Package ‘ProAE’

January 20, 2023

Type Package

Title PRO-CTCAE Scoring, Analysis, and Graphical Tools

Version 0.2.12

Author Blake Langlais [aut, cre],
     Brie Noble [ctb],
     Molly Klanderman [ctb],
     Amylou Dueck [aut]

Maintainer Blake Langlais <langlais.blake@mayo.edu>

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).

License GPL-3

Encoding UTF-8

LazyData true

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

Suggests knitr, rmarkdown, testthat

VignetteBuilder knitr

Depends R (>= 4.0.0)

RoxygenNote 7.2.2

NeedsCompilation no

Repository CRAN

Date/Publication 2023-01-20 20:40:02 UTC

R topics documented:

ProAE .......................................................... 2
PROCTCAE_table ............................................ 2
toxAUC ....................................................... 3
This package is a collection of tools for analyzing and plotting PRO-CTCAE data

**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](https://healthcaredelivery.cancer.gov/pro-ctcae/instrument-pro.html)
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**

```r
toxAUC(
  dsn,
  id_var,
  cycle_var,
  baseline_val,
  arm_var = NA,
  cycle_limit = NA,
  y_limit = 4,
  tab_ymin = NA,
  tab_ymax = NA,
  round_dec = 2,
  overwrite_title = NA,
  permute_tests = FALSE,
  permute_n = 2000,
  bootstrap_ci = FALSE,
  bootstrap_ci_alpha = 0.05,
  bootstrap_n = 2000
)
```

**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`.
- **cycle_limit**: A number. Limit the number of cycles to be used 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. Limit the range of the y-axis.
- **tab_ymin** and **tab_ymax**: Minimum and maximum values for the y-axis.
- **round_dec**: Number of decimal places to round the AUC estimates.
- **overwrite_title**: A character string. Title to overwrite the default title.
- **permute_tests**: Logical. Perform permutation tests to assess statistical significance. Defaults to `FALSE`.
- **permute_n**: Number of permutations for the permutation tests. Defaults to 2000.
- **bootstrap_ci**: Logical. Compute bootstrap confidence intervals. Defaults to `FALSE`.
- **bootstrap_ci_alpha**: Level of confidence for the bootstrap confidence intervals. Defaults to 0.05.
- **bootstrap_n**: Number of bootstrap samples to compute the confidence intervals. Defaults to 2000.
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.

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.
overwrite_title A character string. Add main title to plots. Defaults to NA.
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.

Examples

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

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,
    cycles_only = FALSE,
    x_lab_angle = 0,
    x_lab_vjust = 1,
    x_lab_hjust = 0,
    x_label = "Randomized Treatment Assignment",
    footnote_break = 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 grades 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 grade 0 or higher), 2 = sample size (n) within each cycle with present symptoms (symptom grade > 0), 3 = sample size (n) within each cycle with severe symptoms (symptom grade >= 3), 4 = percent of subjects within each cycle with present symptoms (symptom grade > 0), 5 = percent of subjects within each cycle with severe symptoms (symptom grade >= 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.

cycles_only Logical. Only display the longitudinal time points in figures / Suppress the summary measures from plotting. Defaults to FALSE.

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.

footnote_break Logical. Add a line break to the footnote 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.

Examples

```r
## 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 grades.

---

**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 \([1,2,3, ..., 80]\), 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**

```r
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 grade 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 grades. Composite grade variables for respective PRO-CTCAE item groups are created and named as PROCTCAE_##_COMP.


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.

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 severe 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")

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. 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 severe 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 grades.

Examples

toxTables(dsn=ProAE::tox_acute, id_var="id", cycle_var="Cycle", baseline_val=1)
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_78A_SC. PRO-CTCAE frequency item for nosebleeds
• PROCTCAE_78B_SC. PRO-CTCAE severity item for nosebleeds

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_78A_SC. PRO-CTCAE frequency item for nosebleeds
• PROCTCAE_78B_SC. PRO-CTCAE severity item for nosebleeds
**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_78A_SC. PRO-CTCAE frequency item for nosebleeds
- PROCTCAE_78B_SC. PRO-CTCAE severity item for nosebleeds

**tox_cyclic**  
*PRO-CTCAE data reflecting cyclical drug toxicity*

**Description**  
Simulated example data where the drug group experiences cyclicly 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_78A_SC. PRO-CTCAE frequency item for nosebleeds
- PROCTCAE_78B_SC. PRO-CTCAE severity item for nosebleeds
**tox_late**

**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_78A_SC. PRO-CTCAE frequency item for nosebleeds
- PROCTCAE_78B_SC. PRO-CTCAE severity item for nosebleeds

**Author(s)**

Blake Langlais
Index

* data
  PROCTCAE_table, 2
  tox_acute, 11
  tox_chronic, 11
  tox_cumulative, 12
  tox_cyclic, 12
  tox_late, 13

ProAE, 2
PROCTCAE_table, 2

  tox_acute, 11
  tox_chronic, 11
  tox_cumulative, 12
  tox_cyclic, 12
  tox_late, 13
  toxAUC, 3
  toxFigures, 4
  toxScores, 7
  toxSummary, 8
  toxTables, 9