

Package ‘CoRpower’

April 25, 2019

Title Power Calculations for Assessing Correlates of Risk in Clinical Efficacy Trials

Version 1.0.1

URL <https://github.com/mjuraska/CoRpower>

Description Calculates power for assessment of intermediate biomarker responses as correlates of risk in the active treatment group in clinical efficacy trials, as described in Gilbert, Janes, and Huang, Power/Sample Size Calculations for Assessing Correlates of Risk in Clinical Efficacy Trials (2016, Statistics in Medicine). The methods differ from past approaches by accounting for the level of clinical treatment efficacy overall and in biomarker response subgroups, which enables the correlates of risk results to be interpreted in terms of potential correlates of efficacy/protection. The methods also account for inter-individual variability of the observed biomarker response that is not biologically relevant (e.g., due to technical measurement error of the laboratory assay used to measure the biomarker response), which is important because power to detect a specified correlate of risk effect size is heavily affected by the biomarker's measurement error. The methods can be used for a general binary clinical endpoint model with a univariate dichotomous, trichotomous, or continuous biomarker response measured in active treatment recipients at a fixed timepoint after randomization, with either case-cohort Bernoulli sampling or case-control without-replacement sampling of the biomarker (a baseline biomarker is handled as a trivial special case). In a specified two-group trial design, the computeN() function can initially be used for calculating additional requisite design parameters pertaining to the target population of active treatment recipients observed to be at risk at the biomarker sampling timepoint. Subsequently, the power calculation employs an inverse probability weighted logistic regression model fitted by the tps() function in the 'osDesign' package. Power results as well as the relationship between the correlate of risk effect size and treatment efficacy can be visualized using various plotting functions. To link power calculations for detecting a correlate of risk and a correlate of treatment efficacy, a baseline immunogenicity predictor (BIP) can be simulated according to a specified classification rule (for dichotomous or trichotomous BIPs) or correlation with the biomarker response (for continuous BIPs), then outputted along with biomarker response data under assignment to treatment, and clinical endpoint data for both treatment and placebo groups.

BugReports <https://github.com/mjuraska/CoRpower/issues>

Depends R (>= 3.5.0)

License GPL-2

Encoding UTF-8**Imports** survival, osDesign**RoxygenNote** 6.1.1**Suggests** knitr, rmarkdown**VignetteBuilder** knitr**LazyData** true**NeedsCompilation** no**Author** Stephanie Wu [aut],
Michal Juraska [aut, cre],
Peter Gilbert [aut],
Yunda Huang [aut]**Maintainer** Michal Juraska <mjuraska@fredhutch.org>**Repository** CRAN**Date/Publication** 2019-04-25 08:00:32 UTC

R topics documented:

computeN	2
computePower	4
plotPowerCont	15
plotPowerTri	16
plotROCcurveTri	18
plotRRgradVE	19
plotVElatCont	21
Index	24

computeN	<i>Estimation of Size and Numbers of Cases and Controls in the Target Population of Active Treatment Recipients At Risk at the Biomarker Sampling Timepoint</i>
----------	---

Description

If the power calculation is done at the study design stage, the function estimates the size and numbers of cases and controls in the target population of active treatment recipients observed to be at risk at the biomarker sampling timepoint.

Usage

```
computeN(Nrand, tau, taumax, VEtauToTaumax, VE0toTau, risk0, dropoutRisk,
propCasesWithS)
```

Arguments

Nrand	the number of participants randomized to the active treatment group
tau	the biomarker sampling timepoint after randomization
taumax	the time after randomization marking the end of the follow-up period for the clinical endpoint
VEtauToTaumax	the treatment (vaccine) efficacy level between τ and τ_{max}
VE0toTau	the treatment (vaccine) efficacy between 0 and τ
risk0	the overall placebo-group endpoint risk between τ and τ_{max}
dropoutRisk	the risk of participant dropout between 0 and τ_{max}
propCasesWithS	the proportion of observed cases with a measured biomarker response

Details

The function estimates design parameters that are required as input to [computePower](#). If the power calculation is done after the follow-up was completed, the estimates are replaced by the observed counterparts for use as input parameters in [computePower](#).

The calculations include options to account for participant dropout by specifying `dropoutRisk` as well as for incomplete sample storage by specifying `propCasesWithS`.

The estimation procedure considers the standard survival analysis framework with failure and censoring times denoted by T and C , respectively, and makes the following assumptions:

1. T and C are independent.
2. $T|Z = 0$ follows an exponential distribution with rate θ_t and $C|Z = 0$ follows an exponential distribution with rate θ_c
3. $RR_{\tau-\tau_{max}} := P(T \leq \tau_{max} | T > \tau, Z = 1) / P(T \leq \tau_{max} | T > \tau, Z = 0)$ is assumed to be equal to $P(T \leq t | T > \tau, Z = 1) / P(T \leq t | T > \tau, Z = 0)$ for all $t \in (\tau, \tau_{max}]$.

Value

A list with the following components:

- N: the total estimated number of active treatment recipients observed to be at risk at τ
- nCases: the estimated number of clinical endpoint cases observed between τ and τ_{max} in the active treatment group
- nControls: the estimated number of controls observed to complete follow-up through τ_{max} endpoint-free in the active treatment group
- nCasesWithS: the estimated number of clinical endpoint cases observed between τ and τ_{max} in the active treatment group with an available biomarker response

See Also

[computePower](#)

Examples

```

Nrand = 4100
tau = 3.5
taumax = 24
VEtauToTaumax = 0.75
VE0toTau = 0.75/2
risk0 = 0.034
dropoutRisk = 0.1
propCasesWithS = 1
computeN(Nrand, tau, taumax, VEtauToTaumax, VE0toTau, risk0, dropoutRisk, propCasesWithS)

```

computePower

Power Calculations for Assessing Intermediate Biomarkers as Correlates of Risk in the Active Treatment Group in Clinical Efficacy Trials, Accounting for Biomarker's Measurement Error and Treatment Efficacy

Description

Performs a power calculation for assessing a univariate dichotomous, trichotomous, or continuous intermediate biomarker response as a correlate of risk in the active treatment group in a clinical efficacy trial, accounting for the biomarker's measurement error and treatment efficacy. The statistical methods are described in [Gilbert, Janes, and Huang (2016). "Power/Sample Size Calculations for Assessing Correlates of Risk in Clinical Efficacy Trials."] Simulated data sets, extended to include placebo group and baseline immunogenicity predictor data, can be exported for harmonized assessment of biomarker-specific treatment efficacy.

Usage

```

computePower(nCasesTx, nControlsTx, nCasesTxWithS,
  controlCaseRatio = NULL, VEOverall, risk0, VELat0 = seq(0, VEOverall,
  len = 20), VELat1 = rep(VEOverall, 20), VELowest = NULL,
  Plat0 = NULL, Plat2 = NULL, P0 = Plat0, P2 = Plat2,
  PlatVELowest = NULL, sens = NULL, spec = NULL, FP0 = NULL,
  FN2 = NULL, M = 100, alpha = 0.05, sigma2obs = 1, rho = 1,
  biomType = c("continuous", "trichotomous", "dichotomous"),
  cohort = FALSE, p = NULL, tpsMethod = c("PL", "ML", "WL"),
  saveDir = NULL, saveFile = "CoRpower.RData", saveDataDir = NULL,
  saveDataFile = "fullData.RData", corr = NULL, nCasesPla = NULL,
  nControlsPla = NULL, sensBIP = NULL, specBIP = NULL,
  FP0BIP = NULL, FN2BIP = NULL, P0BIP = P0, P2BIP = P2)

```

Arguments

nCasesTx an integer vector specifying the number of clinical endpoint cases observed (or projected) between τ and τ_{max} in the active treatment group. Each value represents a distinct scenario for power assessment.

nControlsTx	an integer vector specifying the number of controls observed (or projected) to complete follow-up through τ_{max} endpoint-free in the active treatment group. Each value represents a distinct scenario for power assessment. The ordering in nCasesTx and nControlsTx must match.
nCasesTxWithS	an integer vector specifying the number of clinical endpoint cases observed (or projected) between τ and τ_{max} in the active treatment group with an available biomarker response. Each value represents a distinct scenario for power assessment. The ordering must match nCasesTx and nControlsTx.
controlCaseRatio	an integer vector specifying the number of controls sampled per case for biomarker measurement in the without replacement case-control sampling design (set to NULL by default). Each value represents a distinct scenario for power assessment.
VEoverall	a numeric value specifying the overall treatment (vaccine) efficacy between τ and τ_{max}
risk0	a numeric value specifying the overall placebo-group endpoint risk between τ and τ_{max}
VElat0	a numeric vector specifying a grid of treatment (vaccine) efficacy levels in the latent lower protected subgroup for a dichotomous or trichotomous biomarker. Each value of VElat0 corresponds to one unique effect size (RR_t). Default ranges from VEoverall (H_0) to 0 (maximal H_1 not allowing harm by treatment).
VElat1	a numeric vector specifying a grid of treatment (vaccine) efficacy levels in the latent medium protected subgroup for a trichotomous biomarker. Each value corresponds to one unique effect size (RR_t). The ordering must match VElat0. Set to VEoverall by default, but must be set to NULL for a dichotomous biomarker.
VElowest	a numeric vector specifying a grid of treatment (vaccine) efficacy levels in the latent lowest-efficacy subgroup for a continuous biomarker. Default ranges from VEoverall (H_0) to 0 (maximal H_1 not allowing harm by treatment).
Plat0	a numeric vector specifying the prevalence of the latent lower protected subgroup for a dichotomous or trichotomous biomarker (set to NULL by default). Each value represents a distinct scenario for power assessment. The ordering in Plat0, Plat2, P0, and P2 must match.
Plat2	a numeric vector specifying the prevalence of the latent higher protected subgroup for a dichotomous or trichotomous biomarker (set to NULL by default). Each value represents a distinct scenario for power assessment. The ordering in Plat0, Plat2, P0, and P2 must match.
P0	a numeric vector specifying the probability of low biomarker response for a dichotomous or trichotomous biomarker (set to Plat0 by default). Each value represents a distinct scenario for power assessment. The ordering in Plat0, Plat2, P0, and P2 must match.
P2	a numeric vector specifying the probability of high biomarker response for a dichotomous or trichotomous biomarker (set to Plat2 by default). Each value represents a distinct scenario for power assessment. The ordering in Plat0, Plat2, P0, and P2 must match.

PlatVElowest	a numeric vector specifying the prevalence of the latent lowest-efficacy subgroup for a continuous biomarker (set to NULL by default). Each value represents a distinct scenario for power assessment.
sens	a numeric vector specifying the sensitivity, i.e., the probability of high biomarker response conditional on membership in the higher protected subgroup, for a dichotomous or trichotomous biomarker. Default is NULL, which indicates the use of 'approach 2'. Each value represents a distinct scenario for power assessment. The ordering in sens, spec, FP0, and FN2 must match.
spec	a numeric vector specifying the specificity, i.e., the probability of low biomarker response conditional on membership in the lower protected subgroup, of a dichotomous or trichotomous biomarker. Default is NULL, which indicates the use of 'approach 2'. Each value represents a distinct scenario for power assessment. The ordering in sens, spec, FP0, and FN2 must match.
FP0	a numeric vector specifying the false positive rate, i.e., the probability of high biomarker response conditional on membership in the lower protected subgroup, for a dichotomous or trichotomous biomarker. Default is NULL, which indicates the use of 'approach 2'. Each value represents a distinct scenario for power assessment. The ordering in sens, spec, FP0, and FN2 must match.
FN2	a numeric vector specifying the false negative rate, i.e., the probability of low biomarker response conditional on membership in the higher protected subgroup, for a dichotomous or trichotomous biomarker. Default is NULL, which indicates the use of 'approach 2'. Each value represents a distinct scenario for power assessment. The ordering in sens, spec, FP0, and FN2 must match.
M	an integer value specifying the number of simulated clinical trials. Default is 100.
alpha	a numeric value specifying the two-sided Wald test type-I error rate. Default is 0.05.
sigma2obs	a numeric value specifying the variance of the observed continuous biomarker or of the dichotomous or trichotomous biomarker simulated using 'approach 2' (set to 1 by default).
rho	a numeric vector specifying distinct protection-relevant fractions of sigma2obs. Each value represents a distinct scenario for power assessment.
biomType	a character string specifying the biomarker type. Default is continuous; other choices are dichotomous and trichotomous.
cohort	a logical value for whether a case-cohort Bernoulli sampling design is to be used. If FALSE (default), the case-control without replacement sampling is used.
p	a numeric vector specifying the probability of sampling into the subcohort in the case-cohort design (NULL by default). Each value represents a distinct scenario for power assessment.
tpsMethod	a character string specifying the estimation method in the inverse probability weighted logistic regression model fit by the tps function in the osDesign package. The options are PL for pseudo-likelihood (default), ML for maximum likelihood, and WL for weighted likelihood.
saveDir	a character string specifying the path for a directory in which the output of the power calculation is to be saved. If NULL (default), the output is returned only.

saveFile	a character vector specifying the name(s) of the .RData file(s) storing the output of the power calculation, used only if saveDir is not NULL. All file names must include ".RData" at the end. Default is CoRpower.RData.
saveDataDir	a character string specifying the path for a directory in which the simulated data, including placebo group and baseline immunogenicity predictor (BIP) data, are to be saved. If NULL (default), the simulated data are not saved.
saveDataFile	a character vector specifying the name(s) of the .RData file(s) in which the simulated data, including placebo group and BIP data, are to be saved; used only if saveDataDir is not NULL. All file names must include ".RData" at the end. Default is fullData.RData.
corr	a numeric vector in $[-1, 1]$ specifying the correlation between a continuous baseline immunogenicity predictor (BIP) and the (underlying) continuous intermediate biomarker response (NULL by default). Each value represents a distinct scenario for power assessment. A useful BIP is highly correlated with the biomarker response at τ . It must be provided if saveDataDir is specified and a trichotomous biomarker under 'approach 2' or a continuous biomarker is considered.
nCasesPla	an integer vector specifying the number of clinical endpoint cases observed (or projected) between τ and τ_{max} in the placebo group. Each value represents a distinct scenario matching nCasesTx. Default is NULL. It must be provided if saveDataDir is specified.
nControlsPla	an integer vector specifying the number of controls observed (or projected) to complete follow-up through τ_{max} endpoint-free in the placebo group. Each value represents a distinct scenario matching nControlsTx. Default is NULL. It must be provided if saveDataDir is specified.
sensBIP	a numeric vector specifying "the sensitivity" of a dichotomous or trichotomous BIP, i.e., the probability of a high value of the BIP conditional on high biomarker response. Default is NULL, which indicates the use of 'approach 2'. It must be provided if saveDataDir is specified and 'approach 1' is to be used. Each value results in generating a separate BIP variable in the output data.
specBIP	a numeric vector specifying "the specificity" of a dichotomous or trichotomous BIP, i.e., the probability of a low value of the BIP conditional on low biomarker response. Default is NULL, which indicates the use of 'approach 2'. It must be provided if saveDataDir is specified and 'approach 1' is to be used. Each value results in generating a separate BIP variable in the output data.
FP0BIP	a numeric vector specifying "the false positive rate" of a dichotomous or trichotomous BIP, i.e., the probability of a high value of the BIP conditional on low biomarker response. Default is NULL, which indicates the use of 'approach 2'. It must be provided if saveDataDir is specified and 'approach 1' is to be used. Each value results in generating a separate BIP variable in the output data.
FN2BIP	a numeric vector specifying "the false negative rate" of a dichotomous or trichotomous BIP, i.e., the probability of a low value of the BIP conditional on high biomarker response. Default is NULL, which indicates the use of 'approach 2'. It must be provided if saveDataDir is specified and 'approach 1' is to be used. Each value results in generating a separate BIP variable in the output data.

P0BIP	a numeric vector specifying the probability of a low value of a dichotomous or trichotomous BIP. If unspecified, it is set to P0. Each value results in generating a separate BIP variable in the output data.
P2BIP	a numeric vector specifying the probability of a high value of a dichotomous or trichotomous BIP. If unspecified, it is set to P2. Each value results in generating a separate BIP variable in the output data.

Details

A number of calling arguments can be specified as vectors with each component specifying a distinct scenario for power assessment (saved in a separate .RData file). These are referred to as "varying arguments." Some varying arguments occur in a group, where the length and order of all specified vectors in the group must match; others are the only varying argument in their group. Only arguments belonging to a single group may be varied at a time; if two or more groups contain vector inputs, the function will treat such inputs as an error. The following are the groups of varying arguments that can be vectorized:

- nCasesTx, nControlsTx, and nCasesTxWithS (together with nCasesPla and nControlsPla if simulated data sets are to be saved)
- Plat0, Plat2, P0, and P2
- sens, spec, FP0, and FN2
- controlCaseRatio
- rho
- p

Arguments independent of biomarker type and sampling design: nCasesTx, nControlsTx, nCasesTxWithS, VEoverall, risk0, M, alpha, tpsMethod, saveDir, saveFile.

Arguments specific to a trichotomous (or dichotomous) biomarker response: VElat0, VElat1, Plat0, Plat2, P0, P2, biomType = "trichotomous" (or "dichotomous")

- Arguments for Approach 1: sens, spec, FP0, FN2
- Arguments for Approach 2: sigma2obs, rho

Arguments specific to a continuous biomarker response: VElowest, PlatVElowest, sigma2obs, rho, biomType = "continuous"

Arguments for a case-control without replacement sampling design: controlCaseRatio

Arguments for a case-cohort Bernoulli sampling design: cohort = TRUE, p

To save output from the power calculations in an .RData file, saveDir must be specified. The default file name is CoRpower.RData; a different file name may be specified by saveFile as a single character string, to which the value of the varying argument(s) will be appended for descriptive file naming purposes, or, alternatively, a character vector may be specified with full file names (a single file will be produced for each value of the varying argument(s)).

To link power calculations for detecting a correlate of risk and a correlate of treatment efficacy, simulated data sets used in the power calculations can be exported with placebo-group data, with a possible extension including BIP data, for harmonized use by methods assessing biomarker-specific treatment efficacy. The vignette "Algorithms for Simulating Placebo Group and Baseline Immuno-genicity Predictor Data" provides more information on the algorithms and underlying assumptions

for simulating placebo-group and BIP data. The exported data sets include treatment and placebo group data in the form of full rectangular data (i.e., disregarding biomarker sub-sampling), which enables the user to employ any preferred biomarker sub-sampling design. To generate and export such data, `saveDataDir`, `nCasesPla`, and `nControlsPla` must be specified. `nCasesPla` and `nControlsPla` must have the same length and order of components as `nCasesTx`, `nControlsTx`, and `nCasesTxWithS`.

If a BIP is to be included in the simulated data export, additional arguments are necessary. If the biomarker is trichotomous and Approach 1 is used, `sensBIP`, `specBIP`, `FP0BIP`, `FN2BIP`, `P0BIP`, and `P2BIP` must be specified; if the biomarker is trichotomous and Approach 2 is used, `corr`, `P0BIP`, and `P2BIP` must be specified; if the biomarker is continuous, `corr` must be specified.

Calling arguments pertaining to the simulation of the BIP in the exported data may also be specified as vectors, independently of the above varying arguments defining the power calculation scenarios for the active treatment group. Each component of these vectors results in the generation of a separate BIP variable, in the same order, in the output data. Some of these arguments occur in a group, where the length and order of all specified vectors in the group must match; others are the sole argument in their group. Only arguments belonging to a single group may be varied at a time; if two or more groups contain vector inputs, the function will treat such inputs as an error. The following are the groups of BIP arguments that can be vectorized:

- `sensBIP`, `specBIP`, `FP0BIP`, `FN2BIP`
- `P0BIP`, `P2BIP`
- `corr`

The default file name for the outputted data sets is `fullData.RData`. A different file name may be specified by `saveDataFile` as a single character string, to which the value of the "varying argument" for the power calculations will be appended for descriptive file naming purposes, or, alternatively, a character vector may be specified with full file names (a single file will be produced for each value of the varying argument(s)). Note: if the "varying argument" is `controlCaseRatio` or `p`, only one file will be generated because these arguments do not affect the simulation of the full data; therefore, `saveDataFile` must be a character string in these cases.

Value

If `saveDir` is specified, an output list (named `pwr`) for each power scenario is saved as an `.RData` file. Otherwise, the function returns a list of lists, where the outer list ranges over specified values of the varying argument(s) whose components denote distinct scenarios, and the inner list is the output list for each power scenario. For a dichotomous or trichotomous biomarker, each output list has the following components:

- `power`: a numeric vector of fractions of simulated trials in which the null hypothesis H_0 is rejected. Each value of the vector corresponds to a value in the grid of treatment (vaccine) efficacies specified by `VElat0` and `VElat1`.
- `RRt`: a numeric vector of correlate-of-risk relative-risk effect sizes. Each value of the vector corresponds to a value in the grid of treatment (vaccine) efficacies specified by `VElat0` and `VElat1`.
- `risk1_2`: a numeric vector of conditional endpoint risks given a high biomarker response in the active treatment group. Each value of the vector corresponds to a value in the grid of treatment (vaccine) efficacies specified by `VElat0` and `VElat1`.

- risk1_0: a numeric vector of conditional endpoint risks given a low biomarker response in the active treatment group. Each value of the vector corresponds to a value in the grid of treatment (vaccine) efficacies specified by VElat0 and VElat1.
- VElat2: a numeric vector specifying a grid of treatment (vaccine) efficacy levels in the latent higher protected subgroup for a dichotomous or trichotomous biomarker
- VElat0: a numeric vector specifying a grid of treatment (vaccine) efficacy levels in the latent lower protected subgroup for a dichotomous or trichotomous biomarker
- Plat2: a numeric value specifying the prevalence of the latent higher protected subgroup for a dichotomous or trichotomous biomarker
- Plat0: a numeric value specifying the prevalence of the latent lower protected subgroup for a dichotomous or trichotomous biomarker
- P2: a numeric value specifying the probability of high biomarker response for a dichotomous or trichotomous biomarker
- P0: a numeric value specifying the probability of low biomarker response for a dichotomous or trichotomous biomarker
- alphaLat: a numeric vector of the log odds of the clinical endpoint in the subgroup of active treatment recipients with the latent $x^* = 0$ (this coefficient estimate applies to a continuous biomarker)
- betaLat: a numeric vector of the log odds ratio of the clinical endpoint comparing two subgroups of active treatment recipients differing in the latent x^* by 1 (this coefficient estimate applies to a continuous biomarker)
- sens: a numeric vector of sensitivities (i.e., the probability of high biomarker response conditional on membership in the higher protected subgroup) of the observed dichotomous or trichotomous biomarker as a function of rho
- spec: a numeric vector of specificities (i.e., the probability of low biomarker response conditional on membership in the lower protected subgroup) of the observed dichotomous or trichotomous biomarker as a function of rho
- FP0: a numeric vector of false positive rates (i.e., the probability of high biomarker response conditional on membership in the lower protected subgroup) of the observed dichotomous or trichotomous biomarker as a function of rho
- FN2: a numeric vector of false negative rates (i.e., the probability of low biomarker response conditional on membership in the higher protected subgroup) of the observed dichotomous or trichotomous biomarker as a function of rho
- NcompleteTx: an integer value specifying $nCasesTx + nControlsTx$, i.e., the number, observed or projected, of active treatment recipients at risk at τ with an observed endpoint or a completed follow-up through τ_{max}
- nCasesTx: an integer value specifying the number of clinical endpoint cases observed (or projected) between τ and τ_{max} in the active treatment group
- nCasesTxWithS: an integer value specifying the number of clinical endpoint cases observed (or projected) between τ and τ_{max} in the active treatment group with an available biomarker response
- controlCaseRatio: an integer specifying the number of controls sampled per case for biomarker measurement in the without replacement case-control sampling design

- VEoverall: a numeric value specifying the overall treatment (vaccine) efficacy between τ and τ_{max}
- risk0: a numeric value specifying the overall placebo-group endpoint risk between τ and τ_{max}
- alpha: a numeric value specifying the two-sided Wald test type-I error rate
- rho: a numeric vector specifying distinct protection-relevant fractions of the variance of the observed biomarker

For a continuous biomarker, each output list has the following components:

- power: a numeric vector of fractions of simulated trials in which the null hypothesis H_0 is rejected. Rows represent calculations for different values of rho or nCasesTx, depending on which is a vector. Columns represent calculations for the grid of treatment (vaccine) efficacy levels in the latent lowest-efficacy subgroup, specified by VElowest.
- RRc: a numeric vector of correlate-or-risk relative-risk effect sizes as a function of the grid of treatment (vaccine) efficacy levels in the latent lowest-efficacy subgroup, specified by VElowest
- betaLat: a numeric vector specifying the log odds ratio of the clinical endpoint comparing two subgroups of active treatment recipients differing in the latent x^* by 1 (this coefficient estimate applies to a continuous biomarker)
- alphaLat: a numeric vector specifying the the log odds of the clinical endpoint in the subgroup of active treatment recipients with the latent $x^* = 0$ (this coefficient estimate applies to a continuous biomarker)
- PlatVElowest: a numeric value specifying the prevalence of the latent lowest-efficacy subgroup for a continuous biomarker
- VElowest: a numeric vector specifying a grid of treatment (vaccine) efficacy levels in the latent lowest-efficacy subgroup for a continuous biomarker
- sigma2obs: a numeric value specifying the variance of the observed continuous biomarker or of the dichotomous or trichotomous biomarker simulated using 'approach 2'
- NcompleteTx: an integer value specifying nCasesTx + nControlsTx, i.e., the number, observed or projected, of active treatment recipients at risk at τ with an observed endpoint or a completed follow-up through τ_{max}
- nCasesTx: an integer value specifying the number of clinical endpoint cases observed (or projected) between τ and τ_{max} in the active treatment group
- nCasesTxWithS: an integer value specifying the number of clinical endpoint cases observed (or projected) between τ and τ_{max} in the active treatment group with an available biomarker response
- controlCaseRatio: an integer value specifying the number of controls sampled per case for biomarker measurement in the without replacement case-control sampling design
- VEoverall: a numeric value specifying the overall treatment (vaccine) efficacy between τ and τ_{max}
- risk0: a numeric value specifying the overall placebo-group endpoint risk between τ and τ_{max}
- alpha: a numeric value specifying the two-sided Wald test type-I error rate

- rho: a numeric vector specifying distinct protection-relevant fractions of the variance of the observed biomarker

If saveDataDir is specified, the simulated data, including placebo group and BIP data, are saved in one or more .RData file(s) containing a list of lists of data frames. The components of the outer list consist each of one Monte-Carlo iteration of simulated data for all values of VElat0 or VElat1 if the biomarker is trichotomous, or of VElowest if the biomarker is continuous. Each data frame corresponds to one simulated trial.

See Also

[computeN](#), [plotPowerTri](#), [plotPowerCont](#)

Examples

```
## Trichotomous biomarker, Approach 1, varying sens and spec ##
## Specify sens, spec, FP0, FN2
nCasesTx <- 32
nControlsTx <- 1000
nCasesTxWithS <- 32
controlCaseRatio <- 5
VEoverall <- 0.75
risk0 <- 0.034
VElat0 <- seq(0, VEoverall, len=20) # 20 data points for the power curve
VElat1 <- rep(VEoverall, 20)
Plat0 <- 0.2
Plat2 <- 0.6
P0 <- Plat0 # different values of P0 can be set
P2 <- Plat2 # different values of P2 can be set
sens <- spec <- c(1, 0.9, 0.8, 0.7)
FP0 <- FN2 <- rep(0, 4)
M <- 5
alpha <- 0.05
biomType <- "trichotomous"
computePower(nCasesTx=nCasesTx, nControlsTx=nControlsTx, nCasesTxWithS=nCasesTxWithS,
             controlCaseRatio=controlCaseRatio, VEoverall=VEoverall,
             risk0=risk0, VElat0=VElat0, VElat1=VElat1, Plat0=Plat0,
             Plat2=Plat2, P0=P0, P2=P2, M=M, alpha=alpha, spec=spec,
             FP0=FP0, sens=sens, FN2=FN2, biomType=biomType)

## Not run:
## Trichotomous biomarker, Approach 2, varying rho ##
## Saving simulated data (including placebo and BIP data)
## Specify rho, sigma2obs, saveDataDir, saveDataFile, corr

nCasesTx <- 32
nControlsTx <- 1000
nCasesTxWithS <- 32
controlCaseRatio <- 5
VEoverall <- 0.75
risk0 <- 0.034
```

```

VElat0 <- seq(0, VEOverall, len=20)
VElat1 <- rep(VEOverall, 20)
Plat0 <- 0.2
Plat2 <- 0.6
P0 <- Plat0
P2 <- Plat2
M <- 5
alpha <- 0.05
sigma2obs <- 1
rho <- c(1, 0.9, 0.7, 0.5)
biomType <- "trichotomous"
saveDataDir <- "~/myDir"
saveDataFile <- "myDataFile.RData"
corr <- 0.7
computePower(nCasesTx=nCasesTx, nControlsTx=nControlsTx, nCasesTxWithS=nCasesTxWithS,
             controlCaseRatio=controlCaseRatio, VEOverall=VEOverall, risk0=risk0,
             VElat0=VElat0, VElat1=VElat1, Plat0=Plat0, Plat2=Plat2, P0=P0, P2=P2,
             M=M, alpha=alpha, sigma2obs=sigma2obs, rho=rho, biomType=biomType,
             saveDataDir=saveDataDir, saveDataFile=saveDataFile, corr=corr)

## dichotomous biomarker, Approach 2, varying rho ##
## Plat0 + Plat2 = 1

nCasesTx <- 32
nControlsTx <- 1000
nCasesTxWithS <- 32
controlCaseRatio <- 5
VEOverall <- 0.75
risk0 <- 0.034
VElat0 <- seq(0, VEOverall, len=20) # 20 data points for the power curve
VElat1 <- rep(0, 20) # will not be used by function
Plat0 <- 0.2
Plat2 <- 1 - Plat0
P0 <- Plat0
P2 <- Plat2
M <- 5
alpha <- 0.05
sigma2obs <- 1
rho <- c(1, 0.9, 0.7, 0.5)
biomType <- "dichotomous"
computePower(nCasesTx=nCasesTx, nControlsTx=nControlsTx, nCasesTxWithS=nCasesTxWithS,
             controlCaseRatio=controlCaseRatio, VEOverall=VEOverall, risk0=risk0,
             VElat0=VElat0, VElat1=VElat1, Plat0=Plat0, Plat2=Plat2, P0=P0, P2=P2,
             M=M, alpha=alpha, sigma2obs=sigma2obs, rho=rho, biomType=biomType)

## Continuous biomarker, varying rho ##

nCasesTx <- 32
nControlsTx <- 1000
nCasesTxWithS <- 32
controlCaseRatio <- 5

```

```

VEoverall <- 0.75
risk0 <- 0.034
PlatVElowest <- 0.2
VElowest <- seq(0, VEoverall, len=20)
M <- 5
alpha <- 0.05
sigma2obs <- 1
rho <- c(1, 0.9, 0.7, 0.5)
biomType <- "continuous"
computePower(nCasesTx=nCasesTx, nControlsTx=nControlsTx, nCasesTxWithS=nCasesTxWithS,
             controlCaseRatio=controlCaseRatio, VEoverall=VEoverall, risk0=risk0,
             PlatVElowest=PlatVElowest, VElowest=VElowest, M=M, alpha=alpha,
             sigma2obs=sigma2obs, rho=rho, biomType=biomType)

## Continuous biomarker, case-cohort sampling design, varying p ##
nCasesTx <- 32
nControlsTx <- 1000
nCasesTxWithS <- 32
VEoverall <- 0.75
risk0 <- 0.034
PlatVElowest <- 0.2
VElowest <- seq(0, VEoverall, len=20)
M <- 5
alpha <- 0.05
sigma2obs <- 1
rho <- 0.9
biomType <- "continuous"
cohort <- TRUE
p <- c(0.01, 0.02, 0.03)
computePower(nCasesTx=nCasesTx, nControlsTx=nControlsTx, nCasesTxWithS=nCasesTxWithS,
             VEoverall=VEoverall, risk0=risk0, PlatVElowest=PlatVElowest,
             VElowest=VElowest, M=M, alpha=alpha, sigma2obs=sigma2obs,
             rho=rho, biomType=biomType, cohort=cohort, p=p)

## Continuous biomarker, saving output, varying sample sizes ##
nCasesTx <- 32
nControlsTx <- 1000
nCasesTxWithS <- 32
controlCaseRatio <- 5
VEoverall <- 0.75
risk0 <- 0.034
PlatVElowest <- 0.2
VElowest <- seq(0, VEoverall, len=20)
M <- 5
alpha <- 0.05
sigma2obs <- 1
rho <- c(1, 0.9, 0.7, 0.5)
biomType <- "continuous"
saveDir <- "~/myDir"
saveFile <- "MyFile.RData"
computePower(nCasesTx=nCasesTx, nCasesTxWithS=nCasesTxWithS, nControlsTx=nControlsTx,

```

```

controlCaseRatio=controlCaseRatio, VEOverall=VEOverall,
risk0=risk0, PlatVElowest=PlatVElowest, VElowest=VElowest,
M=M, alpha=alpha, sigma2obs=sigma2obs, rho=rho,
biomType=biomType, saveDir=saveDir, saveFile=saveFile)

## End(Not run)

```

plotPowerCont	<i>Plotting of Power Curve versus Correlate of Risk Effect Size for Continuous Biomarkers</i>
---------------	---

Description

Plots power (on the y-axis) to detect a correlate of risk effect size (on the x-axis) in the active treatment group for a continuous biomarker. The correlate of risk effect size is quantified as the relative risk ratio of the clinical endpoint per standard deviation increase for a noise-free biomarker.

Usage

```
plotPowerCont(outComputePower, outDir = NULL, legendText)
```

Arguments

outComputePower	either a list of lists containing output from computePower or a character vector specifying the .RData file(s) containing computePower output
outDir	a character vector specifying path(s) to output .RData file(s), necessary if outComputePower is a character vector. Default is NULL.
legendText	a character vector specifying the entirety of the legend text. The order of the elements (i.e., parameter values) must match that of the computePower input parameters in order for legend labels to be accurate.

Details

The function's plot can be interpreted in conjunction with the output of [plotVElatCont](#) by matching the CoR relative risk in the two plots and examining power compared to treatment (vaccine) efficacy. This sheds light on the importance of overall vaccine efficacy on power and allows correlates of risk results to be interpreted in terms of potential correlates of efficacy/protection.

Value

None. The function is called solely for plot generation.

See Also

[computePower](#), [plotVElatCont](#), [plotPowerTri](#)

Examples

```

# Example scenario with continuous biomarker, where values of rho are varied

# Set input parameters for computePower function
nCasesTx <- 10
nControlsTx <- 300
nCasesTxWithS <- 10
controlCaseRatio <- 5
VEoverall <- 0.75
risk0 <- 0.034
PlatVElowest <- 0.2
VElowest <- seq(0, VEoverall, len=5)
Plat0 <- P0 <- 0.2
Plat2 <- P2 <- 0.6
M <- 22
alpha <- 0.05
sigma2obs <- 1
rho <- c(1, 0.7, 0.4)
biomType <- "continuous"

# Output from computePower function is stored in an object as a list of lists
pwr <- computePower(nCasesTx=nCasesTx, nCasesTxWithS=nCasesTxWithS, nControlsTx=nControlsTx,
                    controlCaseRatio=controlCaseRatio, risk0=risk0, VEoverall=VEoverall,
                    PlatVElowest=PlatVElowest, VElowest=VElowest,
                    Plat0=Plat0, Plat2=Plat2, P0=P0, P2=P2, M=M, alpha=alpha,
                    sigma2obs=sigma2obs, rho=rho, biomType=biomType)

# Set parameters for plotPowerCont function
# outComputePower is a list of lists containing output from the computePower function
outComputePower <- pwr
legendText <- paste0("rho = ", c(1, 0.7, 0.4))
plotPowerCont(outComputePower=outComputePower, legendText=legendText)

## Not run:
# Output from computePower function is saved in RData files
computePower(..., saveDir = "myDir", saveFile = "myFile.RData")
# outComputePower is a character string specifying the file containing the
# computePower output
# outDir is a character string specifying the outComputePower file directory
outComputePower <- paste0("myFile_rho_", c(1, 0.7, 0.4), ".RData")
outDir <- "~/myDir"
legendText <- paste0("rho = ", c(1, 0.7, 0.4))
plotPowerCont(outComputePower, outDir=outDir, legendText = legendText)

## End(Not run)

```


Description

Plots power (on the y-axis) to detect a correlate of risk effect size (on the x-axis) in the active treatment group for a dichotomous or trichotomous biomarker. The correlate of risk effect size is quantified as the relative risk ratio of the clinical endpoint comparing subgroups of active treatment recipients with high and low biomarker response.

Usage

```
plotPowerTri(outComputePower, outDir = NULL, legendText)
```

Arguments

outComputePower	either a list of lists containing output from <code>computePower</code> or a character vector specifying the .RData file(s) containing <code>computePower</code> output
outDir	a character vector specifying path(s) to output .RData file(s), necessary if outComputePower is a character vector. Default is NULL.
legendText	a character vector specifying the entirety of the legend text. The order of the elements (i.e., parameter values) must match that of the <code>computePower</code> input parameters in order for legend labels to be accurate.

Details

If multiple levels are specified for the biomarker measurement error input parameters (i.e., for sens/spec or rho) in `computePower`, only the first level is used to determine the RR_t values that are plotted on the x-axis.

Value

None. The function is called solely for plot generation.

See Also

[computePower](#), [plotPowerCont](#)

Examples

```
# Example scenario with trichotomous biomarker, where values of controlCaseRatio are varied

# Set input parameters for computePower function
nCasesTx <- 10
nControlsTx <- 300
nCasesTxWithS <- 10
controlCaseRatio <- c(5,3)
VEoverall <- 0.75
risk0 <- 0.034
VELat0 <- seq(0, VEoverall, len=5)
VELat1 <- rep(VEoverall, 5)
Plat0 <- P0 <- 0.2
Plat2 <- P2 <- 0.6
```

```

sens <- spec <- 0.8
FP0 <- FN2 <- 0
M <- 50
alpha <- 0.05
biomType <- "trichotomous"

# Output from computePower function is stored in an object as a list of lists
pwr <- computePower(nCasesTx=nCasesTx, nControlsTx=nControlsTx, nCasesTxWithS=nCasesTxWithS,
                    controlCaseRatio=controlCaseRatio, risk0=risk0,
                    VOverall=VOverall, Plat0=Plat0, Plat2=Plat2, P0=P0, P2=P2,
                    VElat0=VElat0, VElat1=VElat1, M=M, alpha=alpha, spec=spec,
                    FP0=FP0, sens=sens, FN2=FN2, biomType=biomType)

# Set parameters for plotPowerTri function
# outComputePower is a list of lists containing outputs from the computePower function
outComputePower <- pwr
legendText <- paste0("controls:cases = ", c("5:1","3:1"))
plotPowerTri(outComputePower=outComputePower, legendText=legendText)

## Not run:
# outComputePower is a character vector specifying the files containing computePower output
# outDir is a character vector specifying the outComputePower file directories
outComputePower <- paste0("myFile_controlCaseRatio_", c(5, 3), ".RData")
outDir <- rep("~/myDir", 2)
legendText <- paste0("controls:cases = ", c("5:1","3:1"))
plotPowerTri(outComputePower, outDir=outDir, legendText = legendText)

## End(Not run)

```

plotROCcurveTri

Plotting of ROC Curves for Trichotomous Biomarkers

Description

Plots the receiver operating characteristic (ROC) curve displaying sensitivity and specificity for a range of P2 and P0 values, four values of rho, and four values of Plat2. Illustrates how different levels of measurement error rho map to sensitivity and specificity, depending on the value of Plat2. This function is used to create Figure 1 in the Supplementary Material of [Gilbert, Janes, and Huang (2016). "Power/Sample Size Calculations for Assessing Correlates of Risk in Clinical Efficacy Trials."]

Usage

```
plotROCcurveTri(Plat0, Plat2, P0, P2, rho)
```

Arguments

Plat0 a numeric value specifying the prevalence of the latent lower protected subgroup for a dichotomous or trichotomous biomarker

Plat2	a numeric vector of length four specifying the prevalences of the latent higher protected subgroup for a dichotomous or trichotomous biomarker
P0	a numeric vector specifying a grid of probabilities of low biomarker response for a dichotomous or trichotomous biomarker.
P2	a numeric vector specifying a grid of probabilities of high biomarker response for a dichotomous or trichotomous biomarker.
rho	a numeric vector of length four specifying distinct protection-relevant fractions of sigma2obs.

Value

None. The function is called solely for plot generation.

Examples

```
Plat0 <- 0.2
Plat2 <- c(0.2, 0.3, 0.4, 0.5)
P0 <- seq(0.90, 0.10, len=10)
P2 <- seq(0.10, 0.90, len=10)
rho <- c(1, 0.9, 0.7, 0.5)
plotROccurveTri(Plat0 = Plat0, Plat2 = Plat2, P0 = P0, P2 = P2, rho = rho)
```

plotRRgradVE	<i>Plotting of the Ratio of Relative Risks for Higher/Lower Latent Subgroups against Correlate of Risk Effect Size for Trichotomous Biomarkers</i>
--------------	--

Description

Plots the ratio of relative risks for the higher and lower latent subgroups (on the y-axis) versus the correlate of risk effect size (on the x-axis) in the active treatment group for a trichotomous biomarker. The correlate of risk effect size is quantified as the relative risk ratio of the clinical endpoint comparing subgroups of active treatment recipients with high and low biomarker response.

Usage

```
plotRRgradVE(outComputePower, outDir = NULL, legendText)
```

Arguments

outComputePower	either a list of lists containing output from <code>computePower</code> or a character vector specifying the .RData file(s) containing <code>computePower</code> output
outDir	a character vector specifying path(s) to output .RData file(s), necessary if outComputePower is a character vector. Default is NULL.
legendText	a character vector specifying the entirety of the legend text. The order of the elements (i.e., parameter values) must match that of the <code>computePower</code> input parameters in order for legend labels to be accurate.

Details

When ρ is varied, this plot shows how the relationship between the correlate of risk effect size and the relative risks for the higher and lower latent subgroups changes for different values of ρ . The ratio of relative risks for the higher and lower latent subgroups is a relative vaccine efficacy parameter. When $\rho=1$, a correlate of risk in the vaccine group is equivalent to the relative vaccine efficacy parameter, whereas for imperfectly measured biomarkers with $\rho<1$, the correlate of risk effect size is closer to the null than the relative vaccine efficacy parameter is.

Value

None. The function is called solely for plot generation.

See Also

[computePower](#), [plotPowerTri](#)

Examples

```
# Example scenario with trichotomous biomarker, where values of rho are varied

# Set input parameters for computePower function
nCasesTx <- 10
nControlsTx <- 300
nCasesTxWithS <- 10
controlCaseRatio <- 3
VEoverall <- 0.75
risk0 <- 0.034
VELat0 <- seq(0, VEoverall, len=10)
VELat1 <- rep(VEoverall, 10)
Plat0 <- P0 <- 0.2
Plat2 <- P2 <- 0.6
M <- 20
alpha <- 0.05
sigma2obs <- 1
rho <- c(1, 0.7, 0.4)
biomType <- "trichotomous"

# Output from computePower function is stored in an object as a list
pwr <- computePower(nCasesTx=nCasesTx, nControlsTx=nControlsTx, nCasesTxWithS=nCasesTxWithS,
                    controlCaseRatio=controlCaseRatio, risk0=risk0, VEoverall=VEoverall,
                    Plat0=Plat0, Plat2=Plat2, P0=P0, P2=P2, VELat0=VELat0,
                    VELat1=VELat1, M=M, alpha=alpha, sigma2obs=sigma2obs, rho=rho,
                    biomType=biomType)

# Set parameters for plotPowerCont function
# outComputePower is a list of lists containing output from the computePower function
outComputePower <- pwr
legendText <- paste0("rho = ", c(1, 0.7, 0.4))
plotRRgradVE(outComputePower=outComputePower, legendText=legendText)

## Not run:
```

```

# Output from computePower function is saved in an RData file
computePower(..., saveDir = "myDir", saveFile = "myFile.RData")

# outComputePower is a character string specifying the file containing the computePower output
# outDir is a character string specifying the outComputePower file directory
outComputePower <- paste0("myFile_rho_", c(1, 0.7, 0.4), ".RData")
outDir <- "~/myDir"
legendText <- paste0("rho = ", c(1, 0.7, 0.4))
plotRRgradVE(outComputePower, outDir=outDir, legendText = legendText)

## End(Not run)

```

plotVELatCont	<i>Plotting Treatment (Vaccine) Efficacy Curves for Different Correlate of Risk Relative Risks for Continuous Biomarkers</i>
---------------	--

Description

Plots the treatment (vaccine) efficacy curve for the true latent biomarker for eight different values of the latent correlate of risk relative risk and the lowest vaccine efficacy level for the true biomarker. All curves assume $\rho=1$, and treatment (vaccine) efficacy ranges from 0 to 1. The legend is completely determined by the function.

Usage

```
plotVELatCont(outComputePower, outDir = NULL)
```

Arguments

outComputePower	a list of lists of length 1 containing output from <code>computePower</code> or a character string specifying the .RData file containing <code>computePower</code> output
outDir	a character string specifying path to output .RData file, necessary if outComputePower is a character string. Default is NULL.

Details

`computePower` function input parameter `VELowest` must have length greater than or equal to eight for all eight scenarios to have unique `RRc` and `VELowest`. Otherwise, only `length(VELowest)` unique VE curves will be displayed.

When interpreting the output of the function, the null hypothesis corresponds to a flat curve where vaccine efficacy for all values of the true latent biomarker is equal to the overall vaccine efficacy. Increasing departures from the null hypothesis correspond to increasingly variable and steep VE curves. The output assumes the overall placebo-group endpoint risk between τ and τ_{max} is constant for all values of the latent and observed biomarker and that there is no measurement error ($\rho = 1$). When this is the case, an association of the biomarker with infection risk in the vaccine group (a correlate of risk) is equivalent to an association of the biomarker with treatment (vaccine) efficacy.

The function's plot can also be interpreted in conjunction with the output of the [plotPowerCont](#) function by matching the CoR relative risk in the two plots and examining power compared to VE. This sheds light on the importance of overall VE on power and further enables correlates of risk results to be interpreted in terms of potential correlates of efficacy/protection.

Value

None. The function is called solely for plot generation.

See Also

[computePower](#), [plotPowerCont](#)

Examples

```
# Example scenario with continuous biomarker, where values of rho are varied

# Set input parameters for computePower function
nCasesTx <- 10
nControlsTx <- 300
nCasesTxWithS <- 10
controlCaseRatio <- 3
VEoverall <- 0.75
risk0 <- 0.034
PlatVELowest <- 0.2
VELowest <- seq(0, VEoverall, len=8)
Plat0 <- P0 <- 0.2
Plat2 <- P2 <- 0.6
M <- 13
alpha <- 0.05
sigma2obs <- 1
rho <- 1
biomType <- "continuous"

# Output from computePower function is stored in an object as a list
pwr <- computePower(nCasesTx=nCasesTx, nControlsTx=nControlsTx, nCasesTxWithS=nCasesTxWithS,
                    controlCaseRatio=controlCaseRatio, risk0=risk0, VEoverall=VEoverall,
                    PlatVELowest=PlatVELowest, VELowest=VELowest, Plat0=Plat0, Plat2=Plat2,
                    P0=P0, P2=P2, M=M, alpha=alpha, sigma2obs=sigma2obs, rho=rho, biomType=biomType)

# Set parameters for plotPowerCont function
# outComputePower is a list containing output from the computePower function
outComputePower <- pwr
plotVELatCont(outComputePower=outComputePower)

## Not run:
# Output from computePower function is saved in an RData file
computePower(..., saveDir = "myDir", saveFile = "myFile.RData")
# outComputePower is a character string specifying the file containing the computePower output
# outDir is a character string specifying the outComputePower file directory
outComputePower <- "myFile.RData"
outDir <- "~/myDir"
```

```
plotVELatCont(outComputePower, outDir=outDir)
```

```
## End(Not run)
```

Index

computeN, [2](#), [12](#)
computePower, [3](#), [4](#), [15](#), [17](#), [19–22](#)

plotPowerCont, [12](#), [15](#), [17](#), [22](#)
plotPowerTri, [12](#), [15](#), [16](#), [20](#)
plotROCcurveTri, [18](#)
plotRRgradVE, [19](#)
plotVElatCont, [15](#), [21](#)