

Replicating the CDISC Pilot Submission Project

Using R and PharmaRTF

Agenda

- The CDISC Pilot
- Why R?
- Challenges
- The 'pharmaRTF' package
- Next Steps

CDISC SDTM/ADaM Pilot Project

CDISCPILLOT01

1. TITLE PAGE

Project:	CDISCPILLOT01 – Initial Case Study of the CDISC SDTM/ADaM Pilot Project
Case Study Title:	Safety and Efficacy of the Xanomeline Transdermal Therapeutic System (TTS) in Patients with Mild to Moderate Alzheimer's Disease

The CDISC Pilot

The CDISC Pilot



- Released in 2007
- Full set of SDTM and ADaM data
- Contains 30 Tables presented within an example CSR
- Original outputs generated with SAS
- PHUSE Test Data Factory Project released update of data in 2019

Find our
replication
repository
[here](#)

Goal



Demonstrate R's feasibility as a submission language



Replicate all 30 tables



Generate the tables as RTF outputs

Why R?

- Open Source statistical programming language
- Powerful selection of 3rd party packages
- Advanced visualization capabilities
- Forward-looking dynamic dashboarding capabilities



Challenges



Matching Numbers

R and SAS have different default rounding methods
MMRM Models produced slightly different results



Statistical Methods

Alt. Hypothesis for CMH test only available in the package 'vcdExtra'
(Required a modification to handle sparse matrices)

Excellent paper on the topic [here](#)

Table Package Options

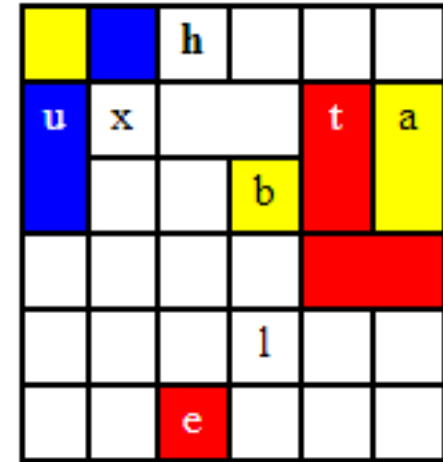
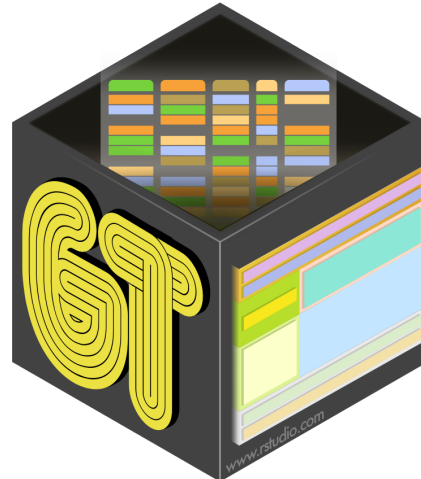


Table Package Options

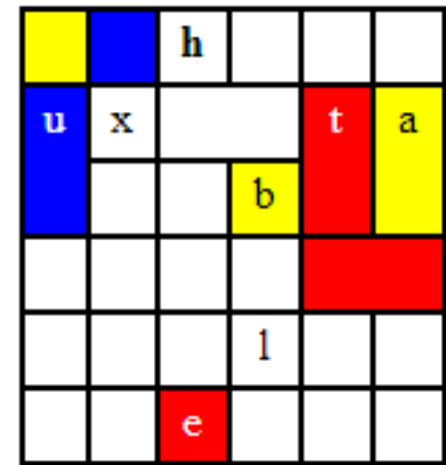
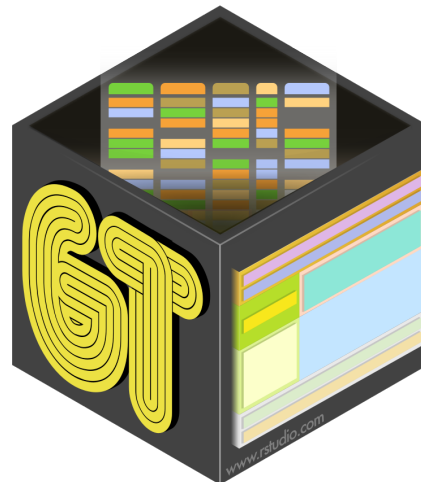
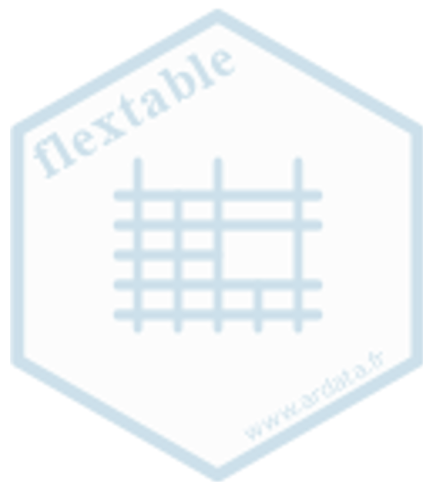
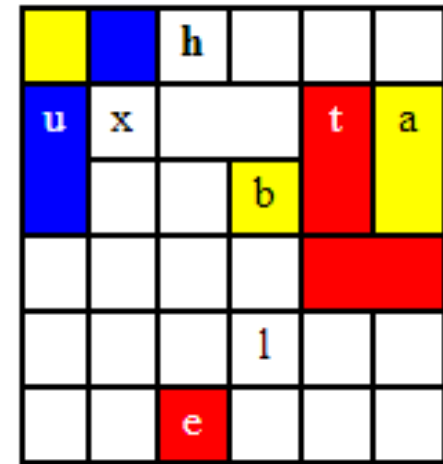
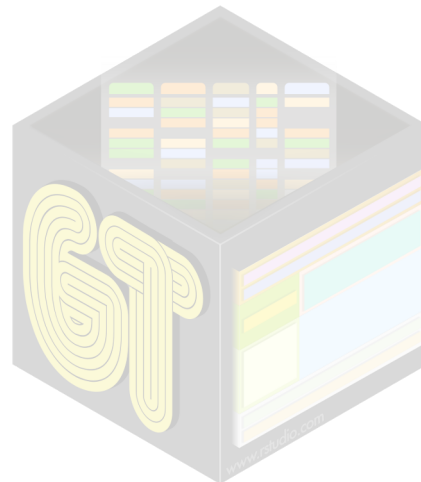
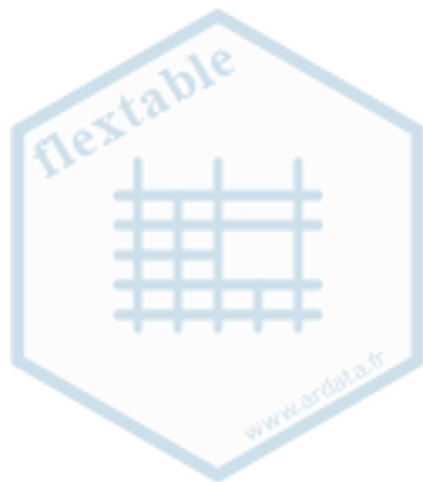


Table Package Options



Huxtable Capabilities



RTF Syntax
support



Easy to work with
abstract of a data-
frame



Supports any
styling of the table
body necessary



Easily flip to other
output types, like
HTML

Issues With RTF



Huxtable writes in-line RTF, no document property support



Table generated won't repeat column headers from page to page



Can't insert headers or footers into a document



RMarkdown's RTF Support doesn't support these features either

pharmaRTF



- Consumes existing table (i.e. Huxtable) and produces output RTF document
- Handles document properties, like page size, orientation, margins, etc.
- Handles headers for titles, footers for footnotes
- Repeats column headers across pages
- Special title/footnote support for page numbering, dates, and file-paths
- MIT License
- Qualification document included

Click me for
the code!

Protocol: CDISCPilot01
Population: Intent-to-Treat

Page 1 of 3

Table 14-2.01
Summary of Demographic and Baseline Characteristics

		Placebo (N=86)	Xanomeline Low Dose (N=84)	Xanomeline High Dose (N=84)	Total (N=254)	p-value [1]
Age (y)	n	86	84	84	254	0.5934
	Mean	75.2	75.7	74.4	75.1	
	SD	8.59	8.29	7.89	8.25	
	Median	76.0	77.5	76.0	77.0	
	Min	52.0	51.0	56.0	51.0	
	Max	89.0	88.0	88.0	89.0	
	<65 yrs	14 (16%)	8 (10%)	11 (13%)	33 (13%)	0.1439
	65-80 yrs	42 (49%)	47 (56%)	55 (65%)	144 (57%)	
	>80 yrs	30 (35%)	29 (35%)	18 (21%)	77 (30%)	
Sex	n	86	84	84	254	0.1409
	Male	33 (38%)	34 (40%)	44 (52%)	111 (44%)	
	Female	53 (62%)	50 (60%)	40 (48%)	143 (56%)	
Race (Origin)	n	86	84	84	254	0.6477
	Caucasian	75 (87%)	72 (86%)	71 (85%)	218 (86%)	
	African Descent	8 (9%)	6 (7%)	9 (11%)	23 (9%)	
	Hispanic	3 (3%)	6 (7%)	3 (4%)	12 (5%)	
	Other	0	0	1 (1%)	1 (<1%)	
MMSE	n	86	84	84	254	0.5947
	Mean	18.0	17.9	18.5	18.1	
	SD	4.27	4.22	4.16	4.21	
	Median	19.5	18.0	20.0	19.0	
	Min	10.0	10.0	10.0	10.0	
	Max	23.0	24.0	24.0	24.0	
Duration of disease	n	86	84	84	254	0.1530
	Mean	42.6	48.7	40.5	43.9	
	SD	30.24	29.58	24.69	28.40	
	Median	35.3	40.2	36.0	36.2	
	Min	7.2	7.8	2.2	2.2	
	Max	183.1	130.8	135.0	183.1	

[1] P-values are results of ANOVA treatment group comparison for continuous variable and Pearson's chisquare test for categorical variables.

NOTE: Duration of disease is computed as months between date of enrollment and date of onset of the first definite symptoms of Alzheimer's disease.

Source: ~/CDISC_pilot_replication/programs/t-14-2-01.R

14:48 Thursday, April 30, 2020

pharmaRTF

Protocol: CDISCPilot01
Population: Intent-to-Treat

Page 1 of 3

Table 14-2.01
Summary of Demographic and Baseline Characteristics

		Placebo (N=86)	Xanomeline Low Dose (N=84)	Xanomeline High Dose (N=84)	Total (N=254)	p-value [1]
Age (yr)	n	86	84	84	254	0.0001
	Mean	75.2	75.7	74.4	75.1	
	SD	8.59	8.29	7.89	8.25	
	Median	76.0	77.5	76.0	77.0	
	Min	52.0	51.0	56.0	51.0	
	Max	89.0	88.0	88.0	89.0	
	<65 yrs	14 (16%)	8 (10%)	11 (13%)	33 (13%)	0.1439
	65-80 yrs	42 (49%)	47 (56%)	55 (65%)	144 (57%)	
	>80 yrs	30 (35%)	29 (35%)	18 (21%)	77 (30%)	
Sex	n	86	84	84	254	0.1409
	Male	33 (38%)	34 (40%)	44 (52%)	111 (44%)	
	Female	53 (62%)	50 (60%)	40 (48%)	143 (56%)	
Race (Origin)	n	86	84	84	254	0.6477
	Caucasian	75 (87%)	72 (86%)	71 (85%)	218 (86%)	
	African Descent	8 (9%)	6 (7%)	9 (11%)	23 (9%)	
	Hispanic	3 (3%)	6 (7%)	3 (4%)	12 (5%)	
	Other	0	0	1 (1%)	1 (<1%)	
MMSE	n	86	84	84	254	0.5947
	Mean	18.0	17.9	18.5	18.1	
	SD	4.27	4.22	4.16	4.21	
	Median	19.5	18.0	20.0	19.0	
	Min	10.0	10.0	10.0	10.0	
	Max	23.0	24.0	24.0	24.0	
Duration of disease	n	86	84	84	254	0.1530
	Mean	42.6	48.7	40.5	43.9	
	SD	30.24	29.58	24.69	28.40	
	Median	35.3	40.2	36.0	36.2	
	Min	7.2	7.8	2.2	2.2	
	Max	182.1	130.8	135.0	182.1	

[1] P-values are results of ANOVA treatment group comparison for continuous variable and Pearson's chisquare test for categorical variables.

NOTE: Duration of disease is computed as months between date of enrollment and date of onset of the first definite symptoms of Alzheimer's disease.

Source: ~/CDISC_pilot_replication/programs/t-14-2-01.R

14:48 Thursday, April 30, 2020



atorus

Your Complete Data Solution

pharmaRTF

Protocol: CDISCPilot01
Population: Intent-to-Treat

Page 1 of 3

Table 14-2.01
Summary of Demographic and Baseline Characteristics

		Placebo (N=86)	Xanomeline Low Dose (N=84)	Xanomeline High Dose (N=84)	Total (N=254)	p-value [1]
Age (yr)	n	86	84	84	254	0.0001
	Mean	75.2	75.7	74.4	75.1	
	SD	8.59	8.29	7.89	8.25	
	Median	76.0	77.5	76.0	77.0	
	Min	52.0	51.0	56.0	51.0	
	Max	89.0	88.0	88.0	89.0	
	<65 yrs	14 (16%)	8 (10%)	11 (13%)	33 (13%)	0.1439
	65-80 yrs	42 (49%)	47 (56%)	55 (65%)	144 (57%)	
	>80 yrs	30 (35%)	29 (35%)	18 (21%)	77 (30%)	
Sex	n	86	84	84	254	0.1409
	Male	33 (38%)	34 (40%)	44 (52%)	111 (44%)	
	Female	53 (62%)	50 (60%)	40 (48%)	143 (56%)	
Race (Origin)	n	86	84	84	254	0.6477
	Caucasian	75 (87%)	72 (86%)	71 (85%)	218 (86%)	
	African Descent	8 (9%)	6 (7%)	9 (11%)	23 (9%)	
	Hispanic	3 (3%)	6 (7%)	3 (4%)	12 (5%)	
	Other	0	0	1 (1%)	1 (<1%)	
MMSE	n	86	84	84	254	0.5947
	Mean	18.0	17.9	18.5	18.1	
	SD	4.27	4.22	4.16	4.21	
	Median	19.5	18.0	20.0	19.0	
	Min	10.0	10.0	10.0	10.0	
	Max	23.0	24.0	24.0	24.0	
Duration of disease	n	86	84	84	254	0.1530
	Mean	42.6	48.7	40.5	43.9	
	SD	30.24	29.58	24.69	28.40	
	Median	35.3	40.2	36.0	36.2	
	Min	7.2	7.8	2.2	2.2	
	Max	183.1	130.8	135.0	183.1	

[1] P-values are results of ANOVA treatment group comparison for continuous variable and Pearson's chisquare test for categorical variables.

NOTE: Duration of disease is computed as months between date of enrollment and date of onset of the first definite symptoms of Alzheimer's disease.

Source: ~/CDISC_pilot_replication/programs/t-14-2-01.R

14:48 Thursday, April 30, 2020

Protocol: CDISCPilot01
Population: Intent-to-Treat

Page 1 of 3

Table 14-2.01
Summary of Demographic and Baseline Characteristics

		Placebo (N=86)	Xanomeline Low Dose (N=84)	Xanomeline High Dose (N=84)	Total (N=254)	p-value [1]
Age (y)	n	86	84	84	254	0.5934
	Mean	75.2	75.7	74.4	75.1	
	SD	8.59	8.29	7.89	8.25	
	Median	76.0	77.5	76.0	77.0	
	Min	52.0	51.0	56.0	51.0	
	Max	89.0	88.0	88.0	89.0	
	<65 yrs	14 (16%)	8 (10%)	11 (13%)	33 (13%)	0.1439
	65-80 yrs	42 (49%)	47 (56%)	55 (65%)	144 (57%)	
	>80 yrs	30 (35%)	29 (35%)	18 (21%)	77 (30%)	
Sex	n	86	84	84	254	0.1409
	Male	33 (38%)	34 (40%)	44 (52%)	111 (44%)	
	Female	53 (62%)	50 (60%)	40 (48%)	143 (56%)	
Race (Origin)	n	86	84	84	254	0.6477
	Caucasian	75 (87%)	72 (86%)	71 (85%)	218 (86%)	
	African Descent	8 (9%)	6 (7%)	9 (11%)	23 (9%)	
	Hispanic	3 (3%)	6 (7%)	3 (4%)	12 (5%)	
	Other	0	0	1 (1%)	1 (<1%)	
MMSE	n	86	84	84	254	0.5947
	Mean	18.0	17.9	18.5	18.1	
	SD	4.27	4.22	4.16	4.21	
	Median	19.5	18.0	20.0	19.0	
	Min	10.0	10.0	10.0	10.0	
	Max	23.0	24.0	24.0	24.0	
Duration of disease	n	86	84	84	254	0.1530
	Mean	42.6	48.7	40.5	43.9	
	SD	30.24	29.58	24.69	28.40	
	Median	35.3	40.2	36.0	36.2	
	Min	7.2	7.8	2.2	2.2	
	Max	183.1	130.8	135.0	183.1	

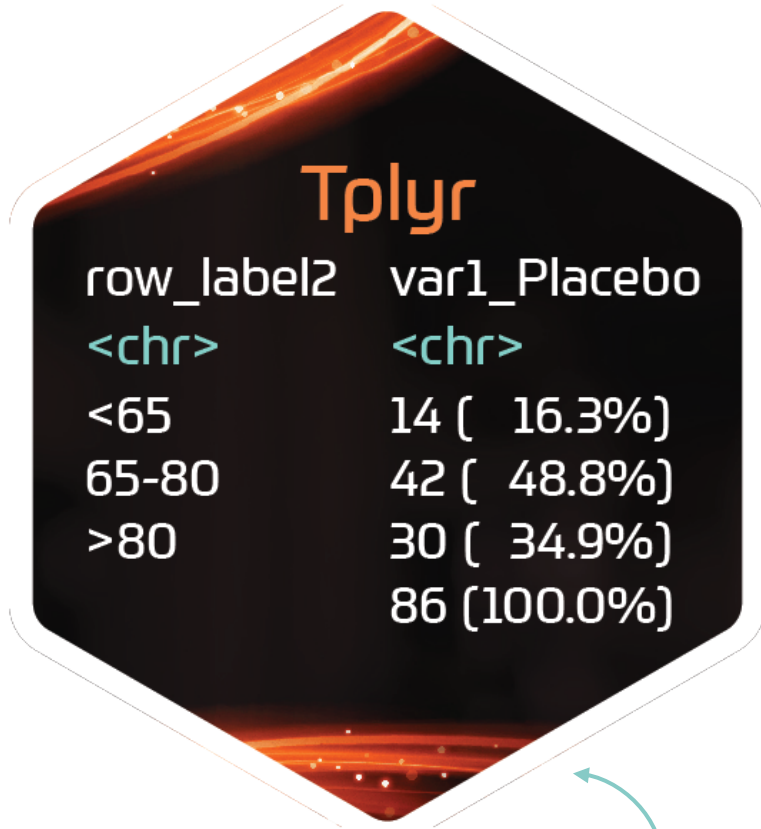
[1] P-values are results of ANOVA treatment group comparison for continuous variable and Pearson's chisquare test for categorical variables.

NOTE: Duration of disease is computed as months between date of enrollment and date of onset of the first definite symptoms of Alzheimer's disease.

Source: ~/CDISC_pilot_replication/programs/t-14-2-01.R

14:48 Thursday, April 30, 2020

Next Steps



Click me for
the code!

- CDISC Pilot scripts were standalone
- Table programming is often redundant, hence macros
- Tplyr breaks tables programming into fundamental pieces
- Program with the output in mind
- MIT License
- Qualification document included

Questions?



Michael Stackhouse

Chief Innovation Officer, Atorus Research, LLC

Mike.Stackhouse@AtorusResearch.com

[LinkedIn](#)

[Twitter](#)



atorus

Your Complete Data Solution