Notes On Interim Analyses

Larry F. León, PhD

October 9, 2020

```
rm(list = ls())
require(gsDesign)
require(knitr)
require(survival)
require(Hmisc)
require(kableExtra)
require(formatR)

source("../R/km_functions_twosample_weighting.R")
source("../R/km_plot_twosample.R")

source("../R/interim_analyses_core-functions_v0.R")
source("../R/interim_datasets_v0.R")
source("../R/get.analyses_functions_v0.R")
source("../R/get.analyses_CR_functions_v0.R")
source("../R/oc.interims_v0.R")
```

1 Configuring `gsDesign` to closely match "Study A" interim OS analysis

In the following we produce a design to closely match "Study A" for the interim OS analyses assuming a significance level of $\alpha = 0.02$ (1-sided).

The design calls for $n = 900$ patients with three interim analyses comparing OS in the ITT population ("all-comer") at alpha-level 0.02 to achieve 88% power. The interims are conducted when 344, 536, and 688 events are observed corresponding to information fractions of 50%, 78%, and 100% (resp.). The hazard ratio crossing boundaries are 0.717, 0.813, and 0.851 with corresponding two-sided p-value boundaries of 0.002, 0.0162, and 0.0348. The expected interim analysis timings are estimated as 30, 42, and 58 months.

We use the `gsDesign` study design parameters as inputs (sample size, target number of looks, and crossing boundaries) to our simulation functions to calculate further details on design operating characteristics. The simulation code is based on two functions. The function `getdata.looks` simulates the interim analysis datasets based on the interim target number of events. For example, with 3 looks based on event triggers (target number of events) d1, d2 and d3 (say), there will be 3 sets of data "Look1", "Look2", and "Look3". For 10,000 simulations (say), Look1 will contain the 10,000 simulated first look analysis statistics (Cox HR estimate, analysis time, 1-sided log-rank Z statistic and p-value, median estimates, and drop-out proportion); and similarly for Look2 and Look3. For the interim analysis trigger, we specify that the target number of events are "strictly met" in the sense that at least the target will be achieved with possible "overage" depending on discretess in the accumlation of events. This is implemented with the following options: Target events are checked within follow-up intervals tau.min and tau.max of length tau.seq (For example, every month (tau.seq=1/12) between 0 (tau.min=0) and 200 months (tau.max=200)). If the target is contained within two intervals, then these intervals are further searched (divided into lengths of dexact.len) to find the exact match or closest to achieve at least the target. Several accrual pattern options are available with the default being analogous to the EAST algorithm with a ramp-up phase. Default censoring is exponential. Survival times are generated according

to Weibull distributions with the default being exponential (but can be according to different Weibull shape and scale parameters for the two treatment arms).

With these interim analysis datasets the function `OC.interims` will then calculate statistics (e.g., crossing probabilities) based on the input crossing boundaries such as OBF per `gsDesign` (or any other boundaries).

Code is here: `https://github.com/larry-leon/interim-analyses`.

See also the comprehensive R package `rpact`. In particular there is a vignette by Wassmer, Pahlke, and Wolbers `https://www.rpact.org/vignettes` comparing rpact with `gsDesign` [So this can be used as an indirect comparison with rpact :)].

1.1 Setting up `gsDesign`

A `gsDesign` configuration that closely matches the "Study A" protocol is described below. The gsDesign specification calls for $n = 906$ so the sample size slightly differs. Assumptions on accrual patterns may slightly differ.

Remark 1. Here we use the template found in the source for the gsDesign R package in `/inst/doc/gsSurvTemplate.rnw`

```
# Test type (one-sided [superiority])
test.type <- 1
# study duration T can be set to NULL if you want to fix enrollment and vary
# study duration
T <- 58
# follow-up duration of last patient enrolled
minfup <- 32
# Enrollment duration will be 58-32 k looks
k <- 3
# timing of interim analyses (k-1 increasing numbers >0 and <1)
timing <- c(0.5, 0.78, 1)
# efficacy bound spending function
sfu <- "OF"
# power
beta <- 1 - 0.88
# Uniform accrual with enrollment duration of 26 months
R <- c(1)  # relative enrollment rates during above periods
gamma <- c(906/26)
# median control time-to-event
median <- 16
# exponential dropout rate per unit of time
eta <- 0.05/12
# hypothesized experimental/control hazard ratio
hr <- 0.78
# null hazard ratio (1 for superiority, >1 for non-inferiority)
hr0 <- 1
# Type I error (1-sided)
alpha <- 0.02

# time units
timename <- "months"
timename1 <- "month"
# endpoint name
ep <- "overall survival"
```

```
# make a string with enrollment rates
# (assumes gamma is a single value or vector)
```

```
x<-x.gsd
nR <- length(x$R)
if (nR==1){enrolrates <- paste("constant at a rate of ",
                          round(gamma,1),"per",timename1,"")
} else{
enrolrates <- paste(c("piecewise constant at rates of ",
 paste(round(as.vector(x$gamma),1)," for ",timename," ",
       cumsum(c(0,x$R[1:(nR-1)])),
       "-",cumsum(x$R),collapse=", "),sep=""),collapse="")
}
```

1.2 Summary of operating characteristics

Analysis	Value	Efficacy
IA 1: 50%	Z	2.9931
N: 906	p (1-sided)	0.0014
Events: 345	˜HR at bound	0.7242
Month: 27	P(Cross) if HR=1	0.0014
	P(Cross) if HR=0.78	0.2462
IA 2: 78%	Z	2.3964
N: 906	p (1-sided)	0.0083
Events: 537	˜HR at bound	0.8131
Month: 39	P(Cross) if HR=1	0.0088
	P(Cross) if HR=0.78	0.6893
Final	Z	2.1164
N: 906	p (1-sided)	0.0172
Events: 689	˜HR at bound	0.8510
Month: 58	P(Cross) if HR=1	0.0200
	P(Cross) if HR=0.78	0.8800

Table 1: One-sided group sequential design with 3 analyses, time-to-event outcome with sample size 906 and 689 events required, 88 percent power, 2 percent (1-sided) Type I error to detect a hazard ratio of 0.78. Enrollment and total study durations are assumed to be 26 and 58 months, respectively. Efficacy bounds derived using a O'Brien-Fleming boundary.

For a comparative trial we consider a 2-arm group sequential design with overall survival as the primary endpoint as shown in Table 1. Timing, number of events, sample size, boundaries (Z-values, nominal p-values, approximate hazard ratios) are shown as well as the probability of crossing study boundaries under the null and alternate hypotheses. The median time-to-event is assumed to be 16 months in the control group. The trial is designed to demonstrate superiority of experimental treatment over control with an assumed hazard ratio of 0.78. The total sample size is 906 and a total of 689 endpoints is required for the final analysis. Planned recruitment duration is 26 months and the minimum follow-up planned is 32 months. Thus, the total expected study duration is 58 months. Enrollment is assumed to be constant at a rate of 34.8 per month . The assumed dropout rate is 0.4166667% per month. .

2 Operating characteristics via simulations

2.1 Code setup

```
k <- 3
# From gsSuru Sample size
```

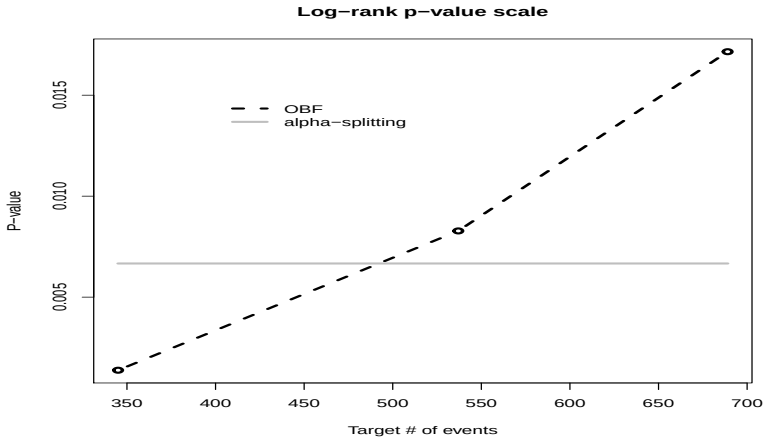
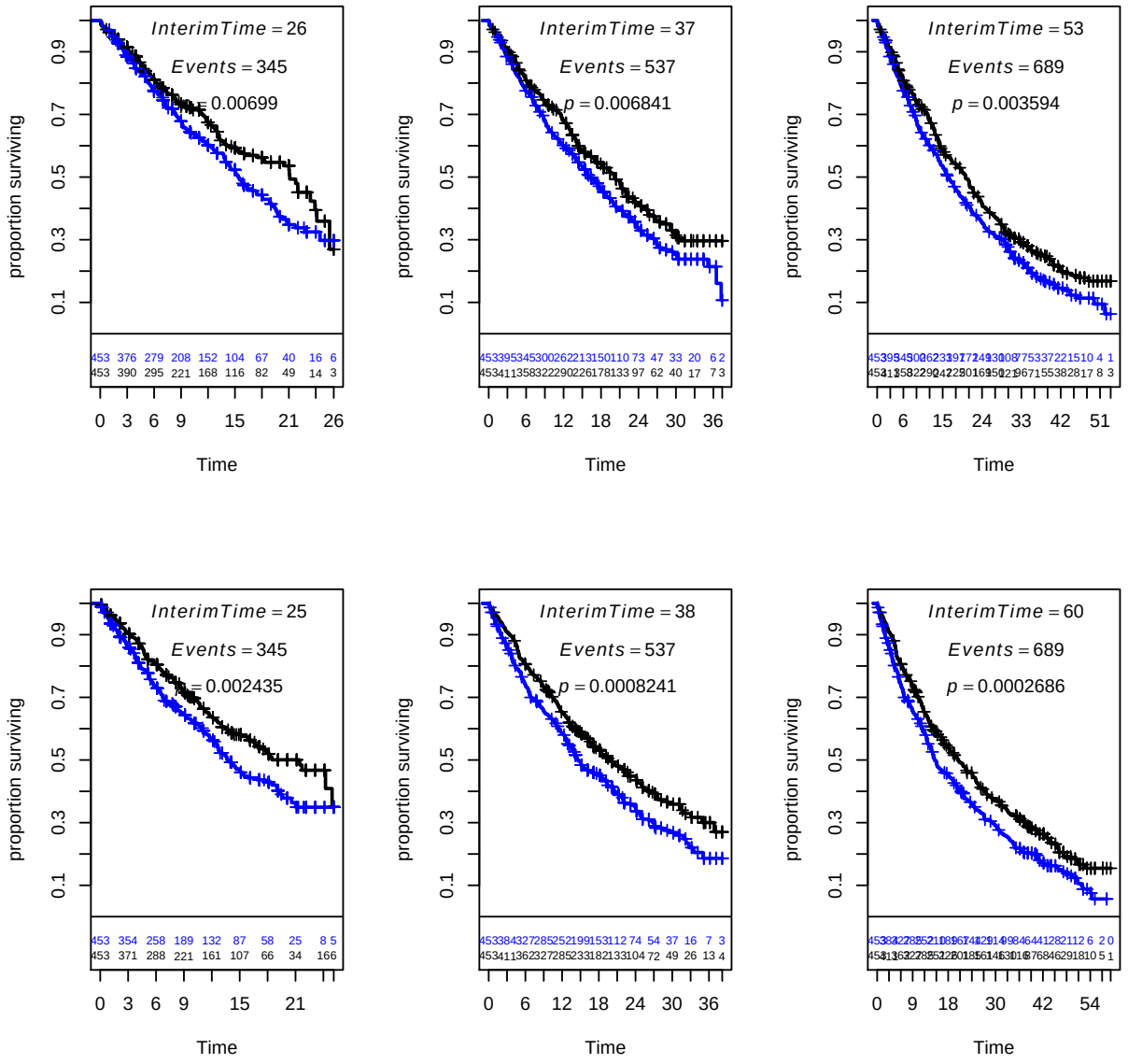


Figure 1: "Study A" Interim Analysis Crossing boundaries: P-value (1-sided) boundaries = $0.0013808, 0.0082787, 0.0171539$; alpha-splitting $= 0.0067, 0.0067, 0.0067$

```
N <- ceiling(sum(x$eNE[k, ] + x$eNC[k, ]))
# Events for final analysis
d.FA <- ceiling(sum(x$eDE[k, ] + x$eDC[k, ]))

# gsDesign crossing boundaries
Z.looks <- c(x$upper$bound)
alpha.looks <- c(1 - pnorm(Z.looks))

# Event-proportion interim triggers; Target # of events
d.looks <- round(d.FA * c(0.5, 0.78, 1), 1)
d.looks <- floor(d.looks + 0.5)

# Median control (in months)
m0 <- 16
```

```
# The first 2 simulations will be plot (specified by option kmplot.nsims=2);
# For each simulation the 3 interim analyses are displayed
options(digits=3)
par(mfrow=c(2,3))
# Simulations under hr=0.78
# Note: 10,000 simulations (n.sims=10000) are previously run and stored

# Projected enrollment (Ramp-up phase)
# Note: Ramp.up needs to be integer
# I'm not aware of the ramp-up assumptions used for "Study A"
# The following seems to approximate the reported timing estimates
Ramp.up<-c(34,35)
# n=34 in first month; n=35 in second month
nRamp.after<-36
# n=36 per month (after 3 months)

accruals<-get.accrual(N=N,rand.ratio=1,Accrue.type="EastProjection",Ramp.up=Ramp.up,nRamp.after=nRamp.a
AC.0<-accruals$AC.0
```

```
AC.1<-accruals$AC.1

n0<-length(AC.0)
n1<-length(AC.1)
i0<-rep(1,n0)
i1<-rep(1,n1)
a0<-sort(AC.0)
a1<-sort(AC.1)
a.both<-c(a0,a1)
i.both<-c(i0,i1)

#plot(sort(a.both),cumsum(i.both),xlab="Weeks",ylab="Cumulative Enrollment",type="s")
#lines(a0,cumsum(i0),type="s",lty=1,lwd=2,col="grey")
#lines(a1,cumsum(i1),type="s",lty=1,lwd=2,col="blue")

accrue.time<-ceiling(max(c(AC.0,AC.1)))

cat("Assumed accrual time with ramp-up (in months)",c(accrue.time),"\n")

## Assumed accrual time with ramp-up (in months) 26

# Here run 10 sims for illustration

sims<-10
# We plot the 3 interim analysis KMs for the first 2 simulations (kmplot.nsims=2)
# and by default the code will output the first dataset)

df.alt1<-getdata.looks(sims=sims,med0=m0,hr=0.78,
d.looks=d.looks,
S.type="Weibull",
sim.status=TRUE,
tau.min=0,tau.max=100,tau.seq=1/12,
max.follow=100,
d.exact=TRUE,dexact.len=100,
quant=0.5,kmplot.nsims=2,
AC1=AC.1, AC0=AC.0,Accrue.type='EastProjection',
drop.type="east",dropout.0=0.05,dropout.1=0.05,tau.drop=12)

## N1= 453
## N0= 453
## Number of looks= 3
## Note: Data-cuts are from tau.first,tau.max= 0 100
## 10% sims done
## Sims and time (mins) so far= 1 0.00783
## # Estimated mins= 0.0783
## # Estimated mins left= 0.0705
```




```
## 50% sims done
## Sims and time (mins) so far= 5 0.0167
## # Estimated mins= 0.0333
## # Estimated mins left= 0.0167
## 90% sims done
## Sims and time (mins) so far= 9 0.0258
## # Estimated mins= 0.0287
## # Estimated mins left= 0.00287
## First 10 sims done
## Sims and time (mins) so far= 10 0.0278
## # Estimated mins= 0.0278
## # Estimated mins per 1000= 2.78
## # Accrual Type EastProjection
## # Accrual Time= 26
```

```
## # Weibull scale parameters (C,T)= 23.1 29.6
## # Weibull shape parameters (C,T)= 1 1
## # Shape parameters in terms of exponential model (C,T)= 0.0433 0.0338
## # Medians (Cntrl,Exp) = 16 20.5
## # HR=,n1,n2=,sims= 0.78 453 453 10
## # Target dropout rate (Cntrl,Exp)= 0.05 0.05
## # Control,Exp drop-out parameters terms of exponential model (C,T)= 0.00427 0.00427
## # Note: these are analogous to etaC and etaE in gsDesign package
## # % drop-out (Control,Exp,All): Final Analysis (no stopping) 0.0929 0.13 0.111
## # Max events=d(tau.max)
## # Mean max events= 795
## Simulations= 10
## Time (min)= 0.0278
## # Estimated minutes per 1,000 simulations= 2.78

# Note: Previous run with 10,000 simulations is stored
#save(df.alt1,file="output/ex1_alt1.Rdata")
```

```
# Under null Not run (eval=FALSE)
sims <- 10000
df.null <- getdata.looks(sims = sims, med0 = m0, hr = 1, d.looks = d.looks, S.type = "Weibull",
    sim.status = TRUE, tau.min = 0, tau.max = 100, tau.seq = 1/12, max.follow = 100,
    d.exact = TRUE, dexact.len = 100, quant = 0.5, kmplot.nsims = 2, AC1 = AC.1,
    AC0 = AC.0, Accrue.type = "EastProjection", drop.type = "east", dropout.0 = 0.05,
    dropout.1 = 0.05, tau.drop = 12)


# save(df.null,file='output/ex1_null.Rdata')
```

```
# loads df.alt1
load("output/ex1_alt1.Rdata")
# loads df.null
load("output/ex1_null.Rdata")

df.names <- c("null", "alt1")

stats <- c("Look", "d", "p(cross)", "mDD", "t(cross)", "t(L)", "t(U)", "Avg(N1)",
    "N1(L)", "N1(U)", "% Reject H0", "X-ByNow")

main.caption <- c("Operating characteristics under standard exponential data generating model; Calculati
                \\label{tab:crm1oc}")

footnote.caption <- c("E denotes the target events;  p* are efficacy crossing boundaries (p-value scale)

names.table <- c("Look", "E", "p*", "mDD", "Timing", "Sample Size", "Rej-H0", "Eff-Yet")

df.oc <- NULL
for (nn in 1:length(df.names)) {
    dfa <- switch(nn, df.null, df.alt1, df.alt2, df.alt3, df.alt4, df.alt5)
    name.dgm <- df.names[nn]
    crossings <- OC.interims(data.looks = dfa, d.looks = d.looks, alpha.looks = alpha.looks,
        direction = "GT", details = FALSE)
```

```
    df.obf <- crossings$result
    df.obf <- df.obf[, c(stats)]
    assign(paste("cross.eff", name.dgm, sep = "."), crossings$result[, c("% Reject H0")])
    assign(paste("dfstop.sizes", name.dgm, sep = "."), crossings$data.stopped[, c("N1.IA",
        "N0.IA", "N.IA")])

    df.oc <- rbind(df.oc, df.obf)
}
rownames(df.oc) <- c()
df.oc <- round(df.oc, 3)
df.oc_new <- df.oc
# Include CIs for N in parenthesis

# Timings
loc1 <- which(colnames(df.oc) == stats[5])
loc2 <- which(colnames(df.oc) == stats[6])
loc3 <- which(colnames(df.oc) == stats[7])

df.oc_new <- df.oc[, -c(loc2, loc3)]

for (jj in 1:nrow(df.oc)) {
    a1 <- df.oc[jj, c(loc1, loc2, loc3)]
    a1 <- round(a1, digits = 1)
    res1 <- paste(a1[1], a1[2], sep = c(" ("))
    res2 <- paste(res1, a1[3], sep = c(","))
    res2 <- paste0(res2, ")")
    df.oc_new[jj, 5] <- c(res2)
}


# Sample size
loc1 <- which(colnames(df.oc_new) == stats[8])
loc2 <- which(colnames(df.oc_new) == stats[9])
loc3 <- which(colnames(df.oc_new) == stats[10])

df.oc_new2 <- df.oc_new[, -c(loc2, loc3)]

for (jj in 1:nrow(df.oc_new)) {
    a1 <- as.numeric(df.oc_new[jj, c(loc1, loc2, loc3)])
    a1 <- round(a1, digits = 0)
    res1 <- paste(a1[1], a1[2], sep = c(" ("))
    res2 <- paste(res1, a1[3], sep = c(","))
    res2 <- paste0(res2, ")")
    df.oc_new2[jj, 6] <- c(res2)
}
```

Note: Minutes for simulations are 20.467 and 20.672 for the alternative and null models, respectively.

Table 2: Operating characteristics under standard exponential data generating model; Calculations are based on 10,000 simulations.

Look	E	p*	mDD	Timing	Sample Size	Rej-H0	Eff-Yet
Under null: HR=1.0							
1	345	0.001	0.723	24.8 (23.6,26)	448 (430,453)	0.001	0.001
2	537	0.008	0.813	35.6 (33.5,37.7)	453 (453,453)	0.007	0.008
3	689	0.017	0.851	51.3 (47.6,55.5)	453 (453,453)	0.011	0.019
Under alternative: Avg HR=0.78							
1	345	0.001	0.724	26.2 (24.9,27.6)	453 (448,453)	0.246	0.246
2	537	0.008	0.813	38.6 (36.2,41)	453 (453,453)	0.446	0.692
3	689	0.017	0.851	57.3 (52.8,62.3)	453 (453,453)	0.191	0.883

Note.

E denotes the target events; p* are efficacy crossing boundaries (p-value scale); mDD denotes the empirical mdd (largest point estimate corresponding to efficacy boundary crossings); Timing (weeks) denotes the interim analysis timing, Avg (95% CI); Sample Size denotes the experimental interim sample size, Avg (95% CI); Rej-H0 denotes the incremental efficacy stopping probabilitiy; Eff-Yet denotes the cumulative efficacy crossing probability (crossing by the time of the look).

Under the null: The overall type-1 error is 1.93%; The average total sample size is 906 (95% CI=906, 906).

Under the alternative: The overall power is 88.3%; The average total sample size is 906 (95% CI=906, 906).

The operating characteristics summarized by gsDesign are —

Analysis	Value	Efficacy
IA 1: 50%	Z	2.9931
N: 906	p (1-sided)	0.0014
Events: 345	˜HR at bound	0.7242
Month: 27	P(Cross) if HR=1	0.0014
	P(Cross) if HR=0.78	0.2462
IA 2: 78%	Z	2.3964
N: 906	p (1-sided)	0.0083
Events: 537	˜HR at bound	0.8131
Month: 39	P(Cross) if HR=1	0.0088
	P(Cross) if HR=0.78	0.6893
Final	Z	2.1164
N: 906	p (1-sided)	0.0172
Events: 689	˜HR at bound	0.8510
Month: 58	P(Cross) if HR=1	0.0200
	P(Cross) if HR=0.78	0.8800

Table 3: One-sided group sequential design with 3 analyses, time-to-event outcome with sample size 906 and 689 events required, 88 percent power, 2 percent (1-sided) Type I error to detect a hazard ratio of 0.78. Enrollment and total study durations are assumed to be 26 and 58 months, respectively. Efficacy bounds derived using a O'Brien-Fleming boundary.

3 Timing and hazard ratio estimation properties

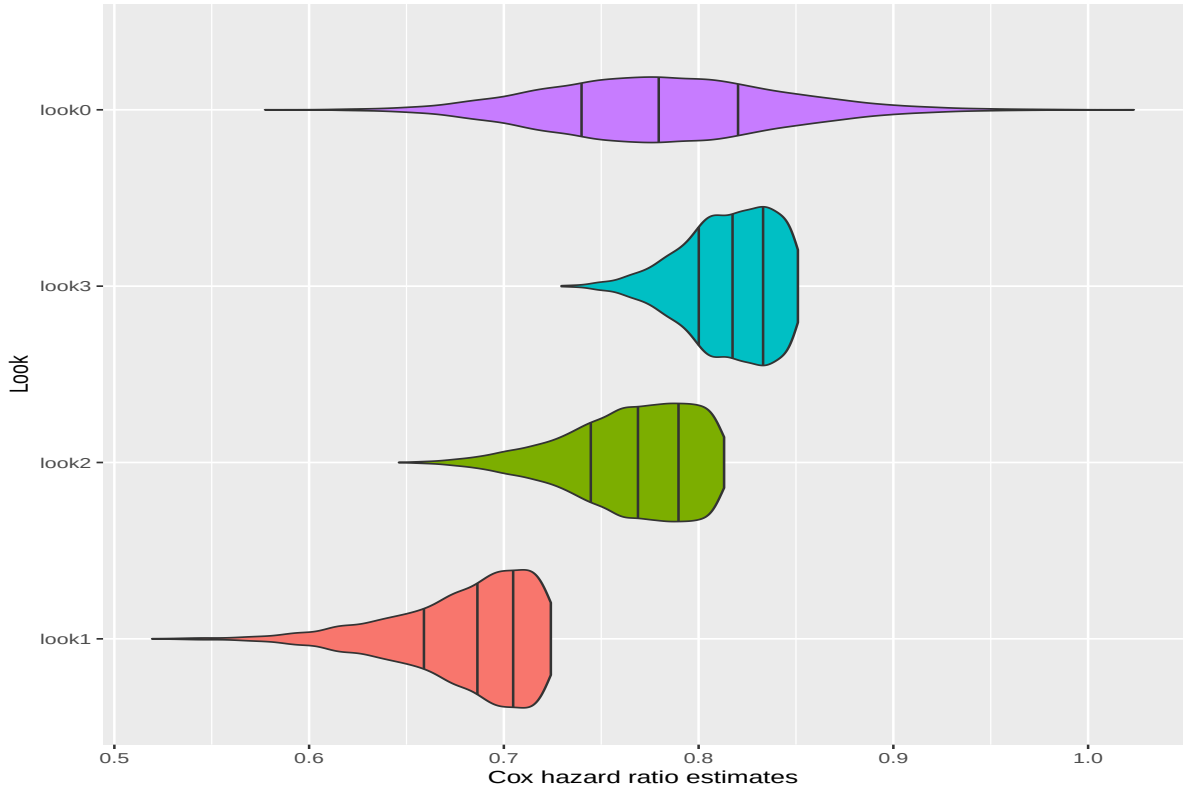


Figure 2: Cox hazard ratio estimates: Look 1 includes look=1 sequences that stopped at look=1 for efficacy; Look 2 includes look=2 sequences that did NOT stop at look=1 (for efficacy or futility) but stopped at look=2 for efficacy; and Look 3 includes look=3 sequences that did NOT stop at look=1 or look=2 but stopped at look=3 for efficacy. Look 0 denotes all final analyses where interim analyses were not conducted (i.e., final analyses for all data sequences). The vertical lines denote the quartiles (25%, 50%, and 75%).

4 Some notes on the target number of events, dropouts, and medians

```
df.pow <- df.alt1

df.look1 <- df.pow$Out$Look1
df.look2 <- df.pow$Out$Look2
df.look3 <- df.pow$Out$Look3
# print(summary(df.look1)) Note: here m.diff.IA and se.diff.IA denote RMST
# estimates; There is an option for this (get.mean=TRUE); BUt default=FALSE since
# timing is increased

df.look1$look = 1
df.look2$look = 2
df.look3$look = 3

df.looks <- rbind(df.look1, df.look2, df.look3)
```

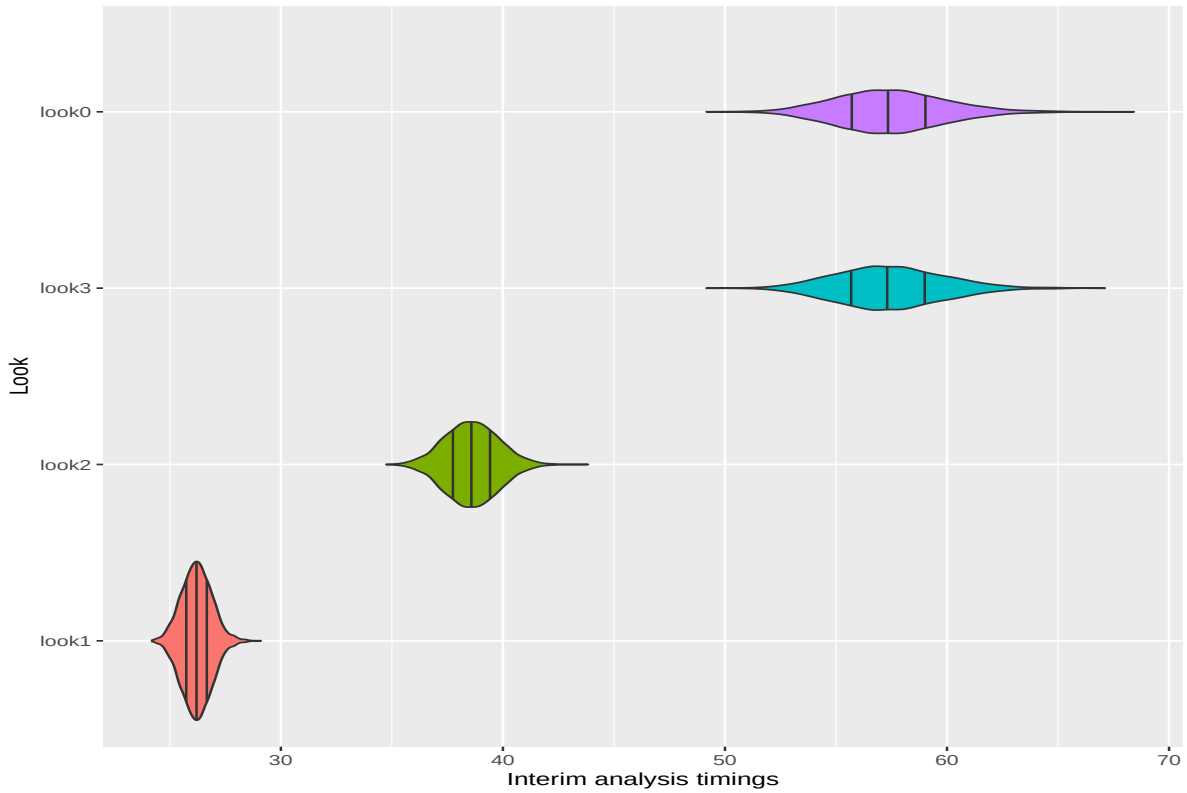


Figure 3: Interim analysis timings: Look 1 includes look=1 sequences that stopped at look=1 for efficacy; Look 2 includes look=2 sequences that did NOT stop at look=1 (for efficacy or futility) but stopped at look=2 for efficacy; and Look 3 includes look=3 sequences that did NOT stop at look=1 or look=2 but stopped at look=3 for efficacy. Look 0 denotes all final analyses where interim analyses were not conducted (i.e., final analyses for all data sequences). The vertical lines denote the quartiles (25%, 50%, and 75%).

4.1 Look=1 details

The target number of events at the first interim look are met at rate $\approx 99.2\%$. The proportion below are 0% and the proportion that exceed are 0.9% (The maximum number of events over target is 1 events). The average drop-out rates for the control and experimental groups by the time of the first look are 9.8% and 13.2%, respectively. The medians for control and experimental are reached (estimable) at rates 100% and 98%, respectively

4.2 Look=2 details

The target number of events at the second interim look (not stopping prior to) are met at rate $\approx 99.6\%$. The proportion below are 0% and the proportion that exceed are 0.4% (The maximum number of events over target is 1 events). The average drop-out rates for the control and experimental groups by the time of the second look (not stopping prior to) are 9.8% and 13.2%, respectively. The medians for control and experimental are reached (estimable) at rates 100% and 100%, respectively

4.3 Look=3 (final analysis) details

The target number of events at the third interim look (not stopping prior ro) are met at rate $\approx 99.8\%$. The proportion below are 0% and the proportion that exceed are 0.2% (The maximum number of events over target is 1 events). The average drop-out rates for the control and experimental groups by the time of the first look are 9.8% and 13.2%, respectively. The medians for control and experimental are reached (estimable) at rates 100% and 100%, respectively

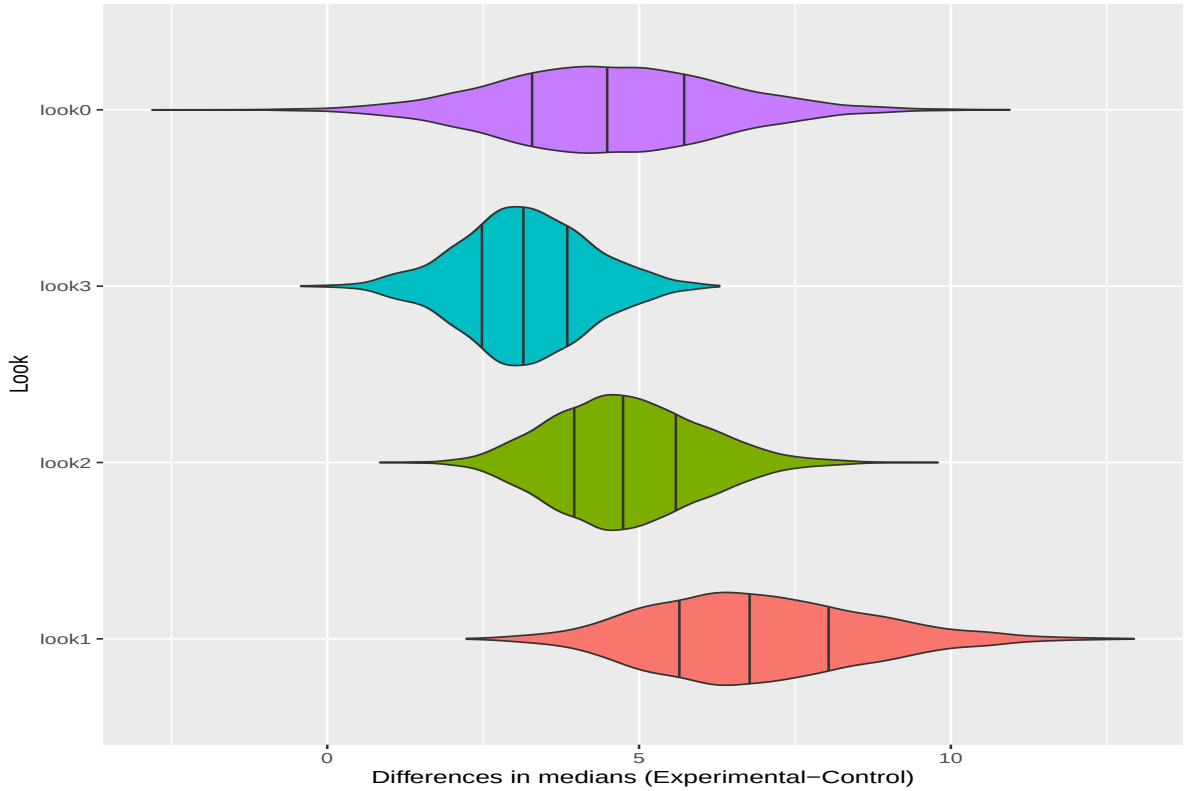


Figure 4: Differences in medians, where estimable (Experimental-Control): Look 1 includes look=1 sequences that stopped at look=1 for efficacy; Look 2 includes look=2 sequences that did NOT stop at look=1 (for efficacy or futility) but stopped at look=2 for efficacy; and Look 3 includes look=3 sequences that did NOT stop at look=1 or look=2 but stopped at look=3 for efficacy. Look 0 denotes all final analyses where interim analyses were not conducted (i.e., final analyses for all data sequences). The vertical lines denote the quartiles (25%, 50%, and 75%).

4.4 Summary of un-conditional (with respect to stopping) statistics across looks

```
summary.looks<-summary(look~tau.IA+Zlr.IA+exp(bhat.IA)+m0.IA+m1.IA+drop0.IA+drop1.IA,
                       method="reverse", overall=FALSE, test=FALSE, data=df.looks)
w<-latex(summary.looks,file="output/summary_looks.tex",digits=2, title='', where="",
caption="Summary details for interim looks. Statistics are un-conditional with respect to stopping:
$\\tau_{IA}$ denotes the interim timming; $Zlr.IA$ the interim z-crossing boundary;
$\\exp(bhat.IA)$ denotes the interim hazard ratio; m0.IA (m1.IA) denotes the median for control (experi
and drop0.IA (drop1.IA) denotes the dropout proportion for control (experimental).",
label="table:summary_looks")
```

5 Some potential differences with gsDesign

I am only recently using the gsDesign package. I suspect there may be differences in the accrual algorithms (gsDesign as specified implements constant accrual of 34.82 in 26 months; My algorithm requires integers and attempts to divide evenly between treatments within accrual windows). It also appears that my current use of gsDesign implements a "hard stop" at 58 months and it does not appear the final target of 689 events is always reached (from the summary below)? But I'm not sure about that. The interim analysis trigger in my code strictly requires the target number of events to be met (and sometimes goes beyond depending on discreteness in terms of event accumulations).

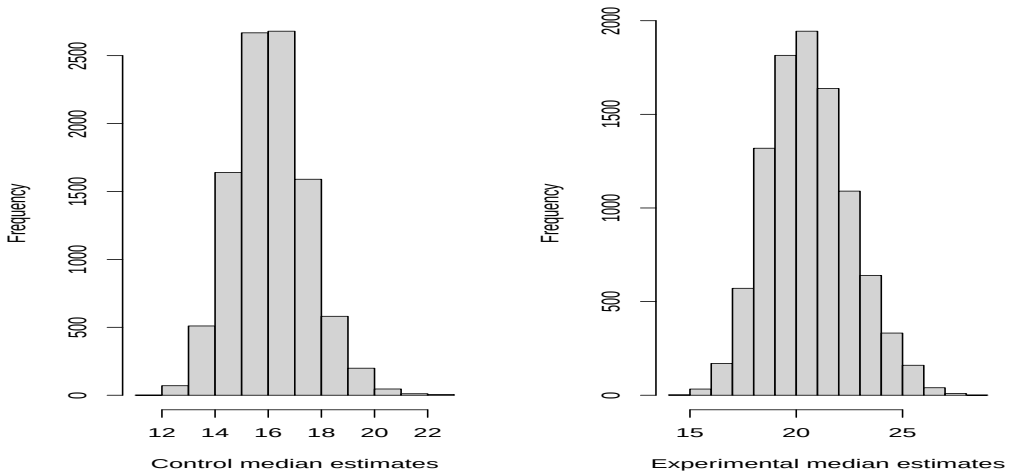


Figure 5: Histograms of Kaplan-Meier medians at look=1

.

Table 4: Summary details for interim looks. Statistics are un-conditional with respect to stopping: τ_{IA} denotes the interim timming; *Zlr.IA* the interim z-crossing boundary; exp(*bhat.IA*) denotes the interim hazard ratio; m0.IA (m1.IA) denotes the median for control (experimental); and drop0.IA (drop1.IA) denotes the dropout proportion for control (experimental).

	N	1	2	3
		$N = 10000$	$N = 10000$	$N = 10000$
τ.IA	30000	26 26 27	38 39 39	56 57 59
Zlr.IA	30000	1.6 2.3 3.0	2.2 2.9 3.5	2.6 3.3 3.9
exp(bhat.IA)	30000	0.72 0.78 0.84	0.74 0.78 0.83	0.74 0.78 0.82
m0.IA	30000	15 16 17	15 16 17	15 16 17
m1.IA	29764	19 20 22	20 21 22	20 21 21
drop0.IA	30000	0.088 0.097 0.108	0.088 0.097 0.108	0.088 0.097 0.108
drop1.IA	30000	0.12 0.13 0.14	0.12 0.13 0.14	0.12 0.13 0.14

$_a\, b\, _c$ represent the lower quartile a, the median b, and the upper quartile c for continuous variables. N is the number of non–missing values.

Whatever the differences, they seem to be mild in terms of how close the key statistics match.

```
## Time to event group sequential design with HR= 0.78
## Equal randomization:          ratio=1
## One-sided group sequential design with
## 88 % power and 2 % Type I Error.
##
##   Analysis  N   Z   Nominal p  Spend
##          1 345 2.99    0.0014 0.0014
##          2 537 2.40    0.0083 0.0074
##          3 689 2.12    0.0172 0.0112
##      Total                    0.0200
##
## ++ alpha spending:
##  O'Brien-Fleming boundary.
##
## Boundary crossing probabilities and expected sample size
## assume any cross stops the trial
##
## Upper boundary (power or Type I Error)
##           Analysis
##   Theta      1      2      3 Total E{N}
##   0.000 0.0014 0.0074 0.0112  0.02  686
##   0.124 0.2462 0.4430 0.1907  0.88  536
##          T   n Events HR efficacy
## IA 1  27.0 905    344       0.724
## IA 2  39.4 905    537       0.813
## Final 58.0 905    688       0.851
## Accrual rates:
##      Stratum 1
## 0-26      34.8
## Control event rates (H1):
##       Stratum 1
## 0-Inf      0.04
## Censoring rates:
##       Stratum 1
## 0-Inf         0
```