

# Template instructions for time-to-event group sequential design

Keaven M. Anderson

February 1, 2016

## 1 Introduction

This vignette provides a instructions for a template for time-to-event sample size calculations for fixed designs using `nSurv` and group sequential designs using `gsSurv`. It may be useful to edit the template (found in the source for the `gsDesign` R package in `/inst/doc/gSurvTemplate.rnw`) so that you can reuse it on a regular basis for sample size calculations for time-to-event study planning. The template only uses the simplest options with a single stratum and exponential failure and dropout rates. The template can be modified to accomodate multiple strata and/or piecewise exponential failure and dropout rates; this was not chosen here since the simplest options are a) often used and b) simplest to learn and apply for beginners. Note that we produce tabular, textual and graphical output; examining the source file to see how this is done will enable you to easily customize to fit your purposes. You will need the `knitr` R package, which I find simpler to use than `Sweave` (although you could make minor edits and use `Sweave`). I have found using `knitr` and the `RStudio` development environment to be a good combination. Within this instruction vignette for the template, we generally show all of the code used to generate results. You will want to suppress this in your own template as demonstrated in `/inst/doc/gSurvTemplate.rnw`.

We apply the Lachin and Foulkes (1986) sample size method and extend it to group sequential design. This method fixes the duration of a study and varies enrollment rates to power a trial. We also use the Lachin and Foulkes (1986) basic power equation to compute sample size along the lines of Kim and Tsiatis (1990) where enrollment rates are fixed and enrollment duration is allowed to vary to enroll a sufficient sample size to power a study. While `nSurv` and `gsSurv` allow stratified populations and piecewise exponential failure rates, we restrict ourselves to a single stratum with an exponential failure distribution here; see the help file for `gsSurv` for examples with a stratified population or piecewise exponential failure.

Some detail in specification comes with the flexibility allowed by the Lachin and Foulkes (1986) method. Thus a template is helpful to simplify use.

## 2 Introduction

We present a relatively simple example of time-to-event sample size calculation for a group sequential design.

### 2.1 Basic assumptions and fixed design sample size

Following is code using the `gsDesign` R package. The first chunk of code sets up enrollment and dropout information.

```
require(gsDesign)
# enrollment period durations; the last of these will be extended if T=NULL below
R <- c(1,2,3,4)
# relative enrollment rates during above periods
gamma<-c(1,1.5,2.5,4)
# study duration
# T can be set to NULL if you want to
# fix enrollment and vary study duration
T <- 36
# follow-up duration of last patient enrolled
minfup <- 12
# randomization ratio, experimental/control
ratio <- 1
```

Next we provide information about the median time to event in the control group, dropout rate, hazard ratios under the null and alternate hypotheses for experimental therapy compared to control, and the desired Type I and II error rates.

```
# median control time-to-event
median <- 12
# exponential dropout rate per unit of time
eta <- .001
# hypothesized experimental/control hazard ratio
hr <- .75
# null hazard ratio (1 for superiority, >1 for non-inferiority)
hr0 <- 1
# Type I error (1-sided)
alpha <- .025
# Type II error (1-power)
beta<- .1
```

Finally, we design a trial with no interim analyses under these assumptions. Note that when calling `nSurv`, we transform the median time-to-event ( $m$ ) to an exponential event rate ( $\lambda$ ) with the formula

$$\lambda = \log(2)/m.$$

```
x <- nSurv(R=R,gamma=gamma,eta=eta,minfup=minfup,
           T=T,lambdaC=log(2)/median,
           hr=hr,hr0=hr0,beta=beta,alpha=alpha)
```

Next we store some text for use in output; only need to edit `timename`, `timename1` and `ep` as the following code is used to format enrollment rate input from above. See the following paragraph below to see where this text is used. Note that this code is not needed for Table 2 produced using `xprint`.

```
# 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)
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="")
}
```

The median time-to-event is assumed to be 12 months in the control group. The trial is designed to demonstrate superiority of experimental treatment over control with an assumed hazard ratio of 0.75 with 90% power and a one-sided Type I error rate of 2.5. The total sample size is 776 and a total of 508 endpoints is required for analysis. Planned recruitment duration is 24 months and the minimum follow-up planned is 12 months. Thus, the total expected study duration is 36 months. Enrollment is assumed to be piecewise constant at rates of 9.3 for months 0 - 1, 13.9 for months 1 - 3, 23.2 for months 3 - 6, 37.1 for months 6 - 24. The assumed dropout rate is 0.1% per month.

### 3 Group sequential design

Now we move on to a group sequential design. We set up the number of analyses, timing and spending function parameters.

```
# number of analyses (interim + final)
k <- 2
# timing of interim analyses (k-1 increasing numbers >0 and <1)
timing <- c(.4)
# efficacy bound spending function
```

```

sfu <- sfHSD
# efficacy bound spending parameter specification
sfupar <- -10
# futility bound spending function
sfl <- sfHSD
# futility bound spending parameter specification
sflpar <- 2

```

Type II error (1-power) may be set up differently than for a fixed design so that more meaningful futility analyses can be performed during the course of the trial.

```

# Type II error=1-Power
beta <- .2

```

Now we are prepared to generate the design.

```

# generate design
x <- gsSurv(k=k,timing=timing,R=R,gamma=gamma,eta=eta,
           minfup=minfup,T=T,lambdaC=log(2)/median,
           hr=hr,hr0=hr0,beta=beta,alpha=alpha,
           sfu=sfu,sfupar=sfupar,sfl=sfl,sflpar=sflpar)
# make a string with enrollment rates
# (assumes gamma is a single value or vector)
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="")
}

```

For a comparative trial we consider a 2-arm group sequential design with overall survival as the primary endpoint as shown in Table 1; an alternate of the table is demonstrated in 2. 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. Bounds are determined by Hwang-Shih-DeCani spending functions with  $\gamma = -10$  ( $\alpha$ -spending) and  $\gamma = 2$  ( $\beta$ -spending). The median time-to-event is assumed to be 12 months in the control group. The trial is designed to demonstrate superiority of experimental treatment over control with an assumed hazard ratio of 0.75. The total sample size is 701 and a total of 459 endpoints is required for the final analysis. Planned recruitment duration is 24 months and the minimum follow-up planned is 12 months. Thus, the total

expected study duration is 36 months. Enrollment is assumed to be piecewise constant at rates of 8.4 for months 0 - 1, 12.6 for months 1 - 3, 21 for months 3 - 6, 33.6 for months 6 - 24. The assumed dropout rate is 0.1% per month. There is a single interim analysis planned after 184 events have accrued which is expected after approximately 19.9 months.

```
print(xtable(x, footnote=footnote, caption=caption,
            label="tab1"),
      include.rownames=F,
      sanitize.text.function=function(x) x)
```

Analysis	Value	Futility	Efficacy
IA 1: 40%	Z-value	0.81	3.84
N: 564	HR	0.89	0.57
Events: 184	p (1-sided)	0.2084	1e-04
19.9 months	P{Cross} if HR=1	0.7916	1e-04
	P{Cross} if HR=0.75	0.1274	0.0293
Final analysis	Z-value	1.96	1.96
N: 702	HR	0.83	0.83
Events: 459	p (1-sided)	0.025	0.025
36 months	P{Cross} if HR=1	0.9801	0.0199
	P{Cross} if HR=0.75	0.2	0.8

P{Cross} is the probability of crossing the given bound (efficacy or futility) at or before the given analysis under the assumed hazard ratio (HR). Design assumes futility bound is discretionary (non-binding), but smaller upper boundary crossing probabilities shown here assume trial stops at first boundary crossing (binding bounds).

Table 1: Overall survival trial design with HR=0.75, 80% power and 2.5% Type 1 error.

```
xprint(xtable(gsBoundSummary(x, logdelta=TRUE),
              digits=4, label="tab2",
              caption=summary(x)))
```

Following are plots of the Z-values (Figure 1) and approximate hazard ratios (Figure 2) at the design bounds.

```
plot(x, cex=.8, xlab="Number of events")
```

```
plot(x, plottype="hr", cex=.8, xlab="Number of events")
```

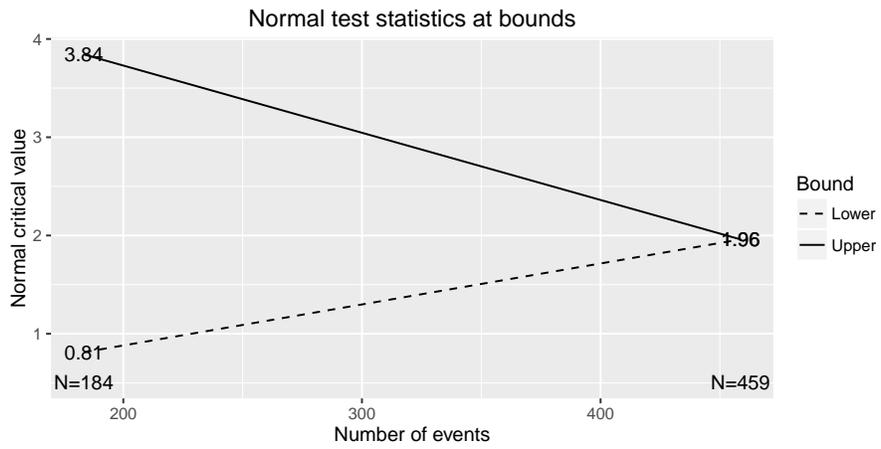


Figure 1: Z-value bound plot

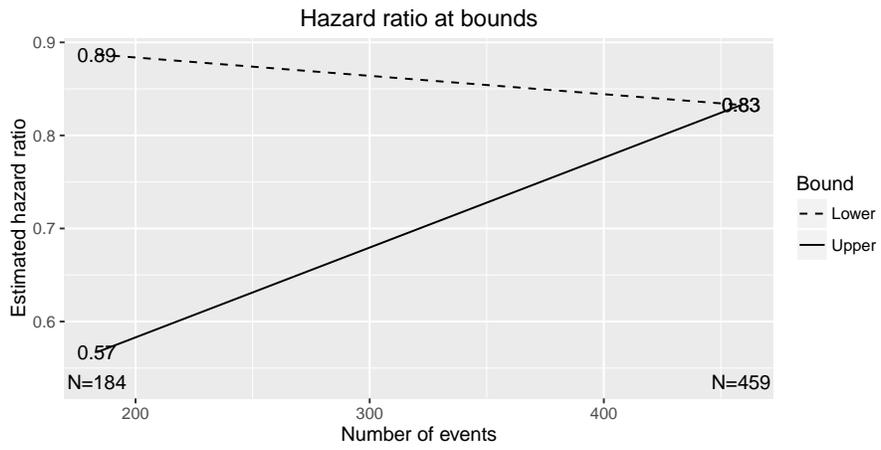


Figure 2: Hazard ratio bound plot

Analysis	Value	Efficacy	Futility
IA 1: 40%	Z	3.8427	0.8119
N: 564	p (1-sided)	0.0001	0.2084
Events: 184	HR at bound	0.5670	0.8870
Month: 20	P(Cross) if HR=1	0.0001	0.7916
	P(Cross) if HR=0.75	0.0293	0.1274
Final	Z	1.9602	1.9602
N: 702	p (1-sided)	0.0250	0.0250
Events: 459	HR at bound	0.8327	0.8327
Month: 36	P(Cross) if HR=1	0.0199	0.9801
	P(Cross) if HR=0.75	0.8000	0.2000

Table 2: Asymmetric two-sided group sequential design with non-binding futility bound, 2 analyses, time-to-event outcome with sample size 702 and 459 events required, 80 percent power, 2.5 percent (1-sided) Type I error to detect a hazard ratio of 0.75. Enrollment and total study durations are assumed to be 24 and 36 months, respectively. Efficacy bounds derived using a Hwang-Shih-DeCani spending function with  $\gamma = -10$ . Futility bounds derived using a Hwang-Shih-DeCani spending function with  $\gamma = 2$ .

## References

- Kyungmann Kim and Anastasios A. Tsiatis. Study duration for clinical trials with survival response and early stopping rule. *Biometrics*, 46:81–92, 1990.
- John M. Lachin and Mary A. Foulkes. Evaluation of sample size and power for analyses of survival with allowance for nonuniform patient entry, losses to follow-up, noncompliance, and stratification. *Biometrics*, 42:507–519, 1986.