# The `OptimPhase2` Package

Bo Huang and Neal Thomas
Pfizer Inc.

December 3, 2008

## Contents

# 1 Introduction

Phase II oncology trials are undertaken to assess the activity of a new treatment with activity most frequently defined in terms of tumor response. With some of the new targeted therapies, it may not be appropriate to use tumor shrinkage to evaluate activity. Instead, the primary endpoint is time to death, time to progression, or some other endpoints that must be evaluated over a longer time period.

Phase II trials are often designed with an interim analysis so they can be stopped early if a drug is ineffective. However, when the primary endpoint requires a longer observation period, interim analyses are challenging because of incomplete follow-up for some patients at the time of the interim analysis. Standard designs such as Simon (1989) require suspension of accrual while patient follow-up is completed. Case and Morgan (2003) presented a two-stage design for a phase II oncology trial with a long-term endpoint that does not suspend accrual while the interim analysis is conducted. They proposed to use the Kaplan-Meier or Nelson-Aalen estimators of the event probability, using methods like those in Lin et al (1996). Estimation at the time of the interim analysis includes patients with partial follow-up without necessitating trial suspension, as also proposed by Jennison and Turnbull (2000). The design minimizes either the expected sample size, expected duration of accrual, or the expected total study length under the hypothesis that the drug is ineffective. The null hypothesis for the new design is an (assumed) known event-free rate within a specified time, which has been judged to represent ineffective treatment. This is similar to the hypothesis in the Simon design, but with much longer specified times for events to occur. Schaid, Wieand, and Therneau (1990) proposed a similar design using the log rank statistic, which also incorporates patients with incomplete follow-up; the log rank statistic is not evaluated here, but could be inlcuded as a future software

option.

Both Lin et al (1996) and Case and Morgan (2003) assume a constant accrual rate throughout the trial, which is not typical in practice. We further investigate the design properties by generalizing the accrual distribution to have different accrual rates in user-specified intervals. As noted in Case and Morgan (2003), when only partial follow-up data are available, the level of the testing procedure can depend on the assumed accrual distribution and the assumed time-to-event distribution under the null hypothesis. We evaluate an optimal design and corresponding analysis that ensure the Type I error rate is below the target level. The reduction in power or corresponding increase in sample size necessary to achieve the conservative type I error rates is also evaluated.

The theoretical derivations of the optimal designs specify a fixed time to end the first stage of accrual and to conduct the interim analysis, with corresponding projected sample size. Case and Morgan (2003) also evaluated a modified interim timing rule that ends the first stage when the projected number of patients has been accrued regardless of the planned interim time. They showed this rule has more robust statistical properties when the accrual rate is mis-specified. We also evaluate this interim timing rule and an additional rule based on the projected patient exposure at the optimal interim time. This rule not only accounts for the number of patients actually accrued, but also the length of time patients have been observed. All of the interim timing rules can be easily applied in practice.

This R package *OptimPhase2* was created to generate the optimal designs and resulting analyses. The package includes code to perform simulations to validate the theoretical calculations, some of which depend on asymptotic approximations. The package also has several options for evaluating a proposed design under conditions that differ from those assumed when the design was created. All these will be elaborated with examples in the subsequent

sections.

# 2 Generate the Optimal Designs

## 2.1 Optimal Design Functions

```
OptimDes(
B.init, m.init, alpha, beta, param, x, target = c("EDA", "ETSL","ES"),
recover=TRUE, control = OptimDesControl(),...)

np.OptimDes(
B.init, m.init, alpha, beta, param, x, n = NULL, pn = NULL,
pt = NULL, target = c("EDA", "ETSL","ES"), recover=TRUE,
control = OptimDesControl(), CMadj=F, ...)
```

*OptimDes* finds a two-stage design for a time to event endpoint with potential stopping for futility after the first stage. It implements the Case and Morgan (2003) generalization of the Simon (1989) two-stage design for comparing a treatment to a known standard. The design minimizes either the expected duration of accrual (EDA), expected sample size (ES), or the expected total study length (ETSL).

The design calculations assume Weibull distributions for the event-free endpoint in the treated group, and for the (assumed known, "Null") control distribution. The function *weibPmatch* (see Section 5) can be used to select Weibull parameters that yield a target event-free rate at a specified time. Estimation is based on the Kaplan-Meier or Nelson-Aalen estimators evaluated at a target time (e.g., 1 year). The treatment and control distributions and the accrual distribution affect power (and alpha level in some settings), see Case and Morgan (2003)).

Accrual rates are specified by the user. These rates can differ across time intervals specified by the user (this generalizes the results in Case and Morgan). The accrual information is controlled by arguments *B.init* and *m.init*.

*OptimDes* also has the capability to choose to recover or not recover the amount of $\alpha$ potentially saved at the interim analysis (analogous to non-binding in group sequential designs).

**Note:** Details of *OptimPhase2* and all subsequent package functions can be found on the help pages.

## 2.2 Examples

Assume the 1-year survival rate of a standard cancer therapy is 0.40 ($H_0$). An improvement to 0.60 would be considered clinically significant ($H_1$). Assume the survival distributions have different shapes and scales under null and the alternative, determined by the weibull parameters $(1, 1.09)$ under $H_0$ and $(2, 1.40)$ under $H_1$. Type I error is 0.05. Type II error is 0.1. It is also assumed that the numbers of patients that can be enrolled in the first 5 years are 15, 20, 25, 20 and 15 respectively.

```
> B.init <- c(1, 2, 3, 4, 5)
> m.init <- c(15, 20, 25, 20, 15)
> alpha <- 0.05
> beta <- 0.1
> param <- c(1, 1.09, 2, 1.4)
> x <- 1
```

The optimal design *object1* minimizing the expected duration of accrual (EDA) after implementing *OptimDes* can then be obtained.

```
> object1 <- OptimDes(B.init, m.init, alpha, beta, param, x, target = "EDA")
> print(object1)
```

```
         Optimal Design Results

H0: S0= 0.4 H1: S1= 0.6

Type I error(1-sided upper): 0.05 type II error: 0.1
Recover alpha: TRUE
Event-free time of interest: 1

              target: EDA
      EDA          ETSL            ES
    2.777         3.243        53.253

      Sample Size at Each Stage
          n1           n.last
          39               70

      Study time at Each Stage
           t1             MTSL
        2.145            4.500

Projected patient exposure at interim analysis:  28.22

Proportion of the total information at the interim analysis:
       Under NuLL     Under Alternative
            0.361                 0.330


      Hypothesis Test Boundaries
           C1               C2
```

```
        0.085             1.601


Approximate Rates Corresponding to Test Boundaries*

      Event-free rate for C1:  0.412
      Event-free rate for C2:  0.5


Single-stage Design (Asymptotic normal calculation)
  Single stage N                  DA                  SL
           64.0                 3.2                 4.2


*Note:  Rates corresponding to test boundaries are a function
of the non-parametric SE computed at the time of the analyses.
The approximate rates are based on the asymptotic SE computed
under the null and alternative hypotheses.
```

A plot function *plot.OptimDes* is used to display the ETSL, ES and EDA for a two-stage design relative to a single-stage design as a function of the combined stage 1 and 2 sample size. It demonstrates the tradeoff between ETSL, EDA and ES as a function of the combined sample size. Robustness of the optimal two-stage design to deviations from the target sample size can be explored. The plot often suggests a compromised design achieving near-optimal results for both EDA and ETSL be a favorable design to the optimal one based on a single criteria. Test boundary values $(C_1, C_2)$, and numerical values of other design parameters, can be obtained for a design selected from the plot using function *np.OptimDes*. Thus, *np.OptimDes* generates the optimal design when the total sample size is fixed.

Using the above case as an example, let's assume the ETSL is minimized this time. Then *object2* is the corresponding optimal design with the optimal plot Figure 1. The optimal design is displayed as the sold circle on the plot. If investigators believe a compromised design with maximum study length ratio = 1.1 ($pt = 1.1$ in *np.OptimDes*) will save some patients while still producing near-optimal results, $pt = 1.1$ can be input into *np.OptimDes* and the adjusted "optimal" design can be created

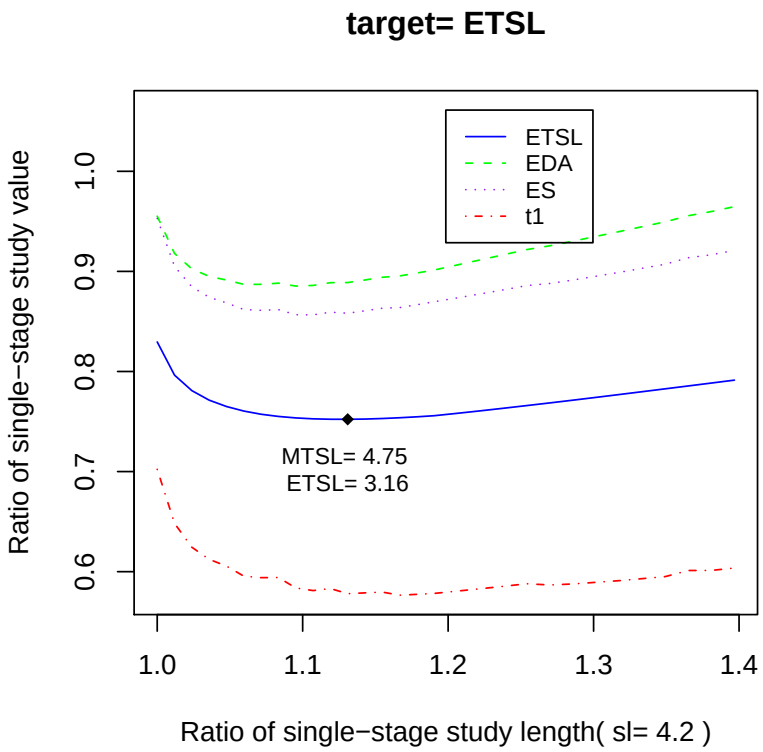


Figure 1: The optimality criteria displayed for a range of maximum sample sizes. The criteria and the maximum sample sizes are expressed as ratios relative to the corresponding value in a single-stage fixed design.

```
> object3 <- np.OptimDes(B.init, m.init, alpha, beta, param, x,
+     pt = 1.1, target = "ETSL")
> print(object3)
```

```
             Optimal Design Results

    H0: S0= 0.4 H1: S1= 0.6

    Type I error(1-sided upper): 0.05 type II error: 0.1
    Recover alpha: TRUE
    Event-free time of interest: 1

                target: ETSL
       EDA           ETSL             ES
     2.835          3.161         54.819

       Sample Size at Each Stage
           n1           n.last
           47               73

       Study time at Each Stage
            t1             MTSL
         2.441            4.650

    Projected patient exposure at interim analysis:  35.03

    Proportion of the total information at the interim analysis:
          Under NuLL     Under Alternative
               0.432                 0.400
```

```
      Hypothesis Test Boundaries
           C1                C2
        0.450             1.572


   Approximate Rates Corresponding to Test Boundaries*

        Event-free rate for C1:  0.454
        Event-free rate for C2:  0.496


   Single-stage Design (Asymptotic normal calculation)
     Single stage N                  DA                  SL
               64.0                 3.2                 4.2


 *Note:  Rates corresponding to test boundaries are a function
 of the non-parametric SE computed at the time of the analyses.
 The approximate rates are based on the asymptotic SE computed
 under the null and alternative hypotheses.
```

# 3 Test Statistics and Decision Rules at Each Stage

```
Test2stage(
Y1, T1, Y2 = NULL, T2 = NULL, p0, x, C1, C2, t1, MTSL =
NULL, printTest = TRUE, cen1=rep(1,length(T1)),cen2=rep(1,length(T2)))
```

The test statistic at the end of each stage is computed and compared to the decision boundaries $C_1$ and $C_2$.

- Stage 1: Accrue $n_1$ patients between time 0 and time $t_1$. Each patient is followed until they have an event or successfully reach time $x$, or until study time $t_1$, whichever is first. Calculate the normalized Z-statistic by *Test2stage*, and denote it by $Z_1(x; t_1)$. If $Z_1(x; t_1) < C_1$, stop the study for futility; otherwise, continue to the next stage. The probability of stopping under the null hypothesis is approximated by $P_s = \Phi(C_1)$, where $\Phi$ is the standard normal cummulative distribution function. $n_1$ is a random variable determined by $t_1$ and the accrual distribution.

- Stage 2: Accrue $n_2$ additional patients between times $t_1$ and maximum duration of accrual ($MDA$). Follow all patients (both stages) until they have an event or successfully reach time $x$, then calculate a second $Z$ statistic at the end of maximum total study length ($MTSL$), denoted by $Z_2(x; MTSL)$, and reject $H_0$ if $Z_2(x; MTSL) > C_2$.

For example, if at the end of Stage 1, the test statistic $Z_1 = 3.391$, $C_1 = 0.085$. Then *Test2stage* will return

```
Z1 >= C1, continue to the second stage
```

# 4 Simulation Studies

```
SimDes(
object,B.init,m.init,weib0,weib1,interimRule='e1',
sim.n=1000,e1conv=1/365,CMadj=F,attainI=1,attainT=1,
FixDes="F", Rseed)
```

The *SimDes* function is a powerful function to simulate experiments to compare the true alpha level and power of two-stage designs from function

*OptimDes* with the targeted nominal values. It can also be used to assess the performance of the optimal design under mis-specification of the design parameters. For example, if the Weibull shape and scale parameters of the time to event distributions are changed, if the accrual rates deviate from the projected ones, or if the interim analysis is conducted differently from the planned one under the more realistic conditions. In addition, the function has the option to determine the timing of the interim analysis by matching the observed information to the expected time, number of patients or patient exposure (interimRule="t1", "n1" or "e1").

## 4.1 Example 1: Optimal Settings

Recall that in Section 2.2 *object1* is the optimal design minimizing the EDA. Under the expected parameter settings, 1000 simulations are conducted by matching the expected patient exposure at the interim

```
> (sim1 <- SimDes(object1))

 The interimRule is  e1

alphaExact  alphaNorm powerExact  powerNorm        eda       etsl         es
 0.0380000  0.0630000  0.9290000  0.9470000  2.8231925  3.3675778 55.6160000
 pstopNull   pstopAlt       aveE  pinfoNull   pinfoAlt         n1         t1
 0.4730000  0.0230000 28.2249407  0.3433120  0.3295983 39.6350000  2.1605113
    phatKl     phatKh     phatRl     phatRh
 0.4081223  0.4081139  0.5035585  0.4893747
```

Details of the returned values can be found from the help pages. For instance, the estimated alpha level using an exact test for the second stage test is 0.038.

## 4.2 Example 2: Differed Accrual Rates

Now suppose the actual numbers of patients that can be accrued in the first 5 years are different from the originally planned for the optimal design (*m.init* below), then the results after 1000 simulated trials become

```
> (sim2 <- SimDes(object1, m.init = c(5, 5, 25, 25, 25)))

 The interimRule is  e1

alphaExact  alphaNorm powerExact  powerNorm        eda       etsl         es
 0.0320000  0.0490000  0.9200000  0.9400000  3.7683860  4.2782850 55.3520000
 pstopNull   pstopAlt       aveE  pinfoNull   pinfoAlt         n1         t1
 0.5090000  0.0360000 28.2250430  0.3220282  0.2925826 41.2330000  3.2325549
    phatKl     phatKh     phatRl     phatRh
 0.4078247  0.4081419  0.5035585  0.4893747
```

## 4.3 Example 3: Differed Interim Timing

If the actual interim time or sample size (depending on *interimRule*,*interimRule*=*"t1"* below) is different from the originally planned for the optimal design (*attainI*=0.8 below), then the results after 1000 simulated trials become

```
> (sim3 <- SimDes(object1, interimRule = "t1", attainI = 0.8))

 The interimRule is  t1

alphaExact  alphaNorm powerExact  powerNorm        eda       etsl         es
 0.0320000  0.0440000  0.9040000  0.9150000  2.5621788  3.0886060 49.6330000
 pstopNull   pstopAlt       aveE  pinfoNull   pinfoAlt         n1         t1
 0.4980000  0.0530000 19.2944758  0.2064760  0.1932355 29.0550000  1.7162918
    phatKl     phatKh     phatRl     phatRh
 0.4091581  0.4143881  0.5035585  0.4893747
```

)

# 5 Survival Curves Based on the Weibull Distribution

```
weibPmatch(x, p0, shape, scale)
```

```
weibull.plot(param, x, l.type = 1:3, l.col = c("blue", "red"), ...)
```

*weibPmatch* and *weibull.plot* are used together to determine the shape and scale parameters of the Weibull distrbution for the survival curves under $H_0$ and $H_1$. The Weibull distribution is flexible enough to cover the majority of scenarios likely to encounter in practice.

*weibPmatch* determines the shape or scale parameter of a Weibull distribution so it has event-free rate $P_0$ at time $x$. If the shape is specified, the scale parameter is computed, and if the scale is specified, the shape parameter is computed.

*weibull.plot* then plots Weibull survival curves with differences at a target time highlighted from the parameters computed from *weibPmatch*. Figure 2 is an example plot implementing the Weibull parameters input to *OptimDes* to create *object1*.

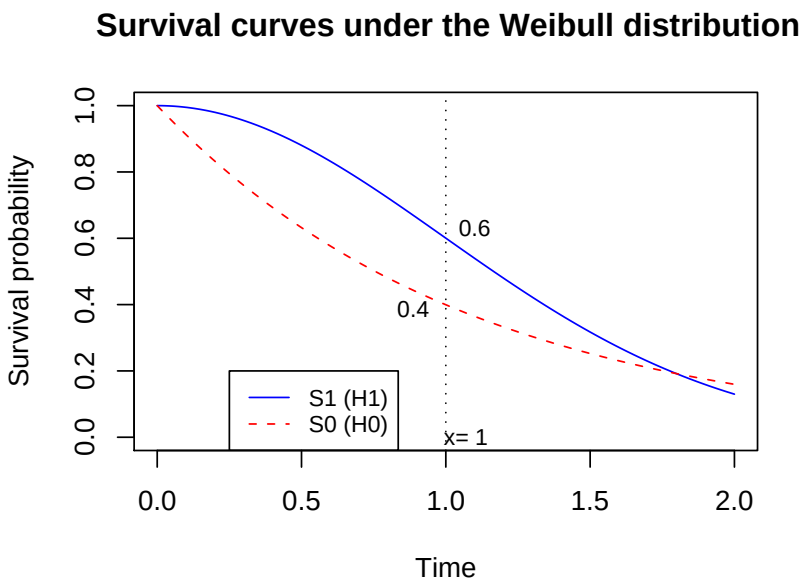


Figure 2: Survival curves under the Weibull distribution

# References

[1] Agresti, A. (2002). Categorical Data Analysis. *Weiley*

[2] Brent, R. (1973). Algorithms for minimization without derivatives. *Englewood Cliffs, Prentice Hall*

[3] Case L.D. and Morgan, T.M. (2003). Design of phase II cancer trials evaluating survival probabilities. *BMC Medical Research Methodology* **3**: 6

[4] Food and Drug Administration (2006). Guidance for Clinical Trial Sponsors: Establishment and operation of clinical trial data monitoring committees. *OMB Control No. 0910-0581*

[5] Jonker DJ et al. (2007). Abstract No. LB-1, *American Association for Cancer Research*

[6] Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of American Statistical Association.* **53**: 457-481

[7] Lin, D.Y., Shen, L., Ying, Z. and Breslow, N.E. (1996). Group sequential designs for monitoring survival probabilities. *Biometrics.* **52**: 1033-1042

[8] Machin, D., Campbell, M., Tan, S.B. and Tan, S.H. (2008). Sample Size Tables for Clinical Studies. *Blackwell*

[9] Nelson, W. (1969). Hazard plotting for incomplete failure data. *J Quality Technology.* **1**: 27-52

[10] Rao S, Cunningham D, de Gramont A, et al (2004). Phase III double-blind placebo-controlled study of farnesyl transferase inhibitor R115777 in patients with refractory advanced colorectal cancer. **22(19)**: 3950-3957

[11] Schaid, D., Wieand, S., and Therneau, T. (1990). Optimal two-stage screening designs for survival comparisons. *Biometrika*. **77**: 507-513

[12] Simon, R. (1989). Optimal two-stage designs for phase II clinical trials. *Controlled Clinical Trials*. **10**: 1-10

[13] Van Cutsem E, Peeters M, et.al. (2007). Open-label phase III trial of panitumumab plus best supportive care compared with best supportive care alone in patients with chemotherapy-refractory metastatic colorectal cancer. *Journal of Clinical Oncology*. **1;25(13)**:1658-1664.