Package: ProAE (via r-universe)

October 5, 2026

Type Package

Title PRO-CTCAE Scoring, Analysis, and Graphical Tools

Version 1.0.4

Description A collection of tools to facilitate standardized analysis and graphical procedures when using the National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) and other PRO measurements.

License GPL-3

Encoding UTF-8

LazyData true

Imports ggplot2, ggnewscale, ggtext, dplyr, Hmisc, magrittr, ggpattern, DescTools, gridExtra, kableExtra

Suggests knitr, rmarkdown, testthat

VignetteBuilder knitr

Depends R (>= 4.0.0)

RoxygenNote 7.2.3

NeedsCompilation no

Author Blake Langlais [aut], Brie Noble [ctb, cre], Mia Truman [ctb], Molly Voss [ctb], Amylou Dueck [aut]

Maintainer Brie Noble <noble.brie@mayo.edu>

Config/pak/sysreqs libabsl-dev cmake libfontconfig1-dev libfreetype6-dev libfribidi-dev libgdal-dev gdal-bin libgeos-dev make libharfbuzz-dev libicu-dev libjpeg-dev libpng-dev libuv1-dev libxml2-dev libssl-dev libproj-dev libsqlite3-dev libudunits2-dev libx11-dev zlib1g-dev

Repository https://bnnoble.r-universe.dev

Date/Publication 2026-01-08 23:10:02 UTC

RemoteUrl https://github.com/cran/ProAE

RemoteRef HEAD

RemoteSha 529b0eee9915b99a3665f84d65c78bb79f689e28

Contents

ProAE . . . 2
PROCTCAE_table . . . 2
tox_acute . . . 3
tox_chronic . . . 3
tox_cumulative . . . 4
tox_cyclic . . . 5
tox_late . . . 5
toxAUC . . . 6
toxFigures . . . 8
toxScores . . . 10
toxSummary . . . 12
toxTables . . . 13

Index **15**

`ProAE` *ProAE.*

Description

This package is a collection of tools for analyzing and plotting PRO-CTCAE data

---

`PROCTCAE_table` *PRO-CTCAE variable / label crosswalk*

Description

A crosswalk / look-up table of expected variable names for associated PRO-CTCAE symptom items.

Format

A data frame with 124 rows and 2 variables

Details

- name. Expected variable name - item number/letter corresponds to the NCI-PRO-CTCAE (English) Item Library Version 1.0
- short_label. Item label including the item symptom group and frequency, severity, interference, or presence component

Author(s)

Blake Langlais

References

https://healthcaredelivery.cancer.gov/pro-ctcae/instrument-pro.html

---

`tox_acute` *PRO-CTCAE data reflecting acute drug toxicity*

Description

Simulated example data where the drug group experiences acute toxicity followed by symptom abatement over the course of treatment.

Format

A data frame with 1400 rows and 5 variables

Details

- id. Subject identifier
- Cycle. Time variable denoting visits/cycles (1-10)
- arm. Treatment groups include drug and placebo
- PROCTCAE_9A_SCL. PRO-CTCAE frequency item for nausea
- PROCTCAE_9B_SCL. PRO-CTCAE severity item for nausea
- PROCTCAE_9_COMP. PRO-CTCAE composite score for nausea

Author(s)

Blake Langlais

`tox_chronic` *PRO-CTCAE data reflecting chronic drug toxicity*

Description

Simulated example data where the drug group experiences chronic toxicity over the course of treatment.

Format

A data frame with 1400 rows and 5 variables

Details

- id. Subject identifier
- Cycle. Time variable denoting visits/cycles (1-10)
- arm. Treatment groups include drug and placebo
- PROCTCAE_9A_SCL. PRO-CTCAE frequency item for nausea
- PROCTCAE_9B_SCL. PRO-CTCAE severity item for nausea
- PROCTCAE_9_COMP. PRO-CTCAE composite score for nausea

Author(s)

Blake Langlais

---

tox_cumulative *PRO-CTCAE data reflecting cumulative drug toxicity*

Description

Simulated example data where drug toxicity is cumulative over the course of treatment.

Format

A data frame with 1400 rows and 5 variables

Details

- id. Subject identifier
- Cycle. Time variable denoting visits/cycles (1-10)
- arm. Treatment groups include drug and placebo
- PROCTCAE_9A_SCL. PRO-CTCAE frequency item for nausea
- PROCTCAE_9B_SCL. PRO-CTCAE severity item for nausea
- PROCTCAE_9_COMP. PRO-CTCAE composite score for nausea

Author(s)

Blake Langlais

tox_cyclic *PRO-CTCAE data reflecting cyclical drug toxicity*

Description

Simulated example data where the drug group experiences cyclically toxicity post-treatment administration

Format

A data frame with 1400 rows and 5 variables

Details

- id. Subject identifier
- Cycle. Time variable denoting visits/cycles (1-10)
- arm. Treatment groups include drug and placebo
- PROCTCAE_9A_SCL. PRO-CTCAE frequency item for nausea
- PROCTCAE_9B_SCL. PRO-CTCAE severity item for nausea
- PROCTCAE_9_COMP. PRO-CTCAE composite score for nausea

Author(s)

Blake Langlais

tox_late *PRO-CTCAE data reflecting late incipient drug toxicity*

Description

Simulated example data where the drug group experiences late incipient toxicity towards the end of the treatment period.

Format

A data frame with 1400 rows and 5 variables

Details

- id. Subject identifier
- Cycle. Time variable denoting visits/cycles (1-10)
- arm. Treatment groups include drug and placebo
- PROCTCAE_9A_SCL. PRO-CTCAE frequency item for nausea
- PROCTCAE_9B_SCL. PRO-CTCAE severity item for nausea
- PROCTCAE_9_COMP. PRO-CTCAE composite score for nausea

Author(s)

Blake Langlais

`toxAUC` *Create longitudinal mean score line plots for PRO-CTCAE data with modified Area Under the Curve (AUC) estimates showing descriptive symptomatic adverse event burden worsening and improvement from baseline.*

Description

Data format should be in 'long' format, where each PRO-CTCAE item is a variable/column. AUC calculations will only include subjects with non- missing baseline values (within each PRO-CTCAE item).

Usage

```
toxAUC(
  dsn,
  id_var,
  cycle_var,
  baseline_val,
  arm_var = NA,
  auc = "above",
  cycle_limit = NA,
  y_limit = 4,
  tab_ymin = NA,
  tab_ymax = NA,
  round_dec = 2,
  permute_tests = FALSE,
  permute_n = 2000,
  bootstrap_ci = FALSE,
  bootstrap_ci_alpha = 0.05,
  bootstrap_n = 2000,
  arm_colors = NA,
  x_label = NA,
  add_item_title = FALSE,
  cycle_label = FALSE,
  cycle_vals = NA,
  cycle_labs = NA
)
```

Arguments

`dsn`	A data.frame object with PRO-CTCAE data
`id_var`	A character string. Name of ID variable differentiating each unique patient.

`cycle_var` A character string. Name of variable differentiating one longitudinal/repeated. PRO-CTCAE survey from another, within an individual ID.

`baseline_val` A number indicating the expected baseline cycle/time point.

`arm_var` A character string. Name of arm variable differentiating treatment groups. Must be character or factor class. Overall frequencies will be reported if no arm/grouping variable is provided. Defaults to `NA`.

`auc` A character string. Specifies the partitioning of area shown. options include: `"above"` = accumulated area above the baseline horizontal can be interpreted as worsening severity from baseline. `"below"` = accumulated area below the baseline horizontal can be interpreted the as the amount of decreased severity from baseline. `"both"` = accumulated area above and below the baseline horizontal are shown. Defaults to `"above"`.

`cycle_limit` A number. Limit the number of cycles to be use to calculate the AUC metrics up to and including a given cycle number. All available cycle time points are used if no cycle number is provided. Defaults to `NA`.

`y_limit` A number. Y axis limit for plots. Defaults to `4`.

`tab_ymin` A number. Y axis coordinate for adjusting the vertical placement of the AUC table within the figure. Defaults to `NA`.

`tab_ymax` A number. Y axis coordinate for adjusting the vertical placement of the AUC table within the figure. Defaults to `NA`.

`round_dec` A number. Number of decimal places to be shown within the AUC table. Defaults to `2`.

`permute_tests` Logical. Calls to calculate p values comparing the difference in AUC between two arms using a permutation test. Typical two- sided null hypothesis for a permutation test is applied. That is, assigning subjects to groups is interchangeable when calculating AUC. Computation time may be extensive depending on data size, number of PRO-CTCAE items, and number of permutations called. Consider staring out an open window or crafting a haiku during this time. Defaults to `FALSE`.

`permute_n` A number. The number of permutations to be used for permutation tests. Defaults to `2000`.

`bootstrap_ci` Logical. Calls to construct alpha-level confidence intervals for the difference in AUC between arms. Similar considerations for computation time as `permute_tests` are recommended here. Defaults to `FALSE`.

`bootstrap_ci_alpha` A number. Specifies the alpha level for bootstrap confidence intervals. Must be between 0 and 1. Defaults to `0.05`.

`bootstrap_n` A number. The number of bootstrap iterations to be used for bootstrap confidence intervals. Defaults to `2000`.

`arm_colors` A column vector of valid colors. Allows the user to define the colors of arms shown in the returned figure. Column vector must have length greater than or equal to the number of arms. Default colors provided.

`x_label` A character string. Label for the x axis of the plot. Defaults to `NA` resulting in x axis labeled with `cycle_var` variable name.

add_item_title	Logical. Adds the item short label to the title of each figure. Defaults to FALSE.
cycle_label	Logical. Assign custom labels to cycles/time point. If TRUE, the cycle_vals and cycle_labs must also be specified.
cycle_vals	Numeric column vector. Vector of values seen within the cycle_var variable. Must be same length of cycle_labs. Defaults to NA.
cycle_labs	Character column vector. Vector of labels to be mapped to the associated cycle_vals. Must be same length of cycle_vals. Defaults to NA.

Examples

```
## Not run:
AUC=toxAUC(dsn = ProAE::tox_acute,
id_var = "id",
cycle_var = "Cycle",
baseline_val = 1)
AUC[[1]]

## End(Not run)
```

---

toxFigures *Create PRO-CTCAE severity frequency distribution figures for individual survey items and composite scores*

Description

Data format should be in 'long' format, where each PRO-CTCAE item is a variable/column.

Usage

```
toxFigures(
  dsn,
  id_var,
  cycle_var,
  baseline_val,
  arm_var = NA,
  plot_limit = NA,
  colors = 1,
  bar_label = 0,
  cycle_label = FALSE,
  cycle_vals = NA,
  cycle_labs = NA,
  summary_only = FALSE,
  summary_highlight = FALSE,
  cycles_only = TRUE,
  x_lab_angle = 0,
  x_lab_vjust = 1,
  x_lab_hjust = 0,
```

```
  x_label = "Randomized Treatment Assignment",
  y_label = "Percent of Total Frequency",
  footnote_break = FALSE,
  suppress_legend = FALSE,
  add_item_title = FALSE
)
```

Arguments

dsn	A data.frame object with PRO-CTCAE data
id_var	A character string.Name of ID variable differentiating each PRO-CTCAE survey/participant entered as a quoted string.
cycle_var	A character string. Name of variable differentiating one longitudinal/repeated PRO-CTCAE survey from another, within an individual ID.
baseline_val	A number indicating the expected baseline cycle/time point.
arm_var	A character string. Name of arm variable differentiating treatment groups. Must be character or factor class. Overall AUC will be reported if no arm/grouping variable is provided. Defaults to NA.
plot_limit	A number. Limit the number of cycles to be plotted up to and including a given cycle number. All available cycle time points are plotted if no cycle number is provided. Defaults to NA.
colors	A number. Specify the coloring scheme of symptom scores within frequency bars. Options include: 1 = Blue and red color shading, 2 = qualitative color shades (color blind friendly), 3 = black and white. Defaults to 1.
bar_label	A number. Label frequency bars with sample size (n) or percent shown on the y-axis. Label options include: 1 = sample size (n) within each cycle (symptom score 0 or higher), 2 = sample size (n) within each cycle with present symptoms (symptom score > 0), 3 = sample size (n) within each cycle with severe symptoms (symptom score >= 3), 4 = percent of subjects within each cycle with present symptoms (symptom score > 0), 5 = percent of subjects within each cycle with severe symptoms (symptom score >= 3). No labels will be applied if not specified. Defaults to NA.
cycle_label	Logical. Assign custom labels to cycles/time point. If TRUE, the cycle_vals and cycle_labs must also be specified.
cycle_vals	Numeric column vector. Vector of values seen within the cycle_var variable. Must be same length of cycle_labs. Defaults to NA.
cycle_labs	Character column vector. Vector of labels to be mapped to the associated cycle_vals. Must be same length of cycle_vals. Defaults to NA.
summary_only	Logical. Only display the summary measures in figures / Suppress the individual time points from plotting. Defaults to FALSE.
summary_highlight	Logical. Add black box around summary measure bar chart. Defaults to FALSE.
cycles_only	Logical. Only display the longitudinal time points in figures / Suppress the summary measures from plotting. Defaults to TRUE.

x_lab_angle	A integer between 0 and 360. Allows the user to rotate the x axis labels in order to fit long arm names (0 or 45 recommended). Defaults to 0.
x_lab_vjust	A number. A ggplot2 object option. Allows the user to vertically adjusts the x axis labels in order to fit arm names. Defaults to 1.
x_lab_hjust	A number. A ggplot2 object option. Allows the user to horizontally adjusts the x axis labels in order to fit arm names. Defaults to 0.
x_label	A character string. Label for the x axis of the plot. Defaults to "Randomized Treatment Assignment" if arm_var is specified, defaults to "Overall" if not arm_var is specified.
y_label	A character string. Label for the y axis of the plot. Defaults to "Percent of Total Frequency".
footnote_break	Logical. Add a line break to the footnote Defaults to FALSE.
suppress_legend	Logical. Suppress the legend from appearing in figure. Defaults to FALSE.
add_item_title	Logical. Add the symptom item name as a title to the figure. Defaults to FALSE.

Value

A list object. The returned object is a (k X 2) or (k x 3) nested list. Where k is the number of PRO-CTCAE item groups (e.g. pain, fatigue, nausea); list[[1 ... i ... k]]. For each list item there are 2 or 3 elements. The 1st element of each list item is the name of the PRO-CTCAE item group returned as a string. The 2nd element is the PRO-CTCAE figure as a ggplot object. These objects can be modified as such.

Examples

```
## Not run:
fig_acute = toxFigures(dsn = ProAE::tox_acute,
 cycle_var = "Cycle",
 baseline_val = 1,
 arm_var = "arm",
 id_var = "id",
 x_lab_angle = -45,
 x_lab_vjust =  .3,
 x_lab_hjust = .2,
 colors = 2)
fig_acute[[1]]

## End(Not run)
```

---

toxScores *Re-code PRO-CTCAE text responses, apply a zero-imputation procedures, and construct PRO-CTCAE composite scores.*

Description

This function takes in a data frame set with PRO-CTCAE survey text fields/responses and returns a data frame with appropriate numerical re-coding. This function will accept 1 or up to all 124 PRO-CTCAE survey fields. All PRO-CTCAE variable names MUST conform to a pre-specified naming structure. PRO-CTCAE variable names are made up of FOUR components: 1)'PROCTCAE', 2) number [1,2,3, ..., i, ..., 80], 3) 'A', 'B', or 'C' component of the i-th PRO-CTCAE field, 4) and 'SCL' (if severity, interference, or frequency) or 'IND' (if yes/no variable). Each component must be delimited by an underscore (_)

Usage

```
toxScores(
  dsn,
  reformat = FALSE,
  impute = FALSE,
  composites = FALSE,
  short_labels = FALSE
)
```

Arguments

dsn	A data.frame object with PRO-CTCAE data
reformat	Reformat PRO-CTCAE text responses to numeric scores. Defaults to FALSE.
impute	Apply zero-imputation where appropriate. Defaults to FALSE.
composites	Construct composite score using available PRO-CTCAE variables within dsn. Defaults to FALSE.
short_labels	Add PRO-CTCAE short labels to available PRO-CTCAE variables within returned object

Details

- [EX1] Question 1 of PRO-CTCAE should be: PROCTCAE_1A_SCL
- [EX2] Question 48 of PRO-CTCAE should be: PROCTCAE_48A_SCL, PROCTCAE_48B_SCL, PROCTCAE_48C_SCL
- [EX3]Question 73 of PRO-CTCAE should be: PROCTCAE_73A_IND

This function also constructs PRO-CTCAE composite scores. Composite score variables for respective PRO-CTCAE item groups are created and named as PROCTCAE_##_COMP.

1. https://healthcaredelivery.cancer.gov/pro-ctcae/pro-ctcae_english.pdf
2. Ethan Basch, et al. Development of a Composite Scoring Algorithm for the National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). ISOQOL 2019
3. Basch E, et al. Composite Grading Algorithm for the National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). Clinical Trials 2020.

Data format should be in 'long' format, where each PRO-CTCAE item is a variable/column.

Value

A data.frame object.

Examples

```
tox_acute_comp = toxScores(dsn = ProAE::tox_acute, composites = TRUE)
```

`toxSummary` *Create patient-level and group-level summary statistics.*

Description

Data format should be in 'long' format, where each PRO-CTCAE item is a variable/column.

Usage

```
toxSummary(
  dsn,
  id_var,
  cycle_var,
  summary_measure,
  baseline_val = NA,
  arm_var = NA
)
```

Arguments

`dsn`	A data.frame object with PRO-CTCAE data.
`id_var`	A character string. Name of ID variable differentiating each PRO-CTCAE survey/participant entered as a quoted string.
`cycle_var`	A character string. Name of variable differentiating one longitudinal/repeated. PRO-CTCAE survey from another, within an individual ID.
`summary_measure`	A character string. Type of summary statistic to be used. Please consult current literature for appropriate interpretations of the summary measure selected and suitable analysis procedures for comparing groups. Options include: `"max"` = Use subjects' maximum score. `"max_post_bl"` = Use subjects' maximum score post-baseline visit. `"bl_adjusted"` = Use subjects' baseline adjusted score over the study period. The baseline adjusted score is derived by the following: If the maximum score post-baseline is more severe than the baseline score, then the use maximum score post-baseline is used as the adjusted score. Otherwise, if the maximum score post-baseline is the same or less serve than the baseline score, then zero (0) is used as the adjusted score. `"toxicity_index"` = Construct patient-level toxicity index. `"AUC_worsening"` = Calculate group-level AUC describing
`baseline_val`	A number indicating the expected baseline cycle/time point.
`arm_var`	A character string. Name of arm variable differentiating treatment arms or other grouping factor. Required for group-level summary measures.

Value

A data.frame with only the id and PRO-CTCAE variables being summarized. Each subject will now only have 1 observation (PRO-CTCAE variables are now the summary measure value).

Examples

```
toxSummary(dsn=ProAE::tox_acute,
id_var="id",
cycle_var="Cycle",
baseline_val=1,
summary_measure = "max")
```

---

toxTables *Create toxicity tables for individual and composite PRO-CTCAE survey items.*

Description

Data format should be in 'long' format, where each PRO-CTCAE item is a variable/column.

Usage

```
toxTables(
  dsn,
  id_var,
  cycle_var,
  baseline_val,
  type = "bl_adjusted",
  test = "c",
  riskdiff = FALSE,
  risk_ci = "wald",
  risk_ci_alpha = 0.05,
  arm_var = NA,
  cycle_limit = NA
)
```

Arguments

dsn	A data.frame object with PRO-CTCAE data.
id_var	A character string. Name of ID variable differentiating each PRO-CTCAE survey/participant entered as a quoted string.
cycle_var	A character string. Name of variable differentiating one longitudinal/repeated. PRO-CTCAE survey from another, within an individual ID.
baseline_val	A number indicating the expected baseline cycle/time point.

type	A character string. Type of summary measure to be be used. Options include: "max_post_bl" = Use subjects' maximum score post-baseline visit. "bl_adjusted" = Use subjects' baseline adjusted score over the study period. "max" = Use subjects' maximum score over the study period. The baseline adjusted score is derived by the following: If the maximum score post-baseline is more severe than the baseline score, then the use maximum score post-baseline is used as the adjusted score. Otherwise, if the maximum score post-baseline is the same or less serve than the baseline score, then zero (0) is used as the adjusted score. Defaults to "bl_adjusted".
test	A character string. Specify the statistical test to apply where comparing rates among arms. Options include: "c" = chi square, "f" = fisher's exact. Defaults to "c".
riskdiff	Logical. Calculates risk differences between two arms. Valid if there are only two arms in the data.frame specified. This option will countermand options called with the test parameter. Defaults to FALSE.
risk_ci	A character string. Specify the confidence interval type to be constructed for risk differences. Options include: "wald", and "agresti-caffo". Defaults to "wald".
risk_ci_alpha	A number between 0 and 1. Specify the alpha level of the risk difference confidence intervals. Defaults to 0.05.
arm_var	A character string. Name of arm variable differentiating treatment groups. Must be character or factor class. Overall frequencies will be reported if no arm/grouping variable is provided. Defaults to NA.
cycle_limit	A number. Limit the data to be analyzed up to and including a given cycle number or time point. Defaults to NA.

Value

A list object with data.frame elements for individual items and composite scores.

Examples

```
toxTables(dsn=ProAE::tox_acute, id_var="id", cycle_var="Cycle", baseline_val=1)
```

Index

∗ **data**
PROCTCAE_table, 2
tox_acute, 3
tox_chronic, 3
tox_cumulative, 4
tox_cyclic, 5
tox_late, 5

ProAE, 2
ProAE-package *(*ProAE*)*, 2
PROCTCAE_table, 2

tox_acute, 3
tox_chronic, 3
tox_cumulative, 4
tox_cyclic, 5
tox_late, 5
toxAUC, 6
toxFigures, 8
toxScores, 10
toxSummary, 12
toxTables, 13