# Power and Sample Size for Dose-Finding Studies With Survival Endpoints Under Model Uncertainty

**Qiqi Deng**[1], **Xiaofei Bai**[1], **Dacheng Liu**[1], **Dooti Roy**[1], **Zhiliang Ying**[2], and **Dan-Yu Lin**[3,*]

[1]Biostatistics and Data Science, Boehringer Ingelheim Pharmaceuticals Inc., 900 Ridgebury Rd, Ridgefield, CT 06811, U.S.A.

[2]Department of Statistics, Columbia University, New York, NY 10027, U.S.A.

[3]Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27599, U.S.A.

## SUMMARY.

Multiple comparison procedures combined with modeling techniques (MCP-Mod) (Bretz et al., 2005) is an efficient and robust statistical methodology for the model-based design and analysis of dose-finding studies with an unknown dose-response model. With this approach, multiple comparison methods are used to identify statistically significant contrasts corresponding to a set of candidate dose-response models, and the best model is then used to estimate the target dose. Power and sample size calculations for this methodology require knowledge of the covariance matrix for the estimators of the (placebo-adjusted) mean responses among the dose groups. In this paper, we consider survival endpoints and derive an analytic form of the covariance matrix for the estimators of the log hazard ratios as a function of the total number of events in the study. We then use this closed-form expression of the covariance matrix to derive the power and sample size formulas. We discuss practical considerations in the application of these formulas. In addition, we provide an illustration with a motivating example on chronic obstructive pulmonary disease. Finally, we demonstrate through simulation studies that the proposed formulas are accurate enough for practical use.

## 1. Introduction

There are two major objectives in early-phase clinical trials: the first one is to demonstrate the evidence of clinical efficacy (with acceptable risk) for the study medication, and the second one is to select the dose(s) to be tested in confirmatory trials. These two objectives are often referred to as proof-of-concept (PoC) and dose-finding steps, respectively. The two steps may be combined in a single trial in order to accelerate the development process.

*lin@bios.unc.edu.

Statistical methods for dose-finding studies can be roughly categorized into multiple comparison procedures (MCP) and modeling techniques. The former approach regards dose as a qualitative factor and evaluates the statistical significance of contrasts between doses while controlling the family-wise error rate (FWER) at a pre-specified level; the statistical inference is restricted to the small set of doses used in the trial. The latter approach assumes a functional relationship between dose and response and uses the fitted dose-response model to estimate a dose required to achieve a desired level of response; the validity of the inference depends strongly on correct model specification.

Bretz et al. (2005) proposed a unified strategy, termed multiple comparison procedures combined with modeling techniques (MCP-Mod), that combines the MCP principle with the modeling approach. MCP-Mod specifies a set of candidate models covering a wide range of dose-response curves. Each model in the candidate set is tested with appropriate contrasts using MCP to preserve the FWER. The best model among the significant ones is chosen to estimate the target dose using a modeling technique. MCP-Mod was recently recommended by the European Medicines Agency (2014) as an efficient statistical methodology for the model-based design and analysis of dose-finding studies under model uncertainty and also received the "fit-for-purpose" designation by the US Food and Drug Administration in 2016.

The MCP-Mod methodology was originally developed for normally distributed, homoscedastic response at a single time point under the parallel-group study design (Bretz et al., 2005; Pinheiro et al., 2006). Recently, Pinheiro et al. (2014) extended this methodology to general parametric models by using generalized least squares estimation. The extension is focused on the analysis stage, where parameter estimates and covariance matrix are obtained from standard statistical software packages. The extension does not deal with calculations of power and sample size, which require knowledge of the covariance matrix before data collection.

In this paper, we show how to perform power calculations and determine sample sizes for the MCP-Mod methodology with survival endpoints. In particular, we derive the covariance matrix for the estimators of the log hazard ratios under the proportional hazards model. We express the covariance matrix in terms of the total number of events (i.e., uncensored observations) in the entire study. This analytic expression of the covariance matrix greatly facilitates power and sample size calculations for MCP-Mod with survival endpoints.

The rest of this paper is organized as follows. In Section 2, we present our methods and address related challenges. In Section 3, we provide a motivating example, together with simulation results. In Section 4, we discuss the use of the analytic form of the covariance matrix in other contexts.

## 2.   Methods

### 2.1.   Parallel Group Survival Studies

Suppose that there are $K$ doses $x_1$, …, $x_K$ besides placebo. Note that Bretz et al. (2005) used $x_1$ to denote placebo. For survival endpoints, we take the placebo-adjusted dose-response modeling, so it is more convenient to use $x_1$ to denote the first dose group. We specify that

the hazard function of survival time $T$ in the $k$th dose group takes the form of the Cox (1972) proportional hazards model

$$\lambda_k(t) = \lambda_0(t)e^{\beta_k}, \quad k = 1, ..., K,$$

where $\lambda_0(t)$ is an arbitrary hazard function for placebo, and $\beta_k$ is the log hazard ratio.

For $k = 0, 1, ..., K$, let $n_k$ denote the number of patients in the $k$th dose group, with $k = 0$ pertaining to the placebo group. Write $n = n_0 + n_1 + ..., n_K$. For $i = 1, ..., n$, let $T_i$ and $C_i$ denote, respectively, the survival time and censoring time for the $i$th patient. For $i = 1, ..., n$ and $k = 1, ..., K$, let $X_{ki}$ indicate, by the values 1 versus 0, whether or not the $i$th patient receives dose $x_k$. The data consist of $(\tilde{T}_i, \Delta_i, X_i)$ ($i = 1, ..., n$), where $\tilde{T}_i = \min(T_i, C_i)$, $\Delta_i = I(T_i \leq C_i)$, $I(\cdot)$ is the indicator function, and $X_i = (X_{1i}, ..., X_{Ki})^T$.

## 2.2. Parameter Estimation and Covariance Matrix

For parameter estimation, it is more convenient to express the $K$ proportional hazards models as a single model: $\lambda\left(t|X_i\right) = \lambda_0(t)e^{\beta^T X_i}$, where $\boldsymbol{\beta} = (\beta_1, ..., \beta_K)^T$. The partial likelihood (Cox, 1975) score function for $\boldsymbol{\beta}$ is given by

$$U(\boldsymbol{\beta}) = \sum_{i=1}^{n} \Delta_i\{X_i - E(\boldsymbol{\beta}, \tilde{T}_i)\},$$

where $E(\boldsymbol{\beta}, t) = \sum_{i=1}^{n} I(\tilde{T}_i \geq t)e^{\beta^T X_i}X_i / \sum_{i=1}^{n} I(\tilde{T}_i \geq t)e^{\beta^T X_i}$. The corresponding information matrix is

$$\mathscr{I}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \Delta_i\left\{\frac{\sum_{j=1}^{n} I(\tilde{T}_j \geq \tilde{T}_i)e^{\beta^T X_j}X_j X_j^T}{\sum_{j=1}^{n} I(\tilde{T}_j \geq \tilde{T}_i)e^{\beta^T X_j}} - E(\boldsymbol{\beta}, \tilde{T}_i)E(\boldsymbol{\beta}, \tilde{T}_i)^T\right\}.$$

Let $\hat{\boldsymbol{\beta}}$ denote the maximum partial likelihood estimator of $\boldsymbol{\beta}$, which is the solution to the score equation $U(\boldsymbol{\beta}) = \boldsymbol{0}$. The estimator $\hat{\boldsymbol{\beta}}$ is asymptotically $K$-variate normal with mean $\boldsymbol{\beta}$ and covariance matrix $\mathscr{I}^{-1}(\hat{\boldsymbol{\beta}})$ (Andersen and Gill, 1982).

The (observed) information matrix can be calculated only after the data are collected. In the design stage, it is necessary to express the information matrix in terms of the quantities that can be pre-specified. Let $D$ denote the total number of events (i.e., uncensored survival times) in the entire study, and let

$$p_k = \frac{\xi_k e^{\beta_k}}{1 + \xi_1 e^{\beta_1} + \cdots + \xi_K e^{\beta_K}}, \quad k = 0, 1, \ldots, K,$$

where $\xi_k = n_k/n_0$ $(k = 0, 1, \ldots, K)$, and $\beta_0 = 0$. In the appendix, we derive an analytic approximation to the information matrix

$$\mathscr{I}(\boldsymbol{\beta}) \approx D$$

$$\begin{bmatrix} p_1(1 - p_1) & -p_1 p_2 & \cdots & -p_1 p_K \\ -p_2 p_1 & p_2(1 - p_2) & \cdots & -p_2 p_K \\ \vdots & \vdots & \vdots & \vdots \\ -p_K p_1 & -p_K p_2 & \cdots & p_K(1 - p_K) \end{bmatrix}.$$

This leads to a very simple closed-form expression for the covariance matrix of $\hat{\boldsymbol{\beta}}$

$$S \approx D^{-1} \begin{bmatrix} p_0^{-1} + p_1^{-1} & p_0^{-1} & \cdots & p_0^{-1} \\ p_0^{-1} & p_0^{-1} + p_2^{-1} & \cdots & p_0^{-1} \\ \vdots & \vdots & \vdots & \vdots \\ p_0^{-1} & p_0^{-1} & \cdots & p_0^{-1} + p_K^{-1} \end{bmatrix}, \quad (1)$$

which, under the null hypothesis $H_0 : \boldsymbol{\beta} = 0$, reduces to

$$S_0 \approx \frac{n}{D} \begin{bmatrix} n_0^{-1} + n_1^{-1} & n_0^{-1} & \cdots & n_0^{-1} \\ n_0^{-1} & n_0^{-1} + n_2^{-1} & \cdots & n_0^{-1} \\ \vdots & \vdots & \vdots & \vdots \\ n_0^{-1} & n_0^{-1} & \cdots & n_0^{-1} + n_K^{-1} \end{bmatrix}. \quad (2)$$

The corresponding correlation matrix is the same as in the case of a normally distributed response variable.

### 2.3. MCPMod

The dose-response model takes the form $f(x, \boldsymbol{\theta}) = a + b f^0(x, \boldsymbol{\theta}^*)$, where $f^0$ is the standardized model function indexed by the parameter vector $\boldsymbol{\theta}^*$ (Bretz et al., 2005; Pinheiro et al., 2006; 2014; Bornkamp et al., 2009). We consider $M$ candidate models $f_m(x, \boldsymbol{\theta}_m) = a_m + b_m f_m^0(x, \boldsymbol{\theta}_m^*)$

($m = 1, \ldots, M$). For each candidate model, we form a contrast of $\widehat{\boldsymbol{\beta}}$. The vector of optimal contrast coefficients under the $m$th model is

$$\boldsymbol{c}_m = \widehat{S}^{-1}\boldsymbol{\beta}_m \quad (3)$$

(Pinheiro et al., 2014), where $\widehat{S} \equiv \mathscr{I}^{-1}(\widehat{\boldsymbol{\beta}})$ is the estimated covariance matrix for $\widehat{\boldsymbol{\beta}}$ and can be obtained from standard software packages, such as R and SAS, and $\boldsymbol{\beta}_m$ is the value of $\beta$ under the $m$th candidate model, such that the $k$th component of $\boldsymbol{\beta}_m$ is $f_m(0, \boldsymbol{\theta}_m) - f_m(x_k, \boldsymbol{\theta}_m) \propto f_m^0(0, \boldsymbol{\theta}_m^*) - f_m^0(x_k, \boldsymbol{\theta}_m^*)$. We use the differences $f_m(0, \boldsymbol{\theta}_m) - f_m(x_k, \boldsymbol{\theta}_m)$ since longer survival times corresponds to lower hazards. Indeed, the proportional hazards model $\lambda(t|X_i) = \lambda_0(t)e^{\beta^{\mathrm{T}}X_i}$ can be expressed as a linear transformation model

$$\log \Lambda_0\left(T_i\right) = -\beta^{\mathrm{T}}X_i + \epsilon_i,$$

where $\Lambda_0(t) = \int_0^t \lambda_0(s)ds$, and $\epsilon_i$ has the extreme-value distribution. Thus, the dose-response curves for immediate normal responses (Bretz et al., 2005; Pinheiro et al., 2006; 2014; Bornkamp et al., 2009) can be applied to proportional hazards models with survival endpoints, but with an appropriate change of the sign.

The test statistics take the form

$$Z_m = \frac{\boldsymbol{c}_m^{\mathrm{T}}\widehat{\boldsymbol{\beta}}}{\left(\boldsymbol{c}_m^{\mathrm{T}}\widehat{S}\boldsymbol{c}_m\right)^{1/2}}, \quad m = 1, \ldots, M. \quad (4)$$

Under $H_0 : \boldsymbol{\beta} = \boldsymbol{0}$, $Z_m$ is asymptotically standard normal. In addition, the random vector $(Z_1, \ldots, Z_M)^{\mathrm{T}}$ is asymptotically $M$-variate normal with mean $\boldsymbol{0}$ and covariance matrix

$$\left\{ \frac{\boldsymbol{c}_m^{\mathrm{T}}S_0\boldsymbol{c}_l}{\left(\boldsymbol{c}_m^{\mathrm{T}}S_0\boldsymbol{c}_m\boldsymbol{c}_l^{\mathrm{T}}S_0\boldsymbol{c}_l\right)^{1/2}}; \quad m, l = 1, \ldots, M \right\}. \quad (5)$$

This joint distribution is used to determine the multiplicity-adjusted critical value for MCP-Mod. Specifically, the FWER of $\alpha$ is achieved if the critical value $q_{1-\alpha}$ satisfies

$$Pr(\max_{l = 1, \ldots, M} Z_l \geq q_{1-\alpha} \,\Big|\, \boldsymbol{\beta} = \boldsymbol{0}) = \alpha. \quad (6)$$

We solve equation (6) for $q_{1-\alpha}$ through multivariate normal integration.

### 2.4. Power and Sample Size Calculations

Under the alternative hypothesis $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_a$, $Z_m$ is asymptotically normal with mean $c_m^{\mathrm{T}} \boldsymbol{\beta}_a / (c_m^{\mathrm{T}} S c_m)^{1/2}$ and variance 1. In addition, the random vector $(Z_1, \ldots, Z_M)^{\mathrm{T}}$ is asymptotically $M$-variate normal with covariance matrix

$$\left\{ \frac{c_m^{\mathrm{T}} S c_l}{\left( c_m^{\mathrm{T}} S c_m c_l^{\mathrm{T}} S c_l \right)^{1/2}}; \quad m, l = 1, \ldots, M \right\}. \quad (7)$$

The power under the $m$th candidate model is given by

$$\pi_m(D, \xi_1, \ldots, \xi_K; \boldsymbol{\beta}_m) = Pr\left( \max_{l = 1, \ldots, M} Z_l \geq q_{1 - \alpha} \middle| \boldsymbol{\beta} = \boldsymbol{\beta}_m \right). \quad (8)$$

The sample size to achieve a desired level of power can be determined accordingly.

To be specific, suppose that patients are recruited uniformly over the time period $[0, R]$ and the survival time for the $k$th dose group follows the exponential distribution with hazard $\lambda_k$. Then the number of patients in the $k$th group to achieve $D_k$ number of events is

$$D_k / \left\{ 1 - \frac{\exp^{-\lambda_k(\tau - R)} - \exp^{-\lambda_k \tau}}{\lambda_k R} \right\},$$

where $\tau$ is the endpoint of the study. This number needs to be adjusted upward to account for patient dropout (Lachin and Foulkes, 1986).

When evaluating (6) and (8), we need to evaluate the covariance matrices of the test statistics given in (5) and (7), respectively. Both covariance matrices involve $c_m$, which, as shown in (3), involves the estimated covariance matrix $\widehat{S}$. In the design stage, we replace $\widehat{S}$ in $c_m$ by $S$ for both (5) and (7). In the analysis stage, we use $\widehat{S}$ from the output of a standard software package.

Note that the covariance matrix $S_0$ is used to determine the critical value whereas the covariance matrix $S$ is used to calculate the optimal contrast coefficients and power. As explained in the last paragraph of the appendix, we replace $\boldsymbol{\beta}$ in the $p_k$'s by $\boldsymbol{\beta}/2$ when evaluating $S$, so as to obtain a more accurate power formula.

The standard two-sample comparison with a survival endpoint has the well-known events-driven property that the power of the log-rank test or Cox regression is driven by the total number of events (rather than the total number of patients) in the two groups. We have discovered that, in the same vein, the power of MCP-Mod for a survival endpoint is determined by the total number of events in the entire study rather than the number of events

in each dose group. This event-driven property is very useful in practice because only the total number of events in the entire study is observable and controllable before un-blinding.

## 3. Motivating Example and Simulation

We illustrate the proposed method for power calculation with a motivating example on chronic obstructive pulmonary disease (COPD). The primary endpoint was time to the first COPD exacerbation. There were four once-daily doses of the study medication: 5 mg, 25 mg, 50 mg and 100 mg, plus placebo, with an equal number of patients in each dose group. The median survival time for the placebo group was 0.5 years, and the hazard ratio on the optimal dose was 0.6. The desired one-sided type I error was set at 0.05.

We considered six dose-response curves, including linear, Emax, exponential, logistic, and betaMod; see Figure 1. We used the formulas given in Table 1 of Bornkamp et al. (2009) but changed the plus sign after $E_0$ to the minus sign because longer survival times corresponds to lower hazards. The mean response was defined as the log hazard for the time to COPD exacerbation. For the placebo group, the mean response was therefore $\log\{\log(2)/0.5\} \approx 0.327$. The mean response for the optimal dose within the dose range Was $\log\{\log(2)/(0.5/0.6)\} \approx -0.184$, corresponding to a treatment difference of 0.511. The model betMod captures a non-monotone dose-response relationship. Although such behavior is not common, it may occur occasionally for a variety of reasons. For example, a high dose may induce severe toxicities and thus result in non-compliance with the study medication. In this case, the dose-response relationship may be non-monotone if all patients are included in the analysis regardless of their compliance with the study medication. Using formula (8), we found that a total of 242 events would be required to achieve an average power of 85% across the six dose-response models.

To assess the accuracy of the approximation behind the power formula, we conducted a simulation study with 10,000 trials that mimic the motivating example. Specifically, we enrolled 75 patients in each dose group and generated times to COPD exacerbations from exponential distributions. Each trial was terminated once the required 242 total number of events had been reached, and the remaining patients were censored at that time point. There was no drop-out. For each of 10,000 simulated trials, we obtained $\hat{\beta}$ and $\hat{S}$ from the *coxph* function in R. We then calculated the test statistics given in (4) and obtained the proportion of rejections under each dose-response model.

Table 1 compares the theoretical and empirical values of the power under each of the six dose-response models, as well as the average power. To make the comparisons more comprehensive, we include the results for the hazard ratios of 0.4 and 0.8 in addition to 0.6. We also display the empirical type I error for the three scenarios under $\beta = 0$. In all three scenarios, the type I error is very close to the nominal level of 0.05. In addition, the empirical power is remarkably close to its theoretical counterpart, the two values being identical up the second decimal place in most cases. Thus, the proposed methods for power and sample-size calculations are accurate enough for practical use.

## 4. Discussion

In this paper, we present the MCP-Mod methodology for designing dose-finding studies with survival endpoints. Our main contribution lies in the derivation of an explicit form of the covariance matrix for the estimators of the log hazard ratios such that the power of MCP-Mod can be evaluated analytically. A second contribution is the assessment of the performance of the corresponding power and sample size formulas in realistic settings. An R code implementing the proposed methods is available upon request.

For each contrast test, the optimal contrast coefficients involve the covariance matrix of the parameter estimators. In the design stage, we use the analytic expression $S$. In the analysis stage, we typically use the covariance matrix estimated from the observed data. In small studies with rare events for some of the arms, the estimated covariance matrix may be unstable, in which case we recommend using $S$ instead. Note that the choice of the contrast coefficients affects the power, but not the validity of MCP-Mod.

We have focused on power and sample size calculations for hypothesis testing in a dose-ranging study. A related task is to identify the dose-response relationship. Specifically, it may be of interest to consider proper estimation of the dose-response curve or target doses (Bornkamp et al., 2007). The precision for such estimation depends on the covariance matrix of the parameter estimators (Pinheiro et al., 2014). In addition, when it comes to selecting the number or location of doses, the optimal design also depends on the covariance matrix or equivalently the information matrix (Bretz et al., 2010). The analytic form of the covariance matrix provided in this paper will be highly useful in all such circumstances.

The analytic form of the covariance matrix will be useful beyond dose-finding studies. For example, investigators may simply be interested in comparing $K$ different treatments, in which case our covariance matrix formula is directly applicable. A related problem is factorial designs (Lin et al., 2016), where the techniques given in the appendix of this paper can be used to derive power and sample size formulas.

## 5. Supplementary Materials

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

## Appendix

## Derivation of the Covariance Matrix

We adopt the counting-process martingale formulation: for $i = 1, \ldots, n$, let $N_i(t) = \Delta_i I(\tilde{T}_i \leq t)$, $Y_i(t) = I(\tilde{T}_i \geq t)$, and

$$M_i(t) = N_i(t) - \int_0^t Y_i(s) e^{\beta^T X_i} d\Lambda_0(s),$$

where $\Lambda_0(t) = \int_0^t \lambda_0(s) ds$ (Andersen and Gill, 1982). The $k$th component of $U(\beta)$ is

$$U_k(\beta) = \sum_{i=1}^n \int_0^\infty \{X_{ki} - E_k(\beta, t)\} dN_i(t),$$

where

$$E_k(\beta, t) = \sum_{i=1}^n Y_i(t) e^{\beta^T X_i} X_{ki} / \sum_{i=1}^n Y_i(t) e^{\beta^T X_i}.$$

Assume that the censoring distributions are approximately the same among the $(K+1)$ dose groups. Then by the law of large numbers,

$$E_k(\beta, t) \approx \frac{n_k e^{\beta_k} e^{-e^{\beta_k} \Lambda_0(t)}}{n_0 e^{-\Lambda_0(t)} + n_1 e^{\beta_1} e^{-e^{\beta_1} \Lambda_0(t)} + \cdots + n_K e^{\beta_K} e^{-e^{\beta_K} \Lambda_0(t)}}.$$

The right side can be approximated by

$$\frac{n_k e^{\beta_k}}{n_0 + n_1 e^{\beta_1} + \cdots + n_K e^{\beta_K}}$$

provided that the $\beta_k$'s are close to 0 or $\Lambda_0(t)$ is close to 0 (i.e., the event rate is low). Thus, $E_k(\beta, t) \approx p_k$.

Clearly,

$$U_k(\beta) = \sum_{i=1}^n \int_0^\infty \{X_{ki} - E_k(\beta, t)\} dM_i(t), \quad k = 1, \ldots, K.$$

By the counting-process martingale theory (Andersen and Gill, 1982), the variance of $U_k(\beta)$ is the expectation of its predictable variation

$$\sum_{i=1}^n \int_0^\infty \{X_{ki} - E_k(\beta, t)\}^2 Y_i(t) e^{\beta^T X_i} d\Lambda_0(t),$$

which can be written as

$$\sum_{m=0}^{K} e^{\beta_m} \sum_{i \in A_m} \int_0^{\infty} \{X_{ki} - E_k(\boldsymbol{\beta}, t)\}^2 Y_i(t) d \Lambda_0(t),$$

where $A_m$ denotes the set of patients in the $m$th group ($m = 0, 1, \ldots, K$). Upon the replacement of $E_k(\boldsymbol{\beta}, t)$ by $p_k$, the above expression becomes

$$\sum_{m=0}^{K} e^{\beta_m} \sum_{i \in A_m} \int_0^{\infty} (X_{ki} - p_k)^2 Y_i(t) d \Lambda_0(t),$$

which is equal to

$$(1 - p_k)^2 e^{\beta_k} \sum_{i \in A_k} \int_0^{\infty} Y_i(t) d \Lambda_0(t) + p_k^2 \sum_{m=0, m \neq k}^{K} e^{\beta_m} \sum_{i \in A_m} \int_0^{\infty} Y_i(t) d \Lambda_0(t).$$

By using the arguments for showing $E_k(\boldsymbol{\beta}, t) \approx p_k$, we can show that $\sum_{i \in A_k} Y_i(t) \approx \xi_k \sum_{i \in A_0} Y_i(t)$ ($k = 1, \ldots, K$). Thus,

$$\sum_{i \in A_k} \int_0^{\infty} Y_i(t) d \Lambda_0(t) \approx \xi_k \sum_{i \in A_0} \int_0^{\infty} Y_i(t) d \Lambda_0(t).$$

Since $E\{M_i(t)\} = 0$ for all $t$, $\sum_{i \in A_0} \int_0^{\infty} Y_i(t) d \Lambda_0(t)$ has the same expectation as $D_0$, where $D_0$ is the total number of events in the placebo group. Therefore,

$$Var\{U_k(\boldsymbol{\beta})\} \approx D_0 \left[ (1 - p_k)^2 \xi_k e^{\beta_k} + p_k^2 \left( 1 + \xi_1 e^{\beta_1} + \cdots + \xi_K e^{\beta_K} - \xi_k e^{\beta_k} \right) \right],$$

which, by the definition of $p_k$, is equal to

$$D_0 \left( 1 + \xi_1 e^{\beta_1} + \cdots + \xi_K e^{\beta_K} \right) \left\{ (1 - p_k)^2 p_k + p_k^2 (1 - p_k) \right\}.$$

Because $D_0 \xi_k e^{\beta_k}$ is approximately the number of events in the $k$th dose group, $D_0(1 + \xi_1 e^{\beta_1} + \cdots + \xi_K e^{\beta_K})$ is approximately the total number of events in the entire study. Hence,

$$Var\{U_k(\boldsymbol{\beta})\} \approx D p_k (1 - p_k).$$

The covariance between $U_k(\boldsymbol{\beta})$ and $U_l(\boldsymbol{\beta})$ $(k \neq l)$ is the expectation of their predictable covariation

$$\sum_{i=1}^{n} \int_0^\infty \{X_{ki} - E_k(\boldsymbol{\beta}, t)\}\{X_{li} - E_l(\boldsymbol{\beta}, t)\} Y_i(t) e^{\boldsymbol{\beta}^{\mathrm{T}} X_i} d\Lambda_0(t).$$

It then follows from the arguments in the previous paragraph that

$$\mathrm{Cov}\{U_k(\boldsymbol{\beta}), U_l(\boldsymbol{\beta})\} \approx p_k p_l \sum_{m=0, m \neq k, l}^{K} \sum_{i \in A_m} e^{\beta_m} \int_0^\infty Y_i(t) d\Lambda_0(t)$$

$$-(1 - p_k) p_l \sum_{i \in A_k} e^{\beta_k} \int_0^\infty Y_i(t) d\Lambda_0(t) - (1 - p_l) p_k \sum_{i \in A_l} e^{\beta_l} \int_0^\infty Y_i(t) d\Lambda_0(t).$$

The right side is approximately equal to

$$D_0\left\{p_k p_l\left(1 + \xi_1 e^{\beta_1} + \cdots + \xi_K e^{\beta_K} - \xi_k e^{\beta_k} - \xi_l e^{\beta_l}\right) - (1 - p_k) p_l \xi_k e^{\beta_k} - (1 - p_l) p_k \xi_l e^{\beta_l}\right\}$$

or

$$D\{p_k p_l(1 - p_k - p_l) - (1 - p_k) p_k p_l - (1 - p_l) p_k p_l\},$$

which is $-D p_k p_l$. Thus, $\mathscr{I}(\boldsymbol{\beta}) \approx D\{a_{kl}; k, l = 1, \ldots, K\}$, where $a_{kk} = p_k(1 - p_k)$, and $a_{kl} = -p_k p_l$ $(k \neq l)$. It can be verified by direct matrix multiplication that $\mathscr{I}^{-1}(\boldsymbol{\beta}) \approx D^{-1}\{b_{kl}; k, l = 1, \ldots, K\}$, where $b_{kk} = p_0^{-1} + p_k^{-1}$, and $b_{kl} = p_0^{-1}$ $(k \neq l)$.

The information matrix is used primarily in the estimation of the covariance matrix of the maximum (partial) likelihood estimator. It is also connected to hypothesis testing via the likelihood-ratio statistic, which is asymptotically $\chi^2$ distributed. Recall that the likelihood-ratio statistic for testing the null hypothesis that $\boldsymbol{\beta} = \mathbf{0}$ is $-2\log\{L(\mathbf{0})/L(\widehat{\boldsymbol{\beta}})\}$, where $L(\boldsymbol{\beta})$ is the (partial) likelihood for $\boldsymbol{\beta}$. Because the first derivative of $\log L(\boldsymbol{\beta})$ at $\boldsymbol{\beta} = \widehat{\boldsymbol{\beta}}$ is $\mathbf{0}$, the second-order Taylor series expansion yields

$$-2\log\{L(\mathbf{0})/L(\widehat{\boldsymbol{\beta}}))\} = \widehat{\boldsymbol{\beta}}^{\mathrm{T}} \mathscr{I}(\boldsymbol{\beta}^*) \widehat{\boldsymbol{\beta}},$$

where $\boldsymbol{\beta}^*$ lies between $\mathbf{0}$ and $\widehat{\boldsymbol{\beta}}$. Thus, evaluating the information matrix at $\boldsymbol{\beta}^*$ for the Wald statistic will result in a more accurate approximation to the desired $\chi^2$ distribution than evaluating it at $\mathbf{0}$ or $\widehat{\boldsymbol{\beta}}$. In the design stage, it is best to evaluate the information matrix at an intermediate value between $\mathbf{0}$ and $\boldsymbol{\beta}$, which is sensibly set to $\boldsymbol{\beta}/2$.

# References

Andersen PK and Gill RD (1982). Cox's regression model for counting processes: a large sample study. The Annals of Statistics 10, 1100–1120.

Bornkamp B, Bretz F, Dmitrienko A, Enas G, Gaydos B, Hsu CY, König F, Krams M, Liu Q, Neuenschwander B, Parke T, Pinheiro J, Roy A, Sax R, and Shen F (2007). Innovative approaches for designing and analyzing adaptive dose-ranging trials. Journal of Biopharmaceutical Statistics 17, 965–995. [PubMed: 18027208]

Bornkamp B, Pinheiro J, and Bretz F (2009). MCPMod: An R package for the design and analysis of dose-finding studies. Journal of Statistical Software 29, 1–23.

Bretz F, Dette H, and Pinheiro J (2010). Practical considerations for optimal designs in clinical dose finding studies. Statistics in Medicine 29, 731–742. [PubMed: 20213708]

Bretz F, Pinheiro JC, and Branson M (2005). Combining multiple comparisons and modeling techniques in dose-response studies. Biometrics 61, 738–748. [PubMed: 16135025]

Cox DR (1972). Regression models and life-tables (with discussion). Journal of the Royal Statistical Society Series B 34, 187–220.

Cox DR (1975). Partial likelihood. Biometrika 62, 269–276.

European Medicines Agency (2014). Qualification opinion of MCP-Mod as an efficient statistical methodology for model-based design and analysis of phase II dose finding studies under model uncertainty. http://goo.gl/imT7IT.

Lachin JM and Foulkes MA (1986). 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. [PubMed: 3567285]

Lin DY, Gong J, Gallo P, Bunn PH, and Couper D (2016). Simultaneous inference on treatment effects in survival studies with factorial designs. Biometrics 72, 1078–1085. [PubMed: 26991149]

Pinheiro J, Bornkamp B, and Bretz F (2006). Design and analysis of dose-finding studies combining multiple comparisons and modeling procedures. Journal of Bio-pharmaceutical Statistics 16, 639–656.

Pinheiro J, Bornkamp B, Glimm E, and Bretz F (2014). Model-based dose finding under model uncertainty using general parametric models. Statistics in Medicine 33, 1646–1661. [PubMed: 24302486]
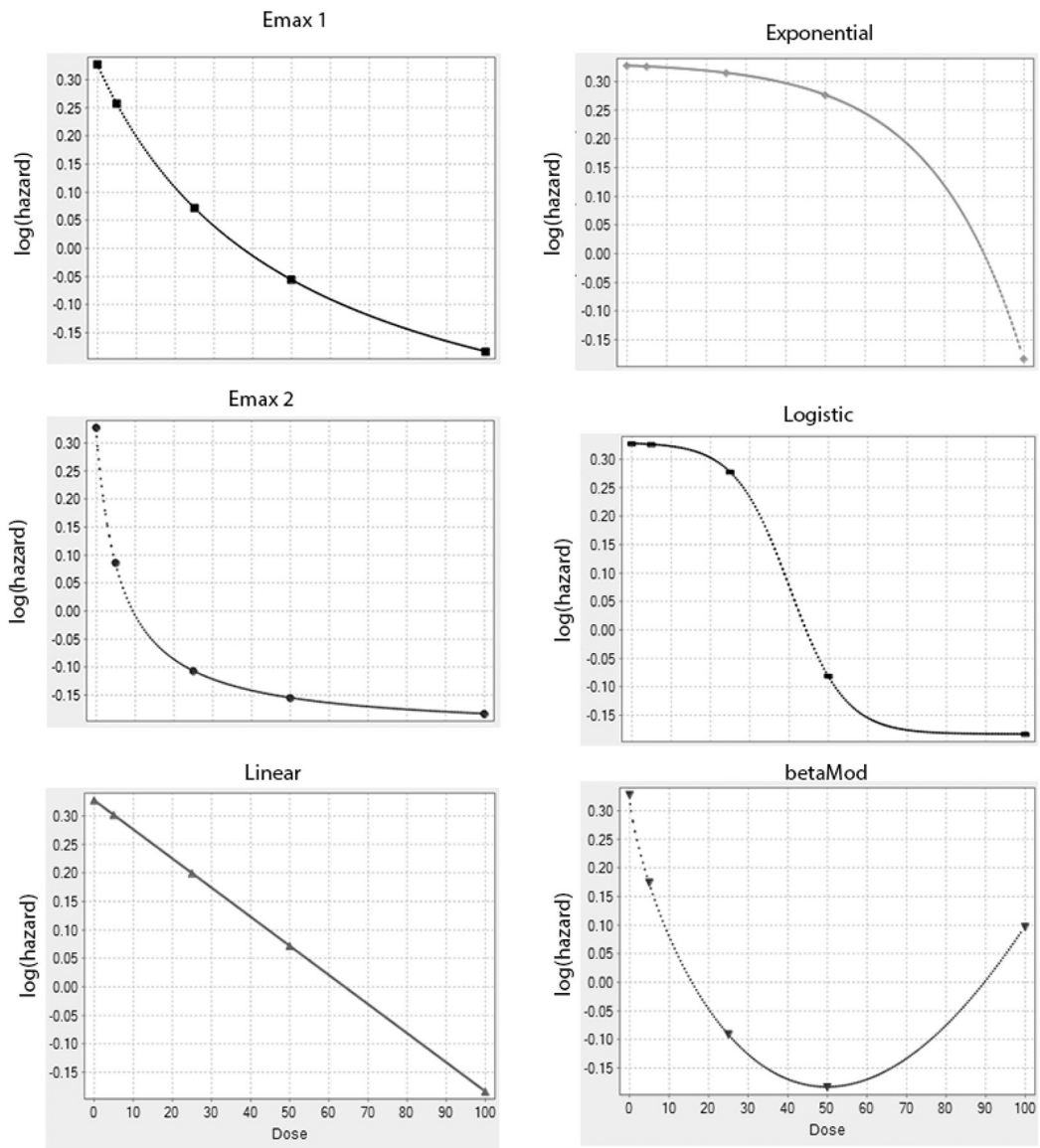
**Figure 1:**
Candidate dose-response relationships for the COPD dose-finding study: the six plots correspond to Emax model with $ED_{50} = 50$ (Emax 1), Emax model with $ED_{50} = 6.25$ (Emax 2), linear model, exponential model with $\delta = 22.756$, logistic model with $ED_{50} = 40.3287$ and $\delta = 6.9764$, and betaMod model with $D = 120$, $\delta_1 = 0.7489$, and $\delta_2 = 1.0485$. The figure was created in ADDPLAN DF, with the label customized for the display.

**Table 1.**

Theoretical and Empirical Power of the MCP-Mod

|  | Hazard ratio | | |
| --- | --- | --- | --- |
|  | **0.4** | **0.6** | **0.8** |
| Number of patients per group | 27 | 75 | 357 |
| Total number of events | 79 | 242 | 1240 |
| Empirical type 1 error | 0.049 | 0.048 | 0.054 |
| Emax1 |  |  |  |
|     Theoretical power | 0.873 | 0.863 | 0.859 |
|     Empirical power | 0.870 | 0.857 | 0.859 |
| Emax2 |  |  |  |
|     Theoretical power | 0.906 | 0.881 | 0.862 |
|     Empirical power | 0.897 | 0.872 | 0.854 |
| Linear |  |  |  |
|     Theoretical power | 0.823 | 0.827 | 0.833 |
|     Empirical power | 0.815 | 0.827 | 0.843 |
| Exponential |  |  |  |
|     Theoretical power | 0.778 | 0.811 | 0.836 |
|     Empirical power | 0.787 | 0.828 | 0.834 |
| Logistic |  |  |  |
|     Theoretical power | 0.913 | 0.917 | 0.921 |
|     Empirical power | 0.914 | 0.920 | 0.923 |
| BetaMod |  |  |  |
|     Theoretical power | 0.823 | 0.805 | 0.796 |
|     Empirical power | 0.806 | 0.804 | 0.803 |
| Average |  |  |  |
|     Theoretical power | 0.853 | 0.851 | 0.851 |
|     Empirical power | 0.848 | 0.851 | 0.853 |