# Package 'BDP2'

July 21, 2025

Title Bayesian Adaptive Designs for Phase II Trials with Binary

Type Package

BDP2-package

Index 16

BDP2-package Bayesian Adaptive Designs for Phase II Trials with Binary Endpoint

#### **Description**

Tools and workflow to choose design parameters in Bayesian adaptive single-arm phase II trial designs with binary endpoint (response, success) with possible stopping for efficacy and futility at interim analyses. Also contains routines to determine and visualize operating characteristics. See Kopp-Schneider et al. (2018) <doi:10.1002/bimj.201700209>.

#### **Details**

#### The DESCRIPTION file:

Package: BDP2
Type: Package

Title: Bayesian Adaptive Designs for Phase II Trials with Binary Endpoint

Version: 0.1.3 Date: 2018-07-31

Author: Annette Kopp-Schneider, Manuel Wiesenfarth, Ulrich Abel

Maintainer: Manuel Wiesenfarth < m.wiesenfarth@dkfz.de>

Description: Tools and workflow to choose design parameters in Bayesian adaptive single-arm phase II trial designs with

License: GPL-2

Depends: rmarkdown, shiny, shinyBS

LazyData: TRUE Suggests: knitr VignetteBuilder: knitr

#### Index of help topics:

BDP2 Operating characteristics of a single-arm trial

with a binary endpoint

BDP2-package Bayesian Adaptive Designs for Phase II Trials

with Binary Endpoint

BDP2\_simulate Simulated operating characteristics of a

single-arm trial with a binary endpoint for  $% \left( 1\right) =\left( 1\right) \left( 1\right) \left($ 

different types of futility criteria

BDP2workflow Shiny app for workflow

of futility criteria

pFstop Operating characteristics of a single-arm trial

with a binary endpoint with futility stopping

pFstopEcall operating characteristics of a single-arm trial

with a binary endpoint with futility stopping

BDP2 3

and calling efficacy at interim

pFstopEstop Operating characteristics of a single-arm trial

with a binary endpoint with stopping for

futility and stopping for efficacy

plotBDP2 Plots

Tools and workflow to choose design parameters in Bayesian adaptive single-arm phase II trial designs with binary endpoint (response, success) with possible stopping for efficacy and futility at interim analyses. Also contains routines to determine and visualize operating characteristics.

Main functions: BDP2workflow, BDP2 and plotBDP2. BDP2workflow provides an interactive shiny app which also generates Word/pdf/html reports.

#### Author(s)

Annette Kopp-Schneider, Manuel Wiesenfarth, Ulrich Abel

Maintainer: Manuel Wiesenfarth < m.wiesenfarth@dkfz.de>

#### References

Kopp-Schneider, A., Wiesenfarth, M., Witt, R., Edelmann, D., Witt, O. and Abel, U. (2018). Monitoring futility and efficacy in phase II trials with Bayesian posterior distributions - a calibration approach. *Biometrical Journal*, to appear.

#### **Examples**

```
## Not run:
# Starts Shiny app
BDP2workflow()
## End(Not run)
```

BDP2

Operating characteristics of a single-arm trial with a binary endpoint

#### **Description**

Determines the operating characteristics of a single-arm trial with a binary endpoint (response, success) and interim efficacy and futility analyses. Declaration of efficacy and futility (including possibly early stopping) is based on the posterior probability that the true response rate is at least pE, pF respectively.

BDP2

#### Usage

```
BDP2(n, interim.at, ptrue,
    eff.stop = FALSE,
    pF, cF, pE = NULL, cE = NULL,
    type="PostProb", alpha=0.05,
    shape1F, shape2F, shape1E = NULL, shape2E = NULL,
    simulate = FALSE, nsim = 10000)
```

#### Arguments

n	sample size at the final analysis
interim.at	vector of sample sizes at the interim analyses
ptrue	true (assumed) response rate used for analytical evaluations or simulating the trial
eff.stop	FALSE: No evaluation of efficacy. "call": no stop for efficacy; in this case the program merely calculates the probability that the efficiency criterion is satisfied (possibly triggering a notification of the DMC and the start of the planning of a subsequent trial). "stop": the study ends if the efficacy criterion is reached at an interim analysis.
pF	response rate used for the futility criterion (may be identical to pE)
cF	critical level of posterior probabilities used for declaring futility
рE	response rate used for the efficacy criterion
cE	critical level of posterior probabilities used for declaring efficacy
type	"PostProb" for decisions based on posterior probabilities (default) or "PredictivePower" for decisions based on predictive power (currently only implemented for simulate==TRUE)
alpha	significance level for final test (only for simulate==TRUE & type=="PredictivePower")
shape1F	first parameter of the Beta prior for futility analysis
shape2F	second parameter of the Beta prior for futility analysis
shape1E	first parameter of the Beta prior for efficacy analysis
shape2E	second parameter of the Beta prior for efficacy analysis
simulate	FALSE for analytical evaluation and TRUE for simulation
nsim	number of simulation runs (only used if simulate==TRUE)

#### **Details**

Assumptions: Endpoint (response/no response) data available for all study patients. Beta-binomial model. Prior distribution = Beta(shape1, shape2).

#### Decisions based on posterior probabilities:

The posterior distribution at interim analysis with n.int patients and k.int successes is Beta(k.int + shape1F, n.int + shape2F - k.int) and Beta(k.int + shape1E, n.int + shape2E - k.int), respectively. Efficacy is declared if the posterior probability P(true response rate > pE) is >= cE. Futility is declared if the posterior probability P(true success rate > pF) is < cF, cE translate into futility/efficacy boundaries (maximum number of responses leading to early termination for futility/minimum number of responses leading to declaring of, or early termination for, efficacy).

BDP2workflow 5

**Decisions based on predictive power:** Given the results of the interim analysis, the predictive power at the final analysis (n patients, critical number of successes k.crit) is P(X >= k.crit - k.int), where X follows a beta-binomial distribution with parameters n' = n - n.int, a = k.int + shape1, and b = n.int - k.int + shape2.

Efficacy is declared if the predictive power is >= cE (cE must be high, e.g. 0.70). Futility is declared if the predictive power is < cF (cF must be small, e.g. 0.10). cE, cF translate into futility/efficacy boundaries (maximum number of responses leading to early termination for futility/minimum number of responses leading to declaring of, or early termination for, efficacy).

#### References

Kopp-Schneider, A., Wiesenfarth, M., Witt, R., Edelmann, D., Witt, O. and Abel, U. (2018). Monitoring futility and efficacy in phase II trials with Bayesian posterior distributions - a calibration approach. *Biometrical Journal*, to appear.

#### **Examples**

```
# Operating characteristics with calling for efficacy
BDP2(n=20, interim.at = c(3,9,13,18), ptrue = 0.3,
    eff.stop = "call",
    pF=0.3, cF=0.01, pE=0.12, cE = 0.9,
    type="PostProb",
    shape1F=0.3, shape2F=0.7, shape1E=0.12, shape2E=0.88)

# Operating characteristics with stopping for efficacy
BDP2(n=20, interim.at = c(3,9,13,18), ptrue = 0.3,
    eff.stop = "stop",
    pF=0.3, cF=0.01, pE=0.12, cE = 0.9,
    type="PostProb",
    shape1F=0.3, shape2F=0.7, shape1E=0.12, shape2E=0.88)
```

BDP2workflow

Shiny app for workflow

## Description

Starts a shiny app in the web browser. It provides a workflow to choose design parameters singlearm trial with a binary endpoint (response, success) and interim efficacy and futility analyses as well as routines to determine and visualize operating characteristics. Also Word/pdf/html reports can be generated.

#### Usage

```
BDP2workflow(display.mode = "normal")
```

6 BDP2\_simulate

#### Arguments

```
display.mode display.mode passed to shiny::runApp()
```

#### References

Kopp-Schneider, A., Wiesenfarth, M., Witt, R., Edelmann, D., Witt, O. and Abel, U. (2018). Monitoring futility and efficacy in phase II trials with Bayesian posterior distributions - a calibration approach. *Biometrical Journal*, to appear.

#### **Examples**

```
## Not run:
# Starts Shiny app
   BDP2workflow()
## End(Not run)
```

BDP2\_simulate

Simulated operating characteristics of a single-arm trial with a binary endpoint for different types of futility criteria

#### **Description**

NOTE: Usually function BDP2 will be preferred.

Determines the operating characteristics of a single-arm trial with a binary endpoint (success failure) and interim futility analyses. The user can choose among 10 futility criteria, which are based on predictive or conditional power (the latter either assuming H1 or the MLE), posterior or predictive probabilities, tail probabilities (under H0 or H1), constant success rates, or arbitrary user-defined futility bounds.

Assumptions: Endpoint (success/no success) data available for all study patients. In case of Bayesian analysis: Beta-binomial model. Prior distribution = Beta(shape1,shape2) (uniform if prior.mean=0.5,prior.sampleSize=2) In case of predictive or conditional power, tail probabilities, or rates: One-sided testing in the final analysis.

#### Usage

#### Arguments

```
n sample size at the final analysis
vn.int vector of sample sizes at the interim analyses (the vector may be equal to 1:(n-1)
= continuous monitoring of futility)
p true (assumed) success rate used for simulating the trial
```

BDP2\_simulate 7

p0 success rate corresponding to H0

p1 success rate corresponding to H1 (p1 > p0)

alpha nominal probability of type 1 error used for the final test

crit critical level(s) of predictive/conditional power, (posterior) probabilities, rates,

or patient numbers used for early termination. crit translates into futility boundaries (maximum number of successes leading to early termination). If simple rates are used for monitoring futility (type=7), crit is the critical success rate. Rates or probabilities must be input as percentages (???). In case of type 10 analyses, crit must be a vector of numbers of successes indicating futility bounds at

each (interim or final)analysis.

type type of futility analysis: 1=predictive power; 2=conditional power under H1;

3=conditional power under the MLE; 4=posterior probability of a success rate > p0; 5=posterior probability of a success rate >= p1; 6=predictive probability of reaching a "positive" final study result, where "positive" is defined in terms of the posterior probability of a success rate > p0. (Here, the cirical level must be >= 1-alpha). 7=estimated success rate; 8=p-values; 9="p-values under H1"; 10=user-defined bounds (a vector of critical numbers of successes for each analysis)

nsim number of simulation runs

shape1 shape parameter for prior distribution shape2 shape parameter for prior distribution

#### **Details**

Some methodological details on the 10 types of futility criteria:

Type 1: Predictive power. Given the results of the interim analysis the predictive power at the final analysis (n patients, critical number of successes k.crit) is  $P(X \ge k.crit - k.int)$ , where X follows a beta-binomial distribution with parameters n' = n - n.int, a = k.int + shape 1, and b = n.int - k.int + shape 2.

Type 2,3: Conditional power. The conditional power at the interim analysis is P(X >= k.crit - k.int), where X follows a binomial distribution with parameters n' = n - n.int, and success probability either equal to p1 (futility analysis type 2) or to the estimated success rate (MLE) at the interim analysis (type 3.)

Type 4,5:Posterior probabilities. The posterior distribution at interim analysis with n.int patients and k.int successes is Beta(k.int + shape1, n + shape2 - k.int) Type 4: Futility is declared if the posterior probability P(true success rate > p0) is < crit. (Here, crit must be large,e.g. 70%). Type 5: Futility is declared if the posterior probability P(true success rate >= p1) is < crit. (Here, crit must be small, e.g. 10%).

Type 6: Predictive probability combined with posterior probability. Futility is declared if the posterior predictive probability that the study will be a success is < crit (e.g. 10%). Here, the success is defined by the total number of successes in the trial yielding a posterior probability of at least 1 - alpha (when evaluated in the final analysis) that the true success rate is > p0.

Type 7: Estimated success rates. Futility is declared if the success rate is smaller than a fixed benchmark crit. The final analysis is test-based.

Type 8,9: Tail probabilites under H0,H1. Type 8: The futility criterion uses an alpha level crit that is constant across all interim analysis. The final analysis is test-based. Futility is declared if the p-value (upper tail) is >= crit. Type 8 futility analyses should only be used if the number of patients at

8 crit\_general

the first interim analysis is not too low (say, at least 5 to 10). The value of crit is not identicalt o the alpha level used in the final test. Generally, a fairly high value of crit will be appropriate (e.g. 70%). Type 9: Similar to type 8, but with lower-tail probabilites calculated under H1 ("p-values under H1"). I.e., futility is declared if, under H1, the probability of obtaining at most as many successes as the observed number is < crit ("observed number of successes "too low" to be compatible with H1 at one-sided significance level = crit). Generally, a small value of crit (e.g. 5% or 10%) should be chosen.

Type 10: User-defined boundaries. Here, the futility boundaries (maximum numbers of successes leading to early termination) are directly input by the user. crit ist the vector of these boundaries at each (interim or final) analysis. The study is terminated if the number of successes is at analysis no. m is <= the crit[m].

### **Examples**

```
BDP2_simulate(n=30, vn.int=c(10,20),
p=0.2, p0=0.1, p1=0.3, crit=0.1, type=5, nsim=1000,shape1=1,shape2=1)
```

crit\_general

Calculates CritBoundaries for different types of futility criteria

#### **Description**

Calculates CritBoundaries for different types of futility criteria

#### Usage

```
crit_general(n, p0, p1, vn.int, alpha, crit, type=5, shape1 = 1, shape2 = 1)
```

#### **Arguments**

n	sample size at the final analysis
p0	success rate corresponding to H0
p1	success rate corresponding to H1 $(p1 > p0)$
vn.int	vector of sample sizes at the interim analyses (the vector may be equal to 1:(n-1) = continuous monitoring of futility)
alpha	nominal probability of type 1 error used for the final test
crit	critical level(s) of predictive/conditional power, (posterior) probabilities, rates, or patient numbers used for early termination. crit translates into futility boundaries (maximum number of successes leading to early termination). If simple rates are used for monitoring futility (type=7), crit is the critical success rate. Rates or probabilites must be input as percentages. In case of type 10 analyses, crit must be a vector of numbers of successes indicating futility bounds at each (interim or final)analysis.

pFstop 9

type	see BDP2_simulate for details. Type of futility analysis: 1=predictive power; 2=conditional power under H1; 3=conditional power under the MLE; 4=posterior probability of a success rate > p0; 5=posterior probability of a success rate >= p1; 6=predictive probability of reaching a "positive" final study result, where "positive" is defined in terms of the posterior probability of a success rate > p0. (Here, the ciritcal level must be >= 1- alpha). 7=estimated success rate; 8=p-values; 9="p-values under H1"; 10=user- defined bounds (a vector of critical numbers of successes for each analysis)
shape1	for prior distribution
shape2	for prior distribution
pFstop	Operating characteristics of a single-arm trial with a binary endpoint with futility stopping

## Description

Determines the operating characteristics of a single-arm trial with a binary endpoint (success failure) on the basis of analytical derivations. The design allows for futility stopping. Outputs are

- probability of futility stopping (P.futil) at each interim analysis
- cumulative stopping probability up to the interim (P.futil.cum).

## Usage

```
pFstop(p, vn.int, v.crit)
```

## Arguments

р	true response rate
vn.int	vector of sample sizes at the interim analyses
v.crit	vector of critical boundaries for futility stopping at the interim analyses (stop for futility if number of successes <= boundary). Choose boundary=-1 if no stopping is allowed.

## Examples

```
pFstop(p=0.3, vn.int=c(3,9,13,18,20), v.crit=c(0,1,2,3,3))\\
```

10 pFstopEstop

pFstopEcall	operating characteristics of a single-arm trial with a binary endpoint with futility stopping and calling efficacy at interim
	wun juutity stopping and catting efficacy at interim

#### **Description**

Determines the operating characteristics of a single-arm trial with a binary endpoint (success failure) on the basis of analytical derivations. The design allows for futility stopping and evaluates efficacy at interim. Outputs are

- probability of calling efficacy (P.effic) at each interim analysis
- cumulative probability of calling efficacy up to the interim (P.effic.cum).
- probability of futility stopping (P.futil) at each interim analysis
- cumulative stopping probability up to the interim (P.futil.cum).

#### Usage

```
pFstopEcall(p, vn.int, v.critE, v.critF)
```

#### **Arguments**

р	true response rate
vn.int	vector of sample sizes at the interim analyses
v.critE	vector of critical boundaries for calling efficacy at the interim analyses (call efficacy if number of successes >= boundary)
v.critF	vector of critical boundaries for futility stopping at the interim analyses (stop for futility if number of successes <= boundary). Choose boundary=-1 if no stopping for futility is allowed.

#### **Examples**

```
pFstopEcall(p=0.3,vn.int=c(3,9,13,18,20),v.critE=4:8, v.critF=c(0,1,2,3,3))

pFstopEstop

Operating characteristics of a single-arm trial with a binary endpoint with stopping for futility and stopping for efficacy
```

## Description

Determines the operating characteristics of a single-arm trial with a binary endpoint (success - failure) on the basis of analytical derivations. The design allows for stopping for futility and stopping for efficacy. Outputs are

- probability of stopping for efficacy (P.effic) at each interim analysis
- cumulative probability of stopping for efficacy up to the interim (P.effic.cum).
- probability of futility stopping (P.futil) at each interim analysis
- cumulative stopping probability up to the interim (P.futil.cum).

plot.cE\_vs\_pEcall 11

#### Usage

```
pFstopEstop(p, vn.int, v.critE, v.critF)
```

#### **Arguments**

р	true response rate
vn.int	vector of sample sizes at the interim analyses
v.critE	vector of critical boundaries for calling efficacy at the interim analyses (stop for efficacy if number of successes >= boundary)
v.critF	vector of critical boundaries for futility stopping at the interim analyses (stop for futility if number of successes <= boundary). Choose boundary=-1 if no stopping for futility is allowed.

#### **Examples**

```
pFstopEstop(p=0.3,vn.int=c(3,9,13,18,20),v.critE=4:8, v.critF=c(0,1,2,3,3))
```

#### **Description**

Plot objects returned by plotBDP2 with arguments x="cE" and either y="PEcall" or y="PEstop" which can computationally be relatively expensive.

#### Usage

```
## S3 method for class 'cE_vs_pEcall'
plot(x, ...)
## S3 method for class 'cE_vs_pEstop'
plot(x, ...)
```

#### **Arguments**

```
    x Object returned by plotBDP2 with arguments x="cE" and either y="PEcall" or y="PEstop"
    ... arguments passed to plot.default
```

#### **Examples**

```
col=c("green","red"),cex.sub=.8)
plot(res)
```

plotBDP2

Plots

#### Description

Output of desired plots

#### Usage

#### **Arguments**

Х

character string specifying what is given by the x axis

- "n" for number of patients at final or at first interim, depending on plot
- "ptrue" for the true response rate
- "k" number of successes at interim (only for y=="PredictivePower")
- "cE" for the critical level of posterior probabilities used for declaring efficacy
- "cF" for the critical level of posterior probabilities used for declaring futility

У

character string specifying what is given by the y axis

- "Prob0Successes" for plot of probability of 0 successes out of n at first interim, i.e. (1-p)^n. Two curves are generated: one for p0 and one for p1. Choose x="n" and set n to a vector of number patients at first interim.
- "PostProb0or1Successes" generates two curves of posterior probability that
  reponse rate exceeds cF. One given 0 successes and one given 1 success
  observed in n, i.e. P(p>pFl 0 successes out of n) and P(p>pFl 1 success out
  of n). Choose x="n" and set n to a vector of number patients at first interim.
- "bFbE" for plot of boundaries for futility bF and for efficacy bE in terms of number of successes per n number of patients. Futility stop if number of successes <= bF, call efficacy if number of successes >= bE. Choose x="n" and set n to the maximal number of patients (not a vector).

• "PEcall\_p0\_p1" for plot of probability to call efficacy at final (i.e. P(p>pElData)>=cE), evaluated for data generated with control response rate p0 (corresponds to type I error) and for data generated with target response rate p1 (corresponds to power) for varying number of patients at final. Choose x="n" and set n to a vector of number patients at final.

- "PEstop\_p0\_p1" as above but with stopping for efficacy instead of calling efficacy
- "PFstopEcall" for plotting operating characteristics for a given response rate as a function of number of patients at final, n. Shows one curve (default: in red) for the probability of stopping for futility up to final analysis and one curve (default: in green) for the probability of calling efficacy at final analysis. For the setting ptrue=p0, the curves show the probability of true stopping (default: in red) and type I error (default: in green). For the setting ptrue=p1, the curves show the probability of false stopping (default: in red) and power (default: in green). Choose x="n" and set n to a vector of number patients at final.
- "PFstopEstop" as above but with stopping for efficacy instead of calling efficacy
- "PEcall" for plot of probability to call efficacy at final (i.e. P(p>pElData)>=cE ).
  - For x="ptrue" then this is the power function.
  - For x="cE" this gives plots of probability to call efficacy at final, evaluated for data generated with control response rate p0 (corresponds to type I error) and for data generated with target response rate p1 (corresponds to power).
- "PEstop"
  - For x="ptrue" then this is the power function.
  - For x="cE" this gives plots of probability to call efficacy at final, evaluated for data generated with control response rate p0 (corresponds to type I error) and for data generated with target response rate p1 (corresponds to power).
- "PFstop" for plot of cumulative probability to stop for futility up to final.
  - For x="ptrue" this gives the futility stopping probability as function of ptrue.
  - For x="cF" this gives 2 curves, evaluated for data generated with control response rate p0 (corresponds to true stopping probability, default: in green) and for data generated with target response rate p1 (corresponds to false stopping probability, default: in red).
- "ExpectedNumber" for expected number of patients in the trial. Choose x= "ptrue". Takes stopping for efficacy into account if Estop==TRUE
- "PredictivePower" for predictive power (only for x=="k")

sample size at the final analysis, vector if x=="n"

vector of sample sizes at the interim analyses

true (assumed) response rate used for simulating the trial, vector if x=="ptrue" response rate used for the futility criterion P(p>pF|Data) < cF (may be identical to pE)

interim.at ptrue pF

cF	critical level of posterior probabilities used for declaring futility, vector if x=="cF"
pE	response rate used for the efficacy criterion P(p>pElData) >= cE
cE	critical level of posterior probabilities used for declaring efficacy, vector if x=="cE"
p0	response rate corresponding to H0
p1	response rate corresponding to H1 ( $p1 > p0$ )
shape1F	first parameter of the Beta prior for futility analysis
shape2F	second parameter of the Beta prior for futility analysis
shape1E	first parameter of the Beta prior for efficacy analysis
shape2E	second parameter of the Beta prior for efficacy analysis
Estop	Stop for efficacy? Defaults to FALSE. Only relevant if y=="ExpectedNumber".
col	line color, for some plots vector of length 2.
add	add line to existing plot. Only supported if x=="ptrue"
show	show plot (otherwise computed objects are invisibly returned)
progress	only used by shiny app
cex.legend	size of legend text relative to cex
	additional arguments passed to plot.default()

#### References

Kopp-Schneider, A., Wiesenfarth, M., Witt, R., Edelmann, D., Witt, O. and Abel, U. (2018). Monitoring futility and efficacy in phase II trials with Bayesian posterior distributions - a calibration approach. *Biometrical Journal*, to appear.

#### **Examples**

```
# See vignette for more details and examples
pF=0.3
pE=0.12
shape1F=0.3
shape2F=0.7
shape1E=0.12
shape2E=0.88
cF=0.01
cE=0.9
nvec=c(18:40)
interim.at=c(10,20,30)
# Type I error and probability of true stopping for the uninteresting response rate.
ptrue=0.12
plotBDP2(x="n", y="PFstopEcall",
         n =nvec, interim.at = interim.at,
         pF=pF,cF=cF,pE=pE,cE=cE,ptrue=ptrue,
         shape1F=shape1F, shape2F=shape2F, shape1E=shape1E, shape2E=shape2E)
# Power and probability of false stopping for the target response rate.
```

## **Index**

```
* design
    BDP2, 3
    BDP2workflow, 5
* package
    BDP2-package, 2
BDP2, 3, 3, 6
BDP2-package, 2
BDP2_simulate, 6, 9
BDP2workflow, 3, 5
\verb|crit_general|, 8
pFstop, 9
{\it pFstopEcall}, \\ 10
{\it pFstopEstop}, 10
plot.cE_vs_pEcall, 11
plot.cE_vs_pEstop (plot.cE_vs_pEcall),
         11
plot.default, 11
plotBDP2, 3, 11, 12
```