Package 'Immunotherapy.Design'

May 18, 2020

Title Study design for immunotherapy clinical trials

Version 1.1.0

Date 2020-05-18

Author Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov>, Yongsoek Park <yongpark@pitt.edu> and Bin Zhu <bin.zhu@nih.gov>

Description Perform sample size, power calculation and subsequent analysis for Immuno-oncology (IO) trials composed of responders and nonresponders.

Maintainer Bill Wheeler <wheelerb@imsweb.com>

Depends survival, msm

License GPL-2

NeedsCompilation yes

R topics documented:

Immunotherapy.Design-package

Study design for immunotherapy clinical trials

Description

Perform sample size, power calculation and subsequent analysis for Immuno-oncology (IO) trials composed of responders and nonresponders.

Author(s)

Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov>, Yongsoek Park <yongpark@pitt.edu> and Bin Zhu <bin.zhu@nih.gov>

References

Xu, Z., Park, Y., Liu, K. and Zhu, B. Treating non-responders: pitfalls and implications for cancer immunotherapy trial design. Journal of Hematology & Oncology 13, 20 (2020).

Xu, Z., Zhu, B. and Park, Y. Designing immuno-oncology clinical trials composed of responders and nonresponders. Statistics in Medicine. (Under Revision).

data

Data for examples

Description

Data for examples.

Details

A data frame used in the examples.

Examples

```
data(data, package="Immunotherapy.Design")

# Display some of the data
data[1:5, ]
```

generate_data *Simulated data*

Description

Generate simulated data

Usage

```
generate_data(nmax=500, rand_ratio=0.5, effect_p=0.6, enroll_rate=380*0.25/6,
              lambda1=0.117, HR=0.5, tau=12*5, t1=1)
```

Arguments

nmax	Sample size
rand_ratio	Allocation ratio
effect_p	Proportion of responders in the treatment arm at baseline
enroll_rate	Enrollment rate in subjects per month
lambda1	Baseline hazard in terms of months
HR	Hazard ratio of responders against controls
tau	Total study duration
t1	Delayed duration

Value

A data frame with columns:

Name	Description
id	id variable
trt	treatment allocation: 1 = treatment arm
Z	patient's response status
tau	total study duration
enroll_time	patients' enrollment times
time_to_event	patients' event times
event_status	censoring indicator
X	observational time
t1	delayed duration

Author(s)

Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov>, Yongsoek Park <yongpark@pitt.edu> and Bin Zhu <bin.zhu@nih.gov>

Examples

```
data <- generate_data()
data[1:5, ]
```

getHazard *Compute initial estimates for the baseline hazard*

Description

Calls the coxph function to compute initial estimates for the baseline hazard

Usage

```
getHazard(time, treatment, event_status)
```

Arguments

time	Vector of times.
treatment	Binary vector of treatments (1=subject received treatment).
event_status	Binary vector of event status (1=subject experienced an event).

Value

Vector of baseline hazards ordered by the event times.

Author(s)

Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov> and Bin Zhu <bin.zhu@nih.gov>

See Also

PRIME.EM

Examples

```
data(data, package="Immunotherapy.Design")
lambda0 <- getHazard(data[, "X"], data[, "trt"], data[, "event_status"])
lambda0[1:10]
```

N.Pembedded.P *Sample Size*

Description

Compute the sample size for a given power for the parametric model

Usage

```
N.Pembedded.P(power=0.8, rand_ratio=0.5, effect_p=0.6, enroll_rate=380*0.25/6,
            lambda1=0.117, HR=0.5, tau=12*5, t1=1, maxiter=100000, stopTol=1e-4,
              alpha=0.05, num_rand=1000, nsim=10000, min.N=100, max.N=700,
              tol.power=0.01, tol.N=1, print=1,
              min.sample.size=50, min.n.event=5, min.per.trt=0.25)
```

Arguments

power	The desired power. The default is 0.8.
rand_ratio	Allocation ratio
effect_p	Proportion of responders in the treatment arm
enroll_rate	Enrollment rate in subjects per month
lambda1	Baseline hazard in terms of months
HR	Hazard ratio
tau	Total study duration
t1	Delayed duration
maxiter	Maximum number of iterations in the EM algorithm. The default is 100000.
stopTol	Stopping tolerance in the EM algorithm. The default is 1e-4.
alpha	Significance level. The default is 0.05.
num_rand	Number of replications in the re-randomization test. The default is 1000.
nsim	Number of simulations in computing power (**see Details**). The default is 10000.
min.N	Lower bound for the sample size. The default is 100.
max.N	Upper bound for the sample size. The default is 700.
tol.power	Stopping tolerance for the power. The default is 0.01.
tol.N	Stopping tolerance for the sample size. The default is 1.
print	0 or 1 to print information. The default is 1.
min.sample.size	Minimum sample size. The default is 50.
min.n.event	Minimum number of events. The default is 5.
min.per.trt	Minimum proportion of controls and treated subjects. The default is 0.25.

Details

This uses a bisection method to estimate the sample size. At each iteration, the estimated power `power_est` is computed using `Pow.Pembedded.P` for a given sample size holding all other parameters fixed. The algorithm terminates when `abs(power -power_est) <= tol.power` or when the length of the estimated interval containing the sample size is less than or equal to `tol.N`.

NOTE:
It is important to note that the power for a given sample size is estimated by running a simulation. Thus, by setting a different seed, a different result may be returned. Therefore, to ensure a more precise estimated sample size, set the option `nsim` to a large value and/or run this function several times by setting different seeds and examine the distribution of returned sample sizes.

Value

A list containing the sample size and power.

Author(s)

Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov>, Yongsoek Park <yongpark@pitt.edu> and Bin Zhu <bin.zhu@nih.gov>

See Also

`Pow.Pembedded.P`

N.PRIME.LRT *Sample Size*

Description

Compute the sample size for a given a power accounting for response dichotomy based on the PRIME procedure

Usage

```
N.PRIME.LRT(power=0.8, rand_ratio=0.5, effect_p=0.6, enroll_rate=380*0.25/6,
            lambda1=0.117, HR=0.5, tau=12*5, t1=1, maxiter=100000, stopTol=1e-4,
              alpha=0.05, nsim=10000, min.N=100, max.N=700,
              tol.power=0.01, tol.N=1, print=1,
              min.sample.size=50, min.n.event=5, min.per.trt=0.25)
```

Arguments

power	The desired power. The default is 0.8.
rand_ratio	Allocation ratio
effect_p	Proportion of responders in the treatment arm at baseline
enroll_rate	Enrollment rate in subjects per month
lambda1	Baseline hazard in terms of months
HR	Hazard ratio of responders against controls
tau	Total study duration
t1	Delayed duration
maxiter	Maximum number of iterations in the EM algorithm. The default is 100000.
stopTol	Stopping tolerance in the EM algorithm. The default is 1e-4.
alpha	Significance level. The default is 0.05.
nsim	Number of simulations in computing power (**see Details**). The default is 10000.
min.N	Lower bound for the sample size. The default is 100.
max.N	Upper bound for the sample size. The default is 700.
tol.power	Stopping tolerance for the power. The default is 0.01.
tol.N	Stopping tolerance for the sample size. The default is 1.
print	0 or 1 to print information. The default is 1.
min.sample.size	Minimum sample size. The default is 50.
min.n.event	Minimum number of events. The default is 5.
min.per.trt	Minimum proportion of controls and treated subjects. The default is 0.25.

Details

This uses a bisection method to estimate the sample size. At each iteration, the estimated power `power_est` is computed using `Pow.PRIME.LRT` for a given sample size holding all other parameters fixed. The algorithm terminates when `abs(power -power_est) <= tol.power` or when the length of the estimated interval containing the sample size is less than or equal to `tol.N`.

NOTE:
It is important to note that the power for a given sample size is estimated by running a simulation. Thus, by setting a different seed, a different result may be returned. Therefore, to ensure a more precise estimated sample size, set the option `nsim` to a large value and/or run this function several times by setting different seeds and examine the distribution of returned sample sizes.

Value

A list containing the sample size and power.

Author(s)

Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov>, Yongsoek Park <yongpark@pitt.edu> and Bin Zhu <bin.zhu@nih.gov>

See Also

`N.Pembedded.P`

`Pembedded.EM.P` *EM algorithm*

Description

Parametric EM algorithm

Usage

```
Pembedded.EM.P(data, time.var="X", trt.var="trt", status.var="event_status",
          effect_p=0.6, t1=1, probResponder=NULL,
          stopTol=1e-5, maxiter=100000, print=0)
```

Arguments

`data`	Data frame or matrix containing a time-to-event variable (`time.var`), a treatment variable (`trt.var`), and a censoring variable (`status.var`).
`time.var`	Observational time variable name in `data` (months). The default is "X".
`trt.var`	Binary treatment assignment indicator name in `data` coded as 0 for controls and 1 for treated subjects.
`status.var`	Name of the binary censoring variable in `data` coded as 0 for censored subjects and 1 for subjects that experienced an event.
`effect_p`	Proportion of responders among the treated subjects. The default is 0.6.
`t1`	Delayed duration. The default is 1 (month).

probResponder	NULL or vector of initial probabilities of a treated subject being a responder. The default is NULL so that the initial probability is 0.5 for treated subjects.
stopTol	Stopping tolerance. The default is 1e-5.
maxiter	Maximum number of iterations. The default is 100000.
print	0-2 to print information. Larger values will print more information. The default is 0.

Value

A list containing the objects:

Name	Description
converged	TRUE if EM algorithm converged
lambda	estimated hazard ratio of responders versus controls
baseline	estimated baseline hazard
probResponder	estimated probability of a treated subject being a responder
loglike	log-likelihood value at the final estimates

Author(s)

Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov>, Yongsoek Park <yongpark@pitt.edu> and Bin Zhu <bin.zhu@nih.gov>

See Also

PRIME.EM

Examples

```
data(data, package="Immunotherapy.Design")
ret <- Pembedded.EM.P(data)
ret$lambda
```

Pembedded.ReRandomizationTest.P

Randomization test

Description

Compute a randomization test p-value where test statisitic is calculated based on a parametric model.

Usage

```
Pembedded.ReRandomizationTest.P(data, time.var="X", trt.var="trt", status.var="event_status",
       effect_p=0.6, t1=1, stopTol=1e-5, maxiter=100000, print=0, num_rand=10000,
           min.sample.size=50, min.n.event=5, min.per.trt=0.25)
```

Arguments

data	Data frame or matrix containing a time-to-event variable (time.var), a treatment variable (trt.var), and a censoring variable (status.var).
time.var	Observational variable name in data.
trt.var	Name of treatment assignment indicator in data coded as 0 for control subjects and 1 for treated subjects.
status.var	Name of the binary censoring variable in data coded as 0 for censored subjects and 1 for subjects that experienced an event.
effect_p	Proportion of responders among the treated subjects. The default is 0.6.
t1	Delayed duration. The default is 1 (month).
stopTol	Stopping tolerance in the EM algorithm. The default is 1e-5.
maxiter	Maximum number of iterations in the EM algorithm. The default is 100000.
print	0-2 to print information. Larger values will print more information. The default is 0.
num_rand	The number of replications in the re-randomization test. The default is 10000.
min.sample.size	Minimum sample size. The default is 50.
min.n.event	Minimum number of events. The default is 5.
min.per.trt	Minimum proportion of controls and treated subjects. The default is 0.25.

Details

In each randomization, the treatment label is resampled and then the EM algorithm is called. The final p-value is based on all randomizations in which the EM algorithm converged.

Value

A list containing the objects:

Name	Description
p.val.rerand	re-randomization test p-value
baseline	estimated baseline hazard from observed data
lambda	estimated hazard ratio from observed data

Author(s)

Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov>, Yongsoek Park <yongpark@pitt.edu> and Bin Zhu <bin.zhu@nih.gov>

See Also

PRIME.ReRandomizationTest

Examples

```
data(data, package="Immunotherapy.Design")
set.seed(1)
ret <- Pembedded.ReRandomizationTest.P(data)
ret$p.val.rerand
```

Pow.Pembedded.P *Power*

Description

Compute the power for the parametric model

Usage

```
Pow.Pembedded.P(nmax=500, rand_ratio=0.5, effect_p=0.6, enroll_rate=380*0.25/6,
                lambda1=0.117, HR=0.5, tau=12*5, t1=1, maxiter=100000, stopTol=1e-4,
                        alpha=0.05, num_rand=1000, nsim=10000, print=0,
                        min.sample.size=50, min.n.event=5, min.per.trt=0.25)
```

Arguments

nmax	Sample size
rand_ratio	Probability of assignment to treatment arm
effect_p	Proportion of responders in the treatment arm
enroll_rate	Enrollment rate in subjects per month
lambda1	Baseline hazard in terms of months
HR	Hazard ratio
tau	Total study duration
t1	Delayed duration in months
maxiter	Maximum number of iterations in the EM algorithm. The default is 100000.
stopTol	Stopping tolerance in the EM algorithm. The default is 1e-4.
alpha	Significance level. The default is 0.05.
num_rand	The number of replications in the re-randomization test. The default is 1000.
nsim	The number of simulations. The default is 10000.
print	0 or 1 to print information. The default is 0.
min.sample.size	Minimum sample size. The default is 50.
min.n.event	Minimum number of events. The default is 5.
min.per.trt	Minimum proportion of controls and treated subjects. The default is 0.25.

Details

For each simulation, a simulated data set is created from the generate_data function and then an estimated p-value is computed by calling Pembedded.ReRandomizationTest.P. The power is calculated as the proportion of iterations whose estimated p-value was less than or equal to alpha.

Value

A list containing the power and the number of simulated datasets used in the calculation.

Author(s)

Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov>, Yongsoek Park <yongpark@pitt.edu> and Bin Zhu <bin.zhu@nih.gov>

See Also

N.Pembedded.P

Pow.PRIME.LRT *Power*

Description

Compute the power using PRIME likelihood ratio test

Usage

```
Pow.PRIME.LRT(nmax=500, rand_ratio=0.5, effect_p=0.6, enroll_rate=380*0.25/6,
                lambda1=0.117, HR=0.5, tau=12*5, t1=1, maxiter=100000, stopTol=1e-4,
                       alpha=0.05, nsim=10000, print=0,
                       min.sample.size=50, min.n.event=5, min.per.trt=0.25)
```

Arguments

nmax	Sample size
rand_ratio	Probability of assignment to treatment arm
effect_p	Proportion of responders in the treatment arm at baseline
enroll_rate	Enrollment rate in subjects per month
lambda1	Baseline hazard in terms of months
HR	Hazard ratio of responders against controls
tau	Total study duration
t1	Delayed duration in months
maxiter	Maximum number of iterations in the EM algorithm. The default is 100000.
stopTol	Stopping tolerance in the EM algorithm. The default is 1e-4.
alpha	Significance level. The default is 0.05.
nsim	The number of simulations. The default is 10000.
print	0 or 1 to print information. The default is 0.
min.sample.size	Minimum sample size. The default is 50.
min.n.event	Minimum number of events. The default is 5.
min.per.trt	Minimum proportion of controls and treated subjects. The default is 0.25.

Details

For each simulation, a simulated data set is created from the generate_data function and then an estimated p-value is computed by calling PRIME.LRT. The power is calculated as the proportion of iterations whose estimated p-value was less than or equal to alpha.

Value

A list containing the power and the number of simulated datasets used in the calculation.

Author(s)

Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov>, Yongsoek Park <yongpark@pitt.edu> and Bin Zhu <bin.zhu@nih.gov>

See Also

Pow.Pembedded.P

Pow.PRIME.ReRandomizationTest
Power

Description

Compute the power using PRIME Re-randomization test

Usage

```
Pow.PRIME.ReRandomizationTest(nmax=500, rand_ratio=0.5, effect_p=0.6, enroll_rate=380*0.25/6,
                lambda1=0.117, HR=0.5, tau=12*5, t1=1, maxiter=100000, stopTol=1e-4,
                        alpha=0.05, num_rand=1000, nsim=10000, print=0,
                        min.sample.size=50, min.n.event=5, min.per.trt=0.25)
```

Arguments

nmax	Sample size
rand_ratio	Probability of assignment to treatment arm
effect_p	Proportion of responders in the treatment arm at baseline
enroll_rate	Enrollment rate in subjects per month
lambda1	Baseline hazard in terms of months
HR	Hazard ratio of responders against controls
tau	Total study duration
t1	Delayed duration in months
maxiter	Maximum number of iterations in the EM algorithm. The default is 100000.
stopTol	Stopping tolerance in the EM algorithm. The default is 1e-4.
alpha	Significance level. The default is 0.05.
num_rand	The number of replications in the re-randomization test. The default is 1000.
nsim	The number of simulations. The default is 1000.
print	0 or 1 to print information. The default is 0.
min.sample.size	Minimum sample size. The default is 50.
min.n.event	Minimum number of events. The default is 5.
min.per.trt	Minimum proportion of controls and treated subjects. The default is 0.25.

Details

For each simulation, a simulated data set is created from the generate_data function and then an estimated p-value is computed by calling PRIME.ReRandomizationTest. The power is calculated as the proportion of iterations whose estimated p-value was less than or equal to alpha.

Value

A list containing the power and the number of simulated datasets used in the calculation.

Author(s)

Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov>, Yongsoek Park <yongpark@pitt.edu> and Bin Zhu <bin.zhu@nih.gov>

See Also

Pow.Pembedded.P

PRIME.EM *EM algorithm*

Description

PRIME EM algorithm

Usage

```
PRIME.EM(data, time.var="X", trt.var="trt", status.var="event_status",
          effect_p=0.6, t1=1, lambda0=NULL, probResponder=NULL,
          stopTol=1e-4, maxiter=100000, print=0)
```

Arguments

data	Data frame or matrix containing a time-to-event variable (time.var), a treatment variable (trt.var), and a censoring variable (status.var).
time.var	Time-to-event variable name in data. The default is "X".
trt.var	Binary treatment variable name in data coded as 0 for controls and 1 for subjects that received treatment.
status.var	Name of the binary censoring variable in data coded as 0 for censored subjects and 1 for subjects that experienced an event.
effect_p	Proportion of responders in the treatment arm at baseline. The default is 0.6.
t1	Delayed duration. The default is 1.
lambda0	NULL or vector of initial estimates for the baseline hazards corrsponding to the ordered event times. The default is NULL and will be computed from getHazard.
probResponder	NULL or vector of initial probabilities of a subject being a responder. The default is NULL so that the initial probability is 0.5 for treated subjects and 0 for controls.
stopTol	Stopping tolerance. The default is 1e-4.

maxiter	Maximum number of iterations. The default is 100000.
print	0-2 to print information. Larger values will print more information. The default is 0.

Value

A list containing the objects:

Name	Description
converged	TRUE if EM algorithm converged
logHR	estimated log(hazard ratio) of responders versus controls
baseline	matrix of event times and baseline hazards
probResponder	estimated probability of a subject being a responder
loglike	log-likelihood value at the final estimates
loglike.marg	marginal log-likelihood value at the final estimates

Author(s)

Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov>, Yongsoek Park <yongpark@pitt.edu> and Bin Zhu <bin.zhu@nih.gov>

See Also

getHazard, Pembedded.EM.P

Examples

```
data(data, package="Immunotherapy.Design")
ret <- PRIME.EM(data)
ret$logHR
```

PRIME.LRT *Likelihood ratio test*

Description

PRIME likelihood ratio test

Usage

```
PRIME.LRT(data, time.var="X", trt.var="trt", status.var="event_status",
          effect_p=0.6, t1=1, lambda0=NULL, probResponder=NULL,
          stopTol=1e-4, maxiter=100000, print=0)
```

Arguments

data	Data frame or matrix containing a time-to-event variable (time.var), a treatment variable (trt.var), and a censoring variable (status.var).
time.var	Time-to-event variable name in data. The default is "X".
trt.var	Binary treatment variable name in data coded as 0 for controls and 1 for subjects that received treatment.

status.var	Name of the binary censoring variable in data coded as 0 for censored subjects and 1 for subjects that experienced an event.
effect_p	Proportion of responders in the treatment arm at baseline. The default is 0.6.
t1	Delayed duration. The default is 1.
lambda0	NULL or vector of initial estimates for the baseline hazards corrsponding to the ordered event times. The default is NULL and will be computed from getHazard.
probResponder	NULL or vector of initial probabilities of a subject being a responder. The default is NULL so that the initial probability is 0.5 for treated subjects and 0 for controls.
stopTol	Stopping tolerance. The default is 1e-4.
maxiter	Maximum number of iterations. The default is 100000.
print	0-2 to print information. Larger values will print more information. The default is 0.

Value

A list containing the objects:

Name	Description
p.value	1 df LRT p-value
loglike.max	log-likelihood value at the final estimates
loglike.0	log-likelihood value under the null setting where HR=1 and probability of being a responder is effect_p

Author(s)

Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov>, Yongsoek Park <yongpark@pitt.edu> and Bin Zhu <bin.zhu@nih.gov>

See Also

Pembedded.ReRandomizationTest.P

Examples

```
data(data, package="Immunotherapy.Design")
PRIME.LRT(data)
```

PRIME.ReRandomizationTest

Randomization test

Description

Compute a randomization test p-value where test statistic is calculated based on PRIME strategy.

Usage

```
PRIME.ReRandomizationTest(data, time.var="X", trt.var="trt", status.var="event_status",
       effect_p=0.6, t1=1, stopTol=1e-4, maxiter=100000, print=0, num_rand=1000,
          min.sample.size=50, min.n.event=5, min.per.trt=0.25)
```

Arguments

`data`	Data frame or matrix containing a time-to-event variable (`time.var`), a treatment variable (`trt.var`), and a censoring variable (`status.var`).
`time.var`	Observational time variable name in `data`.
`trt.var`	Name of treatment assignment indicator in `data` coded as 0 for control subjects and 1 for treated subjects.
`status.var`	Name of the binary censoring variable in `data` coded as 0 for censored subjects and 1 for subjects that experienced an event.
`effect_p`	Proportion of responders in the treatment arm at baseline. The default is 0.6.
`t1`	Delayed duration. The default is 1 (month).
`stopTol`	Stopping tolerance in the EM algorithm. The default is 1e-4.
`maxiter`	Maximum number of iterations in the EM algorithm. The default is 100000.
`print`	0-2 to print information. Larger values will print more information. The default is 0.
`num_rand`	The number of replications in the re-randomization test. The default is 1000.
`min.sample.size`	Minimum sample size. The default is 50.
`min.n.event`	Minimum number of events. The default is 5.
`min.per.trt`	Minimum proportion of controls and treated subjects. The default is 0.25.

Details

In each randomization, the treatment label is resampled and then the EM algorithm is called. The final p-value is based on all re-randomizations in which the EM algorithm converged.

Value

A list containing the objects:

Name	Description
p.val.rerand	re-randomization test p-value
baseline	matrix of event times and baseline hazards from observed data
logHR	estimated log(hazard ratio) of responders versus controls

Author(s)

Zhenzhen Xu <Zhenzhen.Xu@fda.hhs.gov>, Yongsoek Park <yongpark@pitt.edu> and Bin Zhu <bin.zhu@nih.gov>

See Also

`Pembedded.ReRandomizationTest.P`

Examples

```
data(data, package="Immunotherapy.Design")
set.seed(1)
# Will take much time to complete
#ret <- PRIME.ReRandomizationTest(data)
```

Index