rpact Exercises

Kaspar Rufibach, Marc Vandemeulebroecke, Gernot Wassmer,
Marcel Wolbers

BBS Training Course
Basel, 13 September 2022

Getting started

- Make sure that you have Rstudio, R version $\geq$ 3.5 and rpact version 3.0 installed on your laptop.
- Load the rpact package:

```
library(rpact)
packageVersion("rpact")
    # version should be version 3.0 or higher
```

Exercise 1 (Continuous endpoint)

A confirmatory, randomized and blinded study of an investigational drug against Placebo is planned in mild to moderate Alzheimer's disease. The primary endpoint is the change from baseline in ADAS-Cog, a neuropsychological test battery measuring cognitive abilities, assessed 6 months after treatment initiation. The ADAS-Cog has a range of 0-70; we reverse its scale so that greater values are good. We consider our primary endpoint as approximately normally distributed, and for simplicity we assume a known standard deviation of 10. We believe that the improvement in the primary endpoint that can be achieved with the investigational drug is at least 4 points better than that under Placebo; and we want to have 80% chance of achieving a significant result if this is indeed the case. However, if the investigational drug is no better than Placebo, we want to have no more than 2.5% chance to claim success. This yields a sample size of approximately $n = 100$ per treatment group for a trial with fixed sample size.

Exercise 1a (the "alpha calculus")

We want to build in a "sanity check" mid-way through the trial. More precisely, we implement an interim analysis using the inverse normal method, with the following characteristics (all with respect to the primary endpoint):

- Stop for futility if the investigational drug appears worse than Placebo
- Stop for efficacy if the investigational drug appears "very significantly better" than Placebo ($p < 0.0001$)

a) Which set of $(\alpha, \alpha_0, \alpha_1, \alpha_2)$ satisfies these conditions?
Hint: Use `typeOfDesign = "asUser"` and an appropriate `userAlphaSpending` vector

b) What regulatory issues could this raise?

Exercise 1b (early stopping and sample size adaptation)

a) At the interim analysis after $n_1 = 50$ patients per group, we observe an average ADAS-Cog improvement of 4 points under the investigational drug and of 1 point under Placebo.

   Should we stop or continue the trial?

   Hint: Use `getDataset` and `getAnalysisResults`.

b) At the same time, there is a change in strategy, and we now want 90% power at an improvement of 4 points over placebo.

   Determine the sample size per treatment group for the second stage of the trial, in light of the interim results.

   Hint: Use `getSampleSizeMeans()` and `getAnalysisResults()`.

Exercise 1c (final inference)

In the second stage of the trial, we observe an average ADAS-Cog improvement of only **3 points** under the investigational drug and of 1 point under Placebo.

a) Can we reject the null hypothesis and claim superiority of the investigational drug over placebo?

b) Compute the overall ("exact") p-value and confidence interval for the adaptive trial.

c) What would a "naïve" z-test have concluded, based on all observations and ignoring the adaptive nature of the trial? What is your interpretation of the situation?

Exercise 2 (Planning of survival design)

A survival trial is planned to be performed with one interim stage and using an O'Brien & Fleming type α-spending approach at $\alpha = 0.025$. The interim is planned to be performed after half of the necessary events were observed. It is assumed that the median survival time is 18 months in the treatment group, and 12 months in the control. Assume that the drop-out rate is 5% after 1 year and the drop-out time is exponentially distributed.

a) The patients should be recruited within 12 months assuming uniform accrual. Assume an additional follow-up time of 12 months, i.e., the study should be conducted within 2 years. Calculate the necessary number of events and patients (total and per month) in order to reach power 90% with the assumed median survival times if the survival time is exponentially distributed. Under the postulated assumption, estimate interim and final analysis time.

Exercise 2 (Planning of survival design)

b) Assume that 25 patients can be recruited each month and that there is uniform accrual. Estimate the necessary accrual time if the planned follow-up time remains unchanged.

c) Assume that accrual stops after 16 months with 25 patients per month, i.e., after 400 patients were recruited. What is the estimated necessary follow-up time?

d) How do the results change if in the first 3 months 15 patients, in the second 3 months 20 patients, and after 6 months 25 patients per month can be accrued?

Hint: Use `getSampleSizeSurvival()`

Exercise 3 (Adaptive survival design)

Assume that the study from Exercise 2 is planned with 257 events and 400 patients under the assumptions that accrual stops after 16 months with 25 patients per month.

a) Verify by simulation the correctness of the results obtained by the analytical formulae.

b) Assume now that a sample size increase up to a ten-fold of the originally planned number of events is foreseen. Conditional power 90% based on the observed hazard ratios is used to increase the number of events. Assess by simulation the magnitute of power increase when using the appropriate method.

c) Simulate the Type I error rate when using the (inappropriate) group sequential and the inverse normal method.

   Make sure that enough subjects are used in the simulation (set `maxNumberOfSubjects = 3000` and no drop-outs)!

Hint: Use `getPowerSurvival()` and `getSimulationSurvival()`

Exercise 4 (Sample size calculation for testing rates)

Suppose a trial should be conducted in 3 stages where at the first stage 50%, at the second stage 75%, and at the final stage 100% of the information should be observed. O'Brien & Fleming boundaries should be used with one-sided $\alpha = 0.025$ and non-binding futility bounds 0 and 0.5 for the first and the second stage, respectively, on the z-value scale.

The endpoints are binary (failure rates) and should be compared in a parallel group design, i.e., the null hypothesis to be tested is $H_0 : \pi_1 - \pi_2 = 0$, which is tested against the alternative $H_1 : \pi_1 - \pi_2 < 0$.

a) What is the necessary sample size to achieve 90% power if the failure rates are assumed to be $\pi_1 = 0.40$ and $\pi_2 = 0.60$? What is the optimum allocation ratio?

b) Illustrate the decision boundaries on different scales.

c) Suppose that $N = 280$ subjects were planned for the study. What is the power if the failure rate in the active treatment group is $\pi_1 = 0.50$?

d) Illustrate power, expected sample size, and early/futility stop for a range of alternative values.

Hint: Use `getSampleSizeRates()` and `getPowerRates()`

Exercise 5 (Sample size reassessment)

Using an adaptive design, the sample size from Example 4 in the last interim can be increased up to a 4-fold of the originally planned sample size for the last stage. Conditional power 90% *based on the observed effect sizes (failure rates)* is used to increase the sample size.

a) Use the inverse normal method to allow for the sample size increase and compare the test characteristics with the group sequential design from Example 4.

b) Illustrate the gain in power when using the adaptive sample size recalculation.

c) Create a histogram for the attained sample size of the study when using the adaptive sample size recalculation. How often will the maximum sample size be achieved?

Hint: Use `getSimulationRates()` and `ggplot2()`

Exercise 6 (Multi-armed design with continuous endpoint)

Suppose a trial is conducted with three active treatment arms (+ one control arm). An adaptive design using the equally weighted inverse normal method with two interim analyses using O'Brien & Fleming boundaries is chosen where in both interim analyses a selection of treatment arms is foreseen ($\alpha = 0.025$ one-sided). It is decided to test the intersection tests in the closed system of hypotheses with Dunnett's test. In the designing stage, it was decided to conduct the study with 20 patients per treatment arm and stage where at interim the sample size can be redefined.

a) Suppose, at the first stage, the following results were obtained:

arm	n	mean	std
1	19	3.11	1.77
2	23	3.87	1.23
3	21	4.12	1.64
control	21	3.02	1.72

Perform the closed test and assess the conditional power in order to decide which treatment arm(s) should be selected and if the sample size should be redefined.

Hint: Use `getDataset()` with `means1 =`, `means2 =`, `means3 =`, `means4 =`, etc., where index 4 refers to the control

Exercise 6 (Multi-armed design with continuous endpoint)

b) Suppose it was decided to drop treatment arm 1 for stage 2 and leave the sample size for the remaining arms unchanged. For the second stage, the following results were obtained:

arm	n	mean	std
2	23	3.66	1.11
3	19	3.98	1.21
control	22	2.99	1.82

Perform the closed test and discuss whether or not to stop the study and determine overall p-values and confidence intervals.

c) Does the Bonferroni and the Simes test provide the same results?