tabulate      N          N    254
i 28 more rows
i Use `print(n = ...)` to see more rows
i 3 more variables: fmt_fun, warning, error
```

The `.attributes=TRUE` call adds column attributes, like labels, to the ARD table, which leads to better defaults by displaying the column



ARD-first Tables

We can simply use the ARD from the previous slide, and pass it to `tbl_ard_summary()` for a summary table.

```
tbl_ard_summary(ard)
```

Characteristic	Overall ¹
Age	77.0 (70.0, 81.0)
Ethnicity	
HISPANIC OR LATINO	12 (4.7%)
NOT HISPANIC OR LATINO	242 (95.3%)
Female	143 (56.3%)
¹ Median (Q1, Q3); n (%)	



{crane}

At Roche, we've created a companion R package to {gtsummary}.

- Functions are thin wrappers for {gtsummary} functions
- Built to tailor defaults to Roche “aesthetics” in tables
- Produces Standard, Hierarchical, Survival Analysis tables and Listings

Notable Functions



- **tbl_baseline_chg()**: Compute Analysis Values and Change from Baseline at each visit
- **tbl_hierarchical_rate_and_count()** : Hierarchical Rates and Counts
- **tbl_hierarchical_rate_by_grade()** : AE Rates by Highest Toxicity Grade
- **tbl_listings()** : Creates a listing from a data frame

The `tbl_roche_summary()` function is a simple wrapper for `tbl_summary()`, but with Roche defaults



Table 1

```
library(crane)

ads1 |>
  dplyr::mutate(ETHNIC = forcats::fct_expand(ETHNIC, "REFUSED")) |>
  tbl_roche_summary(by = ARM, include = c(AGE, ETHNIC), nonmissing = "always")
```

	Placebo (N = 86)	Xanomeline High Dose (N = 84)	Xanomeline Low Dose (N = 84)
Age			
n	86	84	84
Mean (SD)	75 (9)	74 (8)	76 (8)
Median	76	76	78
Min - Max	52 - 89	56 - 88	51 - 88
ETHNIC			
n	86	84	84
HISPANIC OR LATINO	3 (3.5%)	3 (3.6%)	6 (7.1%)
NOT HISPANIC OR LATINO	83 (97%)	81 (96%)	78 (93%)
REFUSED	0	0	0



What else is in {crane}?



Lab values are summarized by visit and include the change from baseline.

This is a simple table that is just a `tbl_merge()` of the `AVAL` summary and the `CHG` summary.

But the general structure appears enough times in our catalog, we make it simple for our programmers to create.

```
pharmaverseadam::adlb |>
  filter(
    SAFFL == "Y", # filter for safety population
    AVISITN < 9000, # filter out invalid visits
    AVISIT %in% c("Baseline", "Week 4", "Week 8"), # keep only visits "Baseline", "Week 4", and
    PARAM %in% "Albumin (g/L)" # keep only 1 parameter
  )
```

What else is in {crane}?



	Placebo (N = 86)		Xanomeline High Dose (N = 84)		Xanomeline Low Dose (N = 84)	
Visit	Value at Visit	Change from Baseline	Value at Visit	Change from Baseline	Value at Visit	Change from Baseline
Baseline						
n	86		84		82	
Mean (SD)	39.84 (2.81)		40.29 (2.84)		39.77 (2.56)	
Median	40.00		40.00		40.00	
Min - Max	32.00 - 46.00		32.00 - 49.00		32.00 - 46.00	
Week 12						
n	67	67	50	50	51	51
Mean (SD)	39.48 (3.49)	-0.34 (2.39)	39.80 (2.45)	-0.70 (2.73)	38.86 (2.18)	-0.88 (2.12)
Median	40.00	0.00	39.50	-1.00	39.00	-1.00
Min - Max	28.00 - 47.00	-6.00 - 6.00	35.00 - 44.00	-7.00 - 6.00	31.00 - 44.00	-7.00 - 4.00
Week 24						
n	57	57	30	30	26	26
Mean (SD)	39.67 (3.34)	-0.04 (2.90)	40.53 (2.10)	-0.57 (2.65)	40.38 (2.52)	0.54 (2.39)
Median	40.00	0.00	41.00	-1.00	40.50	1.00
Min - Max	29.00 - 46.00	-10.00 - 6.00	35.00 - 44.00	-5.00 - 8.00	32.00 - 45.00	-5.00 - 4.00

What else is in {crane}?



`tbl_summary(include=AVAL)`

	Placebo (N = 86)		Xanomeline High Dose (N = 84)		Xanomeline Low Dose (N = 84)	
Visit	Value at Visit	Change from Baseline	Value at Visit	Change from Baseline	Value at Visit	Change from Baseline
Baseline						
n	86		84		82	
Mean (SD)	39.84 (2.81)		40.29 (2.84)		39.77 (2.56)	
Median	40.00		40.00		40.00	
Min - Max	32.00 - 46.00		32.00 - 49.00		32.00 - 46.00	
Week 12						
n	67	67	50	50	51	51
Mean (SD)	39.48 (3.49)	-0.34 (2.39)	39.80 (2.45)	-0.70 (2.73)	38.86 (2.18)	-0.88 (2.12)
Median	40.00	0.00	39.50	-1.00	39.00	-1.00
Min - Max	28.00 - 47.00	-6.00 - 6.00	35.00 - 44.00	-7.00 - 6.00	31.00 - 44.00	-7.00 - 4.00
Week 24						
n	57	57	30	30	26	26
Mean (SD)	39.67 (3.34)	-0.04 (2.90)	40.53 (2.10)	-0.57 (2.65)	40.38 (2.52)	0.54 (2.39)
Median	40.00	0.00	41.00	-1.00	40.50	1.00
Min - Max	29.00 - 46.00	-10.00 - 6.00	35.00 - 44.00	-5.00 - 8.00	32.00 - 45.00	-5.00 - 4.00



What else is in {crane}?



`tbl_summary(include=CHG)`

		Placebo (N = 86)	Xanomeline High Dose (N = 84)		Xanomeline Low Dose (N = 84)	
Visit	Value at Visit	Change from Baseline	Value at Visit	Change from Baseline	Value at Visit	Change from Baseline
Baseline						
n	86		84		82	
Mean (SD)	39.84 (2.81)		40.29 (2.84)		39.77 (2.56)	
Median	40.00		40.00		40.00	
Min - Max	32.00 - 46.00		32.00 - 49.00		32.00 - 46.00	
Week 12						
n	67	67	50	50	51	51
Mean (SD)	39.48 (3.49)	-0.34 (2.39)	39.80 (2.45)	-0.70 (2.73)	38.86 (2.18)	-0.88 (2.12)
Median	40.00	0.00	39.50	-1.00	39.00	-1.00
Min - Max	28.00 - 47.00	-6.00 - 6.00	35.00 - 44.00	-7.00 - 6.00	31.00 - 44.00	-7.00 - 4.00
Week 24						
n	57	57	30	30	26	26
Mean (SD)	39.67 (3.34)	-0.04 (2.90)	40.53 (2.10)	-0.57 (2.65)	40.38 (2.52)	0.54 (2.39)
Median	40.00	0.00	41.00	-1.00	40.50	1.00
Min - Max	29.00 - 46.00	-10.00 - 6.00	35.00 - 44.00	-5.00 - 8.00	32.00 - 45.00	-5.00 - 4.00



What else is in {crane}?



tbl_merge()

		Placebo (N = 86)		Xanomeline High Dose (N = 84)		Xanomeline Low Dose (N = 84)	
Visit		Value at Visit	Change from Baseline	Value at Visit	Change from Baseline	Value at Visit	Change from Baseline
Baseline							
n		86		84		82	
Mean (SD)		39.84 (2.71)		40.29 (2.84)		39.77 (2.56)	
Median		40.00		40.00		40.00	
Min - Max		32.00 - 46.00		32.00 - 49.00		32.00 - 46.00	
Week 12							
n		67	67	50	50	51	51
Mean (SD)		39.48 (3.49)	-0.34 (2.39)	39.80 (2.45)	-0.70 (2.73)	38.86 (2.78)	-0.88 (2.12)
Median		40.00	0.00	39.50	-1.00	39.00	-1.00
Min - Max		28.00 - 47.00	-1.00 - 6.00	35.00 - 44.00	-7.00 - 6.00	32.00 - 44.00	-7.00 - 4.00
Week 24							
n		57	57	30	30	26	26
Mean (SD)		39.67 (3.34)	-0.04 (2.90)	40.53 (2.10)	-0.53 (2.65)	40.38 (2.52)	0.54 (2.39)
Median		40.00	0.00	40.00	-1.00	40.50	1.00
Min - Max		29.00 - 46.00	-10.00 - 6.00	35.00 - 44.00	-5.00 - 8.00	32.00 - 45.00	-5.00 - 4.00



Let's get our hands dirty!

Working environment

- For consistency, we will be working in Posit Cloud
- Everything has been installed and set up for you

Thank you!