Performance Comparison of Optimized Adaptive Enrichment Designs Versus Standard Designs

June 14, 2019

1 Introduction

Four optimized designs are presented in this report:

- 1SEA: the optimal 1-stage design with equal alpha allocation to all hypotheses;
- 1SOA: the optimal 1-stage design with alpha allocation chosen by optimization;
- 2SEA: the optimal 2-stage design with equal alpha allocation to all hypotheses and futility boundaries chosen by optimization;
- 2SOA: the optimal 2-stage design with interim analysis timing, futility boundaries, and alpha allocation all chosen by optimization.

Each design was optimized to minimize the expected sample size weighted by user-supplied weights over a collection of hypothetical treatment effect scenarios, subject to user-supplied constraints on power and Type I error. See Section 1.1 for the full list of user-supplied parameters.

Details of the design space we searched over and the optimization method can be found in the following papers:

- (For one treatment versus one control) Betz, Josh; Steingrimsson, Jon Arni; Qian, Tianchen; and Rosenblum, Michael. COMPARISON OF ADAPTIVE RANDOMIZED TRIAL DESIGNS FOR TIME-TO-EVENT OUTCOMES THAT EXPAND VERSUS RESTRICT ENROLLMENT CRITERIA, TO TEST NON-INFERIORITY (September 2017). Johns Hopkins University, Dept. of Biostatistics Working Papers. Working Paper 289. `http://biostats.bepress.com/jhubiostat/paper289`
- (For two treatments versus one control) Steingrimsson, Jon Arni; Betz, Joshua; Qian, Tiachen; and Rosenblum, Michael. OPTIMIZED ADAPTIVE ENRICHMENT DESIGNS FOR MULTI-ARM TRIALS: LEARNING WHICH SUBPOPULATIONS BENEFIT FROM DIFFERENT TREATMENTS (September 2017). Johns Hopkins University, Dept. of Biostatistics Working Papers. Working Paper 288. `http://biostats.bepress.com/jhubiostat/paper288`

1.1 User-Supplied Parameters

Below is a list of user-supplied parameters, which are inputs to the trial design optimizer.

- Number of study arms: 3.
- Type of outcome data: continuous.
- Subpopulation 1 proportion: 0.49.
- Familywise Type I error rate: 0.05.
- Maximum sample size: 4000.
- Maximum duration: 10.
- Enrollment rate per year (all arms combined): 500.
- Length of followup: 0.1.

- Optimization target: size.
- Censoring rate: 0.
- Minimum clinically important difference (MCID): 15.
- Assume precision gain from adjusting for baseline variables?: FALSE.
- If assumed precision gain from baseline variables, relative efficiency = 1.
- Distribution of outcomes assumed by the user, used to define power requirements and expected sample size and duration: see Tables **??** and **??**.
- Power requirements: see Table 1.2.

Scenario	weight	control_S1	control_S2	treatmentA_S1	treatmentA_S2	treatmentB_S1	treatmentB_S2
1	1.00	0.00	0.00	0.00	0.00	15.00	15.00

Table 1.1: User-supplied outcome mean under each scenario. In column names, "S1" and "S2" indicate subpopulation 1 and 2.

Scenario	control_S1	control_S2	treatmentA_S1	treatmentA_S2	treatmentB_S1	treatmentB_S2
1	3600.00	3600.00	3600.00	3600.00	3600.00	3600.00

Table 1.2: User-supplied variance (VAR) under each scenario. In column names, "S1" and "S2" indicate subpopulation 1 and 2.

Pow_H.0.1A.	Pow_H.0.2A.	Pow_H.0.1B.	Pow_H.0.2B.	Pow_Reject_all
0.00	0.00	0.00	0.00	0.80

Table 1.3: User-supplied power requirements. In column names, numbers 1 and 2 indicate the corresponding subpopulation.

2 1SEA design

2.1 Specification of the 1SEA design

Here we give the specification of the optimal 1SEA design.

Table 2.1 gives sample size and calendar time for each stage.

Table 2.2 gives the alpha allocation.

```
## Error in data.frame(Stage = 1:nrow(my.table), my.table):  arguments imply differing number of rows:
2, 0
```

```
## Error in data.frame(Stage = 1:nrow(alpha.allocation), alpha.allocation):  arguments imply differing
number of rows:  2, 0
```

2.2 Performance of the 1SEA design

Table 2.3 lists the empirical power.

Table 2.4 lists the empirical familywise Type I error (FWER).

Table 2.5 lists the empirical Type I error.

Table 2.6 gives additional performance metrics including bias, variance, mean-squared error, and confidence interval coverage. Under each scenario, for each treatment-subpopulation combination, the table lists the bias, variance (Var), and mean-squared error (MSE) for the estimator. It also lists the confidence interval coverage (CI.COVERAGE) and confidence interval inflation factor (CI.INFLATION), where the 95% nominal confidence interval is obtained by normal approximation. Confidence interval coverage denotes the probability that the truth is contained in the 95% confidence interval. Confidence interval inflation factor is the minimal number that the 95% nominal confidence interval needs to be multiplied by, in order to have actual 95% coverage. (CI.COVERAGE > 1 means that the 95% nominal confidence interval is anti-conservative.)

```
## Error in `colnames<-`(`*tmp*`, value = "Reject_All_False_Null_Hyp"):  attempt to set 'colnames'
on an object with less than two dimensions
## Error in `colnames<-`(`*tmp*`, value = "Power_"):  attempt to set 'colnames' on an object with
less than two dimensions
## Error in data.frame(Scenario = 1:n.scenarios, my.table):  arguments imply differing number of rows:
1, 0
```

```
## Error in `colnames<-`(`*tmp*`, value = "Familywise Type I Error Rate"):  attempt to set 'colnames'
on an object with less than two dimensions
## Error in names(my.table) <- "Familywise Type I error rate":  attempt to set an attribute on NULL
## Error in data.frame(Scenario = 1:n.scenarios, my.table):  arguments imply differing number of rows:
1, 0
```

```
## Error in `colnames<-`(`*tmp*`, value = "TypeIerror_"):  attempt to set 'colnames' on an object
with less than two dimensions
## Error in data.frame(Scenario = 1:n.scenarios, my.table):  arguments imply differing number of rows:
1, 0
```

```
## Error in `$<-.data.frame`(`*tmp*`, scenario, value = c(1L, 1L, 1L, 1L)): replacement has 4 rows,
data has 0
## Error in `$<-.data.frame`(`*tmp*`, subpopulation, value = c(1L, 2L, 1L, :  replacement has 4 rows,
data has 0
## Error in `$<-.data.frame`(`*tmp*`, treatment, value = c("A", "A", "B", :  replacement has 4 rows,
data has 0
## Error in data.frame(Scenario = scenario, Treatment = treatment, Subpop = subpopulation, :  object
'scenario' not found
## Error in display[ints] <- "d":  invalid subscript type 'list'
```

3 1SOA design

3.1 Specification of the 1SOA design

Here we give the specification of the optimal 1SOA design.

Table 3.1 gives sample size and calendar time for each stage.

Table 3.2 gives the alpha allocation.

```
## Error in data.frame(osoa.result$optima$per.stage.sample.sizes$total.n.per.stage):  object 'osoa.result'
not found
## Error in eval(expr, envir, enclos):  object 'osoa.result' not found
## Error in data.frame(Stage = 1:nrow(my.table), my.table):  arguments imply differing number of rows:
2, 0
```

```
## Error in matrix(osoa.result$parameters$alpha.allocation * ui.total.alpha, :  object 'osoa.result'
not found
## Error in data.frame(Stage = 1:nrow(alpha.allocation), alpha.allocation):  arguments imply differing
number of rows:  2, 0
```

3.2 Performance of the 1SOA design

Table 3.3 lists the empirical power.

Table 3.4 lists the empirical familywise Type I error (FWER).

Table 3.5 lists the empirical Type I error.

Table 3.6 gives additional performance metrics including bias, variance, mean-squared error, and confidence interval coverage. Under each scenario, for each treatment-subpopulation combination, the table lists the bias, variance (Var), and mean-squared error (MSE) for the estimator. It also lists the confidence interval coverage (CI.COVERAGE) and confidence interval inflation factor (CI.INFLATION), where the 95% nominal confidence interval is obtained by normal approximation. Confidence interval coverage denotes the probability that the truth is contained in the 95% confidence interval. Confidence interval inflation factor is the minimal number that the 95% nominal confidence interval needs to be multiplied by, in order to have actual 95% coverage. (CI.COVERAGE > 1 means that the 95% nominal confidence interval is anti-conservative.)

```
## Error in colnames(osoa.result$optima$conj.power) <- "Reject_All_False_Null_Hyp":  object 'osoa.result'
not found
## Error in cbind(osoa.result$optima$empirical.power, osoa.result$optima$conj.power):  object 'osoa.result'
not found
## Error in names(x) <- value:  'names' attribute [1] must be the same length as the vector [0]
## Error in data.frame(Scenario = 1:n.scenarios, my.table):  arguments imply differing number of rows:
1, 0
```

```
## Error in eval(expr, envir, enclos):  object 'osoa.result' not found
## Error in names(x) <- value:  'names' attribute [1] must be the same length as the vector [0]
## Error in names(my.table) <- "Familywise Type I error rate":  'names' attribute [1] must be the
same length as the vector [0]
## Error in data.frame(Scenario = 1:n.scenarios, my.table):  arguments imply differing number of rows:
1, 0
```

```
## Error in eval(expr, envir, enclos):  object 'osoa.result' not found
## Error in names(x) <- value:  'names' attribute [1] must be the same length as the vector [0]
## Error in data.frame(Scenario = 1:n.scenarios, my.table):  arguments imply differing number of rows:
1, 0
```

```
## Error in unlist(osoa.result$optima$bias.variance.mse.ci[i.scenario, i.arm.pop]):  object 'osoa.result'
not found
## Error in eval(expr, envir, enclos):  object 'osoa.result' not found
```

```
## Error in `$<-.data.frame`(`*tmp*`, scenario, value = c(1L, 1L, 1L, 1L)): replacement has 4 rows,
data has 0
## Error in `$<-.data.frame`(`*tmp*`, subpopulation, value = c(1L, 2L, 1L, :  replacement has 4 rows,
data has 0
## Error in `$<-.data.frame`(`*tmp*`, treatment, value = c("A", "A", "B", :  replacement has 4 rows,
data has 0
## Error in data.frame(Scenario = scenario, Treatment = treatment, Subpop = subpopulation, :  object
'scenario' not found
## Error in display[ints] <- "d":  invalid subscript type 'list'
```

4 2SEA design

4.1 Specification of the 2SEA design

Here we give the specification of the optimal 2SEA design.

Table 4.1 gives sample size and calendar time for each stage.

Table 4.2 gives the alpha allocation.

Table 4.3 gives the futility boundaries.

```
## Error in data.frame(tsea.result$optima$per.stage.sample.sizes$total.n.per.stage):  object 'tsea.result'
not found
## Error in eval(expr, envir, enclos):  object 'tsea.result' not found
## Error in data.frame(Stage = 1:nrow(my.table), my.table):  arguments imply differing number of rows:
2, 0
```

```
## Error in matrix(tsea.result$parameters$alpha.allocation * ui.total.alpha, :  object 'tsea.result'
not found
## Error in data.frame(Stage = 1:nrow(alpha.allocation), alpha.allocation):  arguments imply differing
number of rows:  2, 0
```

```
## Error in matrix(c(tsea.result$parameters$futility.boundaries, NA, NA), :  object 'tsea.result'
not found
## Error in names(futility.boundaries) <- paste0("Subpop", 1:n.subpopulation):  object 'futility.boundaries
not found
## Error in nrow(futility.boundaries):  object 'futility.boundaries' not found
```

4.2 Performance of the 2SEA design

Table 4.4 lists the empirical power.

Table 4.5 lists the empirical familywise Type I error (FWER).

Table 4.6 lists the empirical Type I error.

Table 4.7 lists the expected sample size.

Table 4.8 lists the expected duration.

Table 4.9 gives additional performance metrics including bias, variance, mean-squared error, and confidence interval coverage. Under each scenario, for each treatment-subpopulation combination, the table lists the bias, variance (Var), and mean-squared error (MSE) for the estimator. It also lists the confidence interval coverage (CI.COVERAGE) and confidence interval inflation factor (CI.INFLATION), where the 95% nominal confidence interval is obtained by normal approximation. Confidence interval coverage denotes the probability that the truth is contained in the 95% confidence interval. Confidence interval inflation factor is the minimal number that the 95% nominal confidence interval needs to be multiplied by, in order to have actual 95% coverage. (CI.COVERAGE > 1 means that the 95% nominal confidence interval is anti-conservative.)

Figure 4.1 gives the distribution of trial sample size and duration. Note that the x-axis scale is not linear, because sample sizes and durations are clustered at specific values (depending on when the trial stops).

```
## Error in colnames(tsea.result$optima$conj.power) <- "Reject_All_False_Null_Hyp":  object 'tsea.result'
not found
## Error in cbind(tsea.result$optima$empirical.power, tsea.result$optima$conj.power):  object 'tsea.result'
not found
## Error in names(x) <- value:  'names' attribute [1] must be the same length as the vector [0]
## Error in data.frame(Scenario = 1:n.scenarios, my.table):  arguments imply differing number of rows:
1, 0
```

```
## Error in eval(expr, envir, enclos):  object 'tsea.result' not found
## Error in names(x) <- value:  'names' attribute [1] must be the same length as the vector [0]
## Error in names(my.table) <- "Familywise Type I error rate":  'names' attribute [1] must be the
same length as the vector [0]
```

```
## Error in data.frame(Scenario = 1:n.scenarios, my.table):  arguments imply differing number of rows:
1, 0
```

```
## Error in eval(expr, envir, enclos):  object 'tsea.result' not found
## Error in names(x) <- value:  'names' attribute [1] must be the same length as the vector [0]
## Error in data.frame(Scenario = 1:n.scenarios, my.table):  arguments imply differing number of rows:
1, 0
```

```
## Error in eval(expr, envir, enclos):  object 'tsea.result' not found
## Error in data.frame(Scenario = c(1:n.scenarios, "average"), ESS = my.table):  arguments imply differing
number of rows:  2, 0
```

```
## Error in eval(expr, envir, enclos):  object 'tsea.result' not found
## Error in data.frame(Scenario = c(1:n.scenarios, "average"), Duration = my.table):  arguments imply
differing number of rows:  2, 0
```

```
## Error in unlist(tsea.result$optima$bias.variance.mse.ci[i.scenario, i.arm.pop]):  object 'tsea.result'
not found
## Error in eval(expr, envir, enclos):  object 'tsea.result' not found
## Error in `$<-.data.frame`(`*tmp*`, scenario, value = c(1L, 1L, 1L, 1L)): replacement has 4 rows,
data has 0
## Error in `$<-.data.frame`(`*tmp*`, subpopulation, value = c(1L, 2L, 1L, :  replacement has 4 rows,
data has 0
## Error in `$<-.data.frame`(`*tmp*`, treatment, value = c("A", "A", "B", :  replacement has 4 rows,
data has 0
## Error in data.frame(Scenario = scenario, Treatment = treatment, Subpop = subpopulation, :  object
'scenario' not found
## Error in display[ints] <- "d":  invalid subscript type 'list'
```

```
## Error in eval(expr, envir, enclos):  object 'tsea.result' not found
## Error in distribution.of.trials[is.na(distribution.of.trials)] <- 0:  object 'distribution.of.trials'
not found
## Error in aggregate(distribution.of.trials$proportion, by = list(distribution.of.trials$scenario,
:  object 'distribution.of.trials' not found
## Error in names(plot.table.ss) <- c("scenario", "sample.size", "proportion"):  object 'plot.table.ss'
not found
## Error in aggregate(distribution.of.trials$proportion, by = list(distribution.of.trials$scenario,
:  object 'distribution.of.trials' not found
## Error in names(plot.table.duration) <- c("scenario", "duration", "proportion"):  object 'plot.table.dura
not found
## Error in ggplot(aes(y = proportion, x = factor(sample.size), fill = factor(scenario)), :  object
'plot.table.ss' not found
## Error in ggplot(aes(y = proportion, x = factor(neaten(duration, 2)), fill = factor(scenario)),
:  object 'plot.table.duration' not found
```

```
## Error in arrangeGrob(...):  object 'tsea.bar.ss' not found
```

5 2SOA design

5.1 Specification of the 2SOA design

Here we give the specification of the optimal 2SOA design.

Table 5.1 gives sample size and calendar time for each stage.

Table 5.2 gives the alpha allocation.

Table 5.3 gives the futility boundaries.

```
## Error in data.frame(tsoa.result$optima$per.stage.sample.sizes$total.n.per.stage):  object 'tsoa.result'
not found
## Error in eval(expr, envir, enclos):  object 'tsoa.result' not found
## Error in data.frame(Stage = 1:nrow(my.table), my.table):  arguments imply differing number of rows:
2, 0
```

```
## Error in matrix(tsoa.result$parameters$alpha.allocation * ui.total.alpha, :  object 'tsoa.result'
not found
## Error in data.frame(Stage = 1:nrow(alpha.allocation), alpha.allocation):  arguments imply differing
number of rows:  2, 0
```

```
## Error in matrix(c(tsoa.result$parameters$futility.boundaries, NA, NA), :  object 'tsoa.result'
not found
## Error in names(futility.boundaries) <- paste0("Subpop", 1:n.subpopulation):  object 'futility.boundaries
not found
## Error in nrow(futility.boundaries):  object 'futility.boundaries' not found
```

5.2 Performance of the 2SOA design

Table 5.4 lists the empirical power.

Table 5.5 lists the empirical familywise Type I error (FWER).

Table 5.6 lists the empirical Type I error.

Table 5.7 lists the expected sample size.

Table 5.8 lists the expected duration.

Table 5.9 gives additional performance metrics including bias, variance, mean-squared error, and confidence interval coverage. Under each scenario, for each treatment-subpopulation combination, the table lists the bias, variance (Var), and mean-squared error (MSE) for the estimator. It also lists the confidence interval coverage (CI.COVERAGE) and confidence interval inflation factor (CI.INFLATION), where the 95% nominal confidence interval is obtained by normal approximation. Confidence interval coverage denotes the probability that the truth is contained in the 95% confidence interval. Confidence interval inflation factor is the minimal number that the 95% nominal confidence interval needs to be multiplied by, in order to have actual 95% coverage. (CI.COVERAGE > 1 means that the 95% nominal confidence interval is anti-conservative.)

Figure 5.1 gives the distribution of trial sample size and duration. Note that the x-axis scale is not linear, because sample sizes and durations are clustered at specific values (depending on when the trial stops).

```
## Error in colnames(tsea.result$optima$conj.power) <- "Reject_All_False_Null_Hyp":  object 'tsea.result'
not found
## Error in cbind(tsea.result$optima$empirical.power, tsea.result$optima$conj.power):  object 'tsea.result'
not found
## Error in names(x) <- value:  'names' attribute [1] must be the same length as the vector [0]
## Error in data.frame(Scenario = 1:n.scenarios, my.table):  arguments imply differing number of rows:
1, 0
```

```
## Error in eval(expr, envir, enclos):  object 'tsoa.result' not found
## Error in names(x) <- value:  'names' attribute [1] must be the same length as the vector [0]
## Error in names(my.table) <- "Familywise Type I error rate":  'names' attribute [1] must be the
same length as the vector [0]
```

```
## Error in data.frame(Scenario = 1:n.scenarios, my.table):  arguments imply differing number of rows:
1, 0
```

```
## Error in eval(expr, envir, enclos):  object 'tsoa.result' not found
## Error in names(x) <- value:  'names' attribute [1] must be the same length as the vector [0]
## Error in data.frame(Scenario = 1:n.scenarios, my.table):  arguments imply differing number of rows:
1, 0
```

```
## Error in eval(expr, envir, enclos):  object 'tsoa.result' not found
## Error in data.frame(Scenario = c(1:n.scenarios, "average"), ESS = my.table):  arguments imply differing
number of rows:  2, 0
```

```
## Error in eval(expr, envir, enclos):  object 'tsoa.result' not found
## Error in data.frame(Scenario = c(1:n.scenarios, "average"), Duration = my.table):  arguments imply
differing number of rows:  2, 0
```

```
## Error in unlist(tsoa.result$optima$bias.variance.mse.ci[i.scenario, i.arm.pop]):  object 'tsoa.result'
not found
## Error in eval(expr, envir, enclos):  object 'tsoa.result' not found
## Error in `$<-.data.frame`(`*tmp*`, scenario, value = c(1L, 1L, 1L, 1L)): replacement has 4 rows,
data has 0
## Error in `$<-.data.frame`(`*tmp*`, subpopulation, value = c(1L, 2L, 1L, :  replacement has 4 rows,
data has 0
## Error in `$<-.data.frame`(`*tmp*`, treatment, value = c("A", "A", "B", :  replacement has 4 rows,
data has 0
## Error in data.frame(Scenario = scenario, Treatment = treatment, Subpop = subpopulation, :  object
'scenario' not found
## Error in display[ints] <- "d":  invalid subscript type 'list'
```

```
## Error in eval(expr, envir, enclos):  object 'tsoa.result' not found
## Error in distribution.of.trials[is.na(distribution.of.trials)] <- 0:  object 'distribution.of.trials'
not found
## Error in aggregate(distribution.of.trials$proportion, by = list(distribution.of.trials$scenario,
:  object 'distribution.of.trials' not found
## Error in names(plot.table.ss) <- c("scenario", "sample.size", "proportion"):  object 'plot.table.ss'
not found
## Error in aggregate(distribution.of.trials$proportion, by = list(distribution.of.trials$scenario,
:  object 'distribution.of.trials' not found
## Error in names(plot.table.duration) <- c("scenario", "duration", "proportion"):  object 'plot.table.dura
not found
## Error in ggplot(aes(y = proportion, x = factor(sample.size), fill = factor(scenario)), :  object
'plot.table.ss' not found
## Error in ggplot(aes(y = proportion, x = factor(neaten(duration, 2)), fill = factor(scenario)),
:  object 'plot.table.duration' not found
```

```
## Error in arrangeGrob(...):  object 'tsoa.bar.ss' not found
```

6 Glossary

- 1SEA: the optimal 1-stage design with equal alpha allocation to all hypotheses;
- 1SOA: the optimal 1-stage design with alpha allocation chosen by optimization;
- 2SEA: the optimal 2-stage design with equal alpha allocation to all hypotheses and futility boundaries chosen by optimization;
- 2SOA: the optimal 2-stage design with interim analysis timing, futility boundaries, and alpha allocation all chosen by optimization.
- A, B, C: treatment arm A, B; control arm C.
- A1: treatment arm A and subpopulation 1. Similar for C1.
- Bias: bias of the estimator at the end of trial.
- CI.COVERAGE: confidence interval coverage, which is the coverage of a nominal 95% confidence interval centered at the estimator at the end of trial, obtained by normal approximation.
- CI.INFLATION: confidence interval inflation factor, which is the smallest positive number γ such that the confidence interval has exact 95% coverage if the interval is inflated by γ.
- Delta1, Delta2: Average treatment effects (difference between mean outcomes under treatment versus control) in subpopulations 1 and 2, respectively.
- ESS: expected sample size.
- FWER: familywise error rate.
- HR: hazard rate.
- MSE: mean squared error of the estimator at the end of trial.
- Scenario: user-supplied outcome distribution under all treatment and subpopulation combinations.
- VAR: variance.
- Subpop: subpopulation.
- Var: variance of the estimator at the end of trial.