

# An example to illustrate boundary re-calculations during the trial

*Marcel Wolbers, Gernot Wassmer, and Friedrich Pahlke*

*Last change: 28 Juni, 2019*

This R markdown file provides an example for updating the group sequential boundaries when using an alpha spending function approach based on observed information rates.

```
# Load rpact
library(rpact)
packageVersion("rpact") # version should be version 2.0.1 or later
```

```
## [1] '2.0.1.9003'
```

Group-sequential designs based on  $\alpha$ -spending functions protect the type I error exactly even if the pre-planned interim schedule is not exactly adhered to. However, this requires re-calculation of the group-sequential boundaries at each interim analysis based on actually observed information fractions. Unless deviations from the planned information fractions are substantial, the re-calculated boundaries are quite similar to the pre-planned boundaries and the re-calculation will affect the actual test decision only on rare occasions.

Importantly, it is not allowed that the timing of future interim analyses is “motivated” by results from earlier interim analyses as this could inflate the type I error rate. Deviations from the planned information fractions should thus only occur due to operational reasons (as it is difficult to hit the exact number of events exactly in a real trial) or due to external evidence.

The general principles for these boundary re-calculation are as follows (see also, Wassmer & Brannath, 2016, p78f):

- Updates at interim analyses prior to the final analysis:
  - Information fractions are updated according to the actually observed information fraction at the interim analysis relative to the **planned** maximum information.
  - The planned  $\alpha$ -spending function is then applied to these updated information fractions.
- Updates at the final analysis in case the observed information at the final analysis is larger (“over-running”) or smaller (“under-running”) than the planned maximum information:
  - Information fractions are updated according to the actually observed information fraction at all interim analyses relative to the **observed** maximum information.  $\Rightarrow$  Information fraction at final analysis is re-set to 1 but information fractions for earlier interim analyses are also changed.
  - The originally planned  $\alpha$ -spending function cannot be applied to these updated information fractions because this would modify the critical boundaries of earlier interim analyses which is clearly not allowed. Instead, one uses the  $\alpha$  that has actually been spent at earlier interim analyses and spends all remaining  $\alpha$  at the final analysis.

This general principle be implemented via a user-defined  $\alpha$ -spending function and is illustrated for an example trial with a survival endpoint below.

## Original trial design

The original trial design for this example is based on a standard O’Brien&Fleming type  $\alpha$ -spending function with planned efficacy interim analyses after 50% and 75% of information as specified below.

```
# Initial design
design <- getDesignGroupSequential(sided = 1, alpha = 0.025, beta = 0.2,
  informationRates = c(0.5, 0.75, 1), typeOfDesign="asOF")
```

```

# Initial sample size calculation
sampleSizeResult <- getSampleSizeSurvival(
  design = design,
  lambda2 = log(2)/60, hazardRatio = 0.75,
  dropoutRate1 = 0.025, dropoutRate2 = 0.025, dropoutTime = 12,
  accrualTime = 0, accrualIntensity = 30,
  maxNumberOfSubjects = 1000)

# Summarize design
simpleBoundarySummary <- function(result) {
  parameters <- list("Stage" = c(1:result$.design$kMax),
    "Information rate" = round(t(result$.design$informationRates), 2),
    "Number of events" = round(t(result$.design$eventsPerStage[, 1]), 1),
    "Analysis time under H1" = round(t(result$.design$analysisTime[, 1]), 1),
    "Cumulative alpha spent" = round(result$.design$alphaSpent, 4),
    "Stage levels" = round(result$.design$stageLevels, 4),
    "Efficacy boundary (Z-value scale)" = round(result$.design$criticalValues, 3),
    "Efficacy boundary (treatment effect scale)" =
      round(t(result$.design$criticalValuesEffectScale[, 1]), 3))
  names(parameters) <- paste0(format(names(parameters)), " ")
  for (paramName in names(parameters)) {
    values <- parameters[[paramName]]
    values <- format(c(" ", values))[2:(length(values) + 1)]
    cat(paramName, values, "\n")
  }
}
simpleBoundarySummary(sampleSizeResult)

```

```

## Stage                1      2      3
## Information rate      0.5    0.75   1
## Number of events     193.4  290.1  386.8
## Analysis time under H1 39.1   52.7   69.1
## Cumulative alpha spent 0.0015 0.0096 0.025
## Stage levels         0.0015 0.0092 0.022
## Efficacy boundary (Z-value scale) 2.963  2.359  2.014
## Efficacy boundary (treatment effect scale) 0.653  0.758  0.815

```

## Boundary update at the first interim analysis

Assume that the first interim was conducted after 205 rather than the planned 194 events.

The updated design is calculated as per the code below. Note that for the calculation of boundary values on the treatment effect scale, we use the function `getPowerSurvival` with the updated design rather than the function `getSampleSizeSurvival` as we are only updating the boundary, not the sample size or the maximum number of events.

```

# Update design using observed information fraction at first interim.
# Information fraction of later interim analyses is not changed.
designUpdate1 <- getDesignGroupSequential(sided = 1, alpha = 0.025, beta = 0.2,
  informationRates = c(205/387, 0.75, 1), typeOfDesign="asOF")

# Recalculate the power to get boundary values on the effect scale
# (Use original maxNumberOfEvents and sample size)
powerUpdate1 <- getPowerSurvival(

```

```

design = designUpdate1,
lambda2 = log(2)/60,hazardRatio = 0.75,
dropoutRate1 = 0.025, dropoutRate2 = 0.025, dropoutTime = 12,
accrualTime = 0,accrualIntensity = 30,
maxNumberOfSubjects = 1000, maxNumberOfEvents = 387, directionUpper = FALSE)
powerUpdate1

```

```

## Design plan parameters and output for survival data:
##
## Design parameters:
##   Significance level           : 0.0250
##   Test                         : one-sided
##
## User defined parameters:
##   Direction upper             : FALSE
##   Lambda (2)                  : 0.0116
##   Hazard ratio                 : 0.750
##   Maximum number of subjects  : 1000.0
##   Maximum number of events    : 387.0
##   Event time                   : 12
##   Accrual intensity           : 30.0
##   Drop-out rate (1)           : 0.025
##   Drop-out rate (2)           : 0.025
##
## Default parameters:
##   Type of computation         : Schoenfeld
##   Theta H0                    : 1
##   Planned allocation ratio    : 1
##   Piecewise survival times    : 0.00
##   Drop-out time               : 12.00
##
## Sample size and output:
##   pi (1)                      : 0.0987
##   pi (2)                      : 0.129
##   Median (1)                  : 80.0
##   Median (2)                  : 60.0
##   Lambda (1)                  : 0.00866
##   Accrual time                : 33.33
##   Follow up time              : 35.81
##   Analysis times [1]          : 40.60
##   Analysis times [2]          : 52.73
##   Analysis times [3]          : 69.14
##   Expected study duration     : 57.75
##   Maximal study duration      : 69.14
##   Number of events by stage [1] : 205.0
##   Number of events by stage [2] : 290.2
##   Number of events by stage [3] : 387.0
##   Expected number of events   : 317.0
##   Number of subjects [1]      : 1000.0
##   Number of subjects [2]      : 1000.0
##   Number of subjects [3]      : 1000.0
##   Expected number of subjects : 1000.0
##   Reject per stage [1]        : 0.210
##   Reject per stage [2]        : 0.329

```

```
## Reject per stage [3] : 0.261
## Overall reject : 0.8
## Early stop : 0.539
## Critical values (effect scale) [1] : 0.670
## Critical values (effect scale) [2] : 0.758
## Critical values (effect scale) [3] : 0.815
## Local one-sided significance levels [1] : 0.002073
## Local one-sided significance levels [2] : 0.008998
## Local one-sided significance levels [3] : 0.021968
##
## Legend:
## (i): values of treatment arm i
## [k]: values at stage k
```

The updated information rates and corresponding boundaries as per the output above are summarized as follows:

## Stage	1	2	3
## Information rate	0.53	0.75	1
## Number of events	205	290.2	387
## Analysis time under H1	40.6	52.7	69.1
## Cumulative alpha spent	0.0021	0.0096	0.025
## Stage levels	0.0021	0.009	0.022
## Efficacy boundary (Z-value scale)	2.867	2.366	2.015
## Efficacy boundary (treatment effect scale)	0.67	0.758	0.815

## Boundary update at the second interim analysis

Assume that the efficacy boundary was not crossed at the first interim analysis and the trial continued to the second interim analysis which was conducted after 285 rather than the planned 291 events. The updated design is calculated in the same way as for the first interim analysis as per the code below. The idea is to use the cumulative alpha spent from the first stage and an updated cumulative alpha that is spent for the second stage. For the second stage, this can be obtained with the original O'Brien & Fleming alpha spending function:

```
# Update design using observed information fraction at first and second interim.
designUpdate2 <- getDesignGroupSequential(sided = 1, alpha = 0.025, beta = 0.2,
  informationRates = c(205/387,285/387,1), typeOfDesign="asOF")

# Recalculate power to get boundary values on effect scale
# (Use original maxNumberOfEvents and sample size)
powerUpdate2 <- getPowerSurvival(
  design = designUpdate2,
  lambda2 = log(2)/60,hazardRatio = 0.75,
  dropoutRate1 = 0.025, dropoutRate2 = 0.025, dropoutTime = 12,
  accrualTime = 0,accrualIntensity = 30,
  maxNumberOfSubjects = 1000, maxNumberOfEvents = 387, directionUpper = FALSE)
powerUpdate2

## Design plan parameters and output for survival data:
##
## Design parameters:
## Significance level : 0.0250
## Test : one-sided
##
```

```

## User defined parameters:
##   Direction upper                : FALSE
##   Lambda (2)                    : 0.0116
##   Hazard ratio                   : 0.750
##   Maximum number of subjects     : 1000.0
##   Maximum number of events       : 387.0
##   Event time                     : 12
##   Accrual intensity              : 30.0
##   Drop-out rate (1)              : 0.025
##   Drop-out rate (2)              : 0.025
##
## Default parameters:
##   Type of computation            : Schoenfeld
##   Theta H0                      : 1
##   Planned allocation ratio       : 1
##   Piecewise survival times       : 0.00
##   Drop-out time                  : 12.00
##
## Sample size and output:
##   pi (1)                        : 0.0987
##   pi (2)                        : 0.129
##   Median (1)                    : 80.0
##   Median (2)                    : 60.0
##   Lambda (1)                    : 0.00866
##   Accrual time                   : 33.33
##   Follow up time                 : 35.81
##   Analysis times [1]             : 40.60
##   Analysis times [2]             : 51.93
##   Analysis times [3]             : 69.14
##   Expected study duration        : 57.82
##   Maximal study duration         : 69.14
##   Number of events by stage [1] : 205.0
##   Number of events by stage [2] : 285.0
##   Number of events by stage [3] : 387.0
##   Expected number of events      : 317.2
##   Number of subjects [1]         : 1000.0
##   Number of subjects [2]         : 1000.0
##   Number of subjects [3]         : 1000.0
##   Expected number of subjects    : 1000.0
##   Reject per stage [1]           : 0.210
##   Reject per stage [2]           : 0.310
##   Reject per stage [3]           : 0.281
##   Overall reject                  : 0.8
##   Early stop                     : 0.52
##   Critical values (effect scale) [1] : 0.670
##   Critical values (effect scale) [2] : 0.753
##   Critical values (effect scale) [3] : 0.815
##   Local one-sided significance levels [1] : 0.002073
##   Local one-sided significance levels [2] : 0.008356
##   Local one-sided significance levels [3] : 0.022154
##
## Legend:
##   (i): values of treatment arm i
##   [k]: values at stage k

```

## Boundary update at the final analysis

Assume that the efficacy boundary was also not crossed at the second interim analysis and the trial continued to the final analysis which was conducted after 393 rather than the planned 387 events. The updated design is calculated as per the code below. The idea here to use the cumulative alpha spent from the first *and* the second stage stage and the final alpha that is spent for the last stage. An updated correlation has to be used and the original O'Brien & Fleming alpha spending function cannot be used anymore. Instead, the alpha spending function needs to be user defined as follows:

```
# Update boundary with information fractions as per actually observed event numbers
# !! use user-defined alpha-spending and spend alpha according to actual alpha spent
# according to the second interim analysis
designUpdate3 <- getDesignGroupSequential(sided = 1, alpha = 0.025, beta = 0.2,
  informationRates = c(205,285,393)/393,
  typeOfDesign = "asUser",
  userAlphaSpending = designUpdate2$alphaSpent)

# Recalculate power to get boundary values on effect scale
# (Use planned sample size and **observed** maxNumberOfEvents)
powerUpdate3 <- getPowerSurvival(
  design = designUpdate3,
  lambda2 = log(2)/60,hazardRatio = 0.75,
  dropoutRate1 = 0.025, dropoutRate2 = 0.025, dropoutTime = 12,
  accrualTime = 0,accrualIntensity = 30,
  maxNumberOfSubjects = 1000, maxNumberOfEvents = 393, directionUpper = FALSE)
powerUpdate3
```

```
## Design plan parameters and output for survival data:
##
## Design parameters:
##   Significance level           : 0.0250
##   Test                         : one-sided
##
## User defined parameters:
##   Direction upper             : FALSE
##   Lambda (2)                  : 0.0116
##   Hazard ratio                 : 0.750
##   Maximum number of subjects   : 1000.0
##   Maximum number of events     : 393.0
##   Event time                   : 12
##   Accrual intensity            : 30.0
##   Drop-out rate (1)            : 0.025
##   Drop-out rate (2)            : 0.025
##
## Default parameters:
##   Type of computation          : Schoenfeld
##   Theta H0                     : 1
##   Planned allocation ratio     : 1
##   Piecewise survival times     : 0.00
##   Drop-out time                : 12.00
##
## Sample size and output:
##   pi (1)                       : 0.0987
##   pi (2)                       : 0.129
##   Median (1)                   : 80.0
```

```

## Median (2) : 60.0
## Lambda (1) : 0.00866
## Accrual time : 33.33
## Follow up time : 36.95
## Analysis times [1] : 40.60
## Analysis times [2] : 51.93
## Analysis times [3] : 70.28
## Expected study duration : 58.37
## Maximal study duration : 70.28
## Number of events by stage [1] : 205.0
## Number of events by stage [2] : 285.0
## Number of events by stage [3] : 393.0
## Expected number of events : 320.1
## Number of subjects [1] : 1000.0
## Number of subjects [2] : 1000.0
## Number of subjects [3] : 1000.0
## Expected number of subjects : 1000.0
## Reject per stage [1] : 0.210
## Reject per stage [2] : 0.310
## Reject per stage [3] : 0.286
## Overall reject : 0.806
## Early stop : 0.52
## Critical values (effect scale) [1] : 0.670
## Critical values (effect scale) [2] : 0.753
## Critical values (effect scale) [3] : 0.816
## Local one-sided significance levels [1] : 0.002073
## Local one-sided significance levels [2] : 0.008356
## Local one-sided significance levels [3] : 0.022021
##
## Legend:
## (i): values of treatment arm i
## [k]: values at stage k

```

## Overview of all boundaries updates

For easier comparison, all discussed boundary updates are summarized more conveniently below. Note that each update only affects boundaries for the current or later analyses, i.e., earlier boundaries are never retrospectively modified.

### Original design

## Stage	1	2	3
## Information rate	0.5	0.75	1
## Number of events	193.4	290.1	386.8
## Analysis time under H1	39.1	52.7	69.1
## Cumulative alpha spent	0.0015	0.0096	0.025
## Stage levels	0.0015	0.0092	0.022
## Efficacy boundary (Z-value scale)	2.963	2.359	2.014
## Efficacy boundary (treatment effect scale)	0.653	0.758	0.815

### Updated boundaries at the first interim analysis

## Stage	1	2	3
## Information rate	0.53	0.75	1

## Number of events	205	290.2	387
## Analysis time under H1	40.6	52.7	69.1
## Cumulative alpha spent	0.0021	0.0096	0.025
## Stage levels	0.0021	0.009	0.022
## Efficacy boundary (Z-value scale)	2.867	2.366	2.015
## Efficacy boundary (treatment effect scale)	0.67	0.758	0.815

### Updated boundaries at the second interim analysis

## Stage	1	2	3
## Information rate	0.53	0.74	1
## Number of events	205	285	387
## Analysis time under H1	40.6	51.9	69.1
## Cumulative alpha spent	0.0021	0.009	0.025
## Stage levels	0.0021	0.0084	0.0222
## Efficacy boundary (Z-value scale)	2.867	2.393	2.011
## Efficacy boundary (treatment effect scale)	0.67	0.753	0.815

### Updated boundaries at the final analysis

## Stage	1	2	3
## Information rate	0.52	0.73	1
## Number of events	205	285	393
## Analysis time under H1	40.6	51.9	70.3
## Cumulative alpha spent	0.0021	0.009	0.025
## Stage levels	0.0021	0.0084	0.022
## Efficacy boundary (Z-value scale)	2.867	2.393	2.014
## Efficacy boundary (treatment effect scale)	0.67	0.753	0.816

---

System: rpact 2.0.1.9003, R version 3.5.3 (2019-03-11), platform: x86\_64-w64-mingw32

To cite package ‘rpact’ in publications use:

Gernot Wassmer and Friedrich Pahlke (2019). rpact: Confirmatory Adaptive Clinical Trial Design and Analysis. R package version 2.0.1.9003. <https://www.rpact.org>

### License

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>.