Statistics
in Medicine

Research Article

Received 13 November 2013, Accepted 1 June 2015 Published online 28 June 2015 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.6569

Optimal treatment allocation for placebo-treatment comparisons in trials with discrete-time survival endpoints

Mirjam Moerbeek[a*†] and Weng-Kee Wong[b]

In many randomized controlled trials, treatment groups are of equal size, but this is not necessarily the best choice. This paper provides a methodology to calculate optimal treatment allocations for longitudinal trials when we wish to compare multiple treatment groups with a placebo group, and the comparisons may have unequal importance. The focus is on trials with a survival endpoint measured in discrete time. We assume the underlying survival process is Weibull and show that values for the parameters in the Weibull distribution have an impact on the optimal treatment allocation scheme in an interesting way. Additionally, we incorporate different cost considerations at the subject and measurement levels and determine the optimal number of time periods. We also show that when many events occur at the beginning of the trial, fewer time periods are more efficient. As an application, we revisit a risperidone maintenance treatment trial in schizophrenia and use our proposed methodology to redesign it and compare merits of our optimal design.

Keywords: Cox regression model; design efficiency; longitudinal study; multiple-objective optimal design

1. Introduction

The optimal allocation of subjects to treatment groups is an important issue in the design of a randomized controlled trial. Many trials use equal allocation, but this is not necessarily the best design. There are various reasons to allow for an unequal allocation scheme to the various treatment groups. A main and frequent reason is because some treatments may be more expensive than others, or they are less readily available [1–3]. The practical implication is that if some treatment protocols are more expensive than others, it is cost-efficient to allocate more subjects to the cheaper protocols, and a larger number of subjects can be recruited at the same trial cost. Other reasons for unbalanced trials are to gain more experience with the new treatment or to expose fewer subjects to a treatment with potentially harmful side effects.

From a statistical point of view, trials with unbalanced group sizes may be preferred if such designs provide narrower sets of confidence intervals for the parameters of interest or more powerful tests of the treatment effects. Optimal design methodology can provide guidance on group sizes that provide the best possible statistical inference at minimum cost. Some papers on optimal allocation schemes for two group randomized trials are available. For example, Allison *et al*. [4], Dette [5], Hsu [6], and Schouten [7] concern individually randomized trials; Liu [8] deals with cluster randomized trials and multisite trials; and Moerbeek and Wong [9] describe trials with clustering in only one of the two treatment arms. An extension to trials with more than two treatment arms is in [10, 11]. All assumed a continuous or dichotomous endpoint that was measured at one point in time after treatment was administered.

In longitudinal trials, subjects are followed over a period of time to study changes to an outcome, which may be their opinion on a subject matter, a behavioral change, health habit, and so on. A typical type of endpoint in longitudinal trials is the survival endpoint, which measures if and when a specific event

[a]*Department of Methodology and Statistics, Utrecht University, Utrecht, The Netherlands*
[b]*Department of Biostatistics, University of California at Los Angeles, U.S.A.*
Correspondence to: Mirjam Moerbeek, Department of Methodology and Statistics, Utrecht University, P.O. Box 80140, 3508 TC Utrecht, The Netherlands.
†*E-mail: M.Moerbeek@uu.nl*

occurs. Such events may include death, smoking initiation, or treatment termination. Data of such trials are analyzed using survival analysis, which is also referred to by the less pejorative term event history analysis. The common statistical model is the Cox regression model [12], and sample size issues for comparing one treatment group with a placebo group given a pre-defined allocation ratio are discussed in [13–17], among others. There are only a few papers that discuss optimal allocation problems for two-arm survival trials [18, 19].

The Cox regression model assumes the timing of the event can be measured very precisely, while this is not always the case in practice. In many cases, survival data are measured discretely rather than continuously. Survival data that are measured discretely are called discrete-time or grouped-time survival data regardless of whether the underlying survival process is continuous or discrete. Such data can be analyzed using a generalized linear model [20, 21]. Optimal designs for such data have only been studied recently and restricted to the comparison of one treatment group with a placebo group. The early focus was on the optimal number of subjects and time periods in trials with an equal allocation of subjects to treatment conditions [22–24]. Subsequent work considered optimal treatment allocations and the loss of efficiency in using a trial with equal-sized treatment groups [25].

A natural extension is to study optimal treatment allocation in trials when we want to compare multiple treatment groups with a placebo group, and each comparison has an unequal emphasis. For example, the comparison of the placebo group with the treatment 1 group is more important than the comparison of the placebo group with the treatment 2 group, which in its turn is more important than the comparison of the placebo group with the treatment 3 group, and so forth. A multiple-objective optimal design provides the optimal allocation scheme of subjects to treatment groups such that more important comparisons have higher user-specified efficiencies. After satisfying all the user-specified efficiency requirements, the multiple-objective optimal design estimates the least important comparison with as high efficiency as possible. Of course, if the efficiencies are set too high, an optimal design will not exist because of the competing nature of the multiple objectives. A recent review on advances in optimal allocation designs for multi-arm clinical trials with multiple objectives is in Sverdlov and Rosenberger [26].

The aim of this paper is to construct multiple-objective optimal designs for trials with discrete-time survival endpoints. As an example, we demonstrate our methodology using the flexible Weibull survival function, which allows increasing, constant, or decreasing risk of event occurrence across time. The focus is on a trial with 12 time periods (i.e., months in a year), but we also provide optimal allocations for trials with fewer time periods. Additionally, we study the effect of the number of time periods on design efficiency when different cost structures are imposed on subject enrollment and measurement costs at each time period. As an application, we revisit a risperidone maintenance treatment trial in schizophrenia and illustrate how optimal design methodology can capture study goals more realistically and, at the same time, provide more accurate inference on the effects of various dose reduction schemes in the trial.

2. Statistical model for discrete time-survival data

The survival time T of a subject in a trial is a continuous random variable, and the corresponding survival function is defined as $S(t) = P(T > t)$. By definition, $S(0) = 1$, $S(t)$ monotonically decreases as time increases, and it cannot be negative. For discrete time survival data, survival status is only measured at the end of each time period at time points $t_k (k = 1, \dots, p)$. This implies that it is only known if a subject experienced the event in period k, not at which time within the period the event occurred. Given that the event has not occurred in previous time periods, the hazard probability for period k is defined as $h(t_k) = P(T = t_k | T \geqslant t_k)$ with T a discrete random variable for time. We observe that the hazard probability is a conditional probability given by $h(t_k) = [S(t_{k-1}) - S(t_k)]/S(t_{k-1})$ with values in the range $[0, 1]$. For a fair comparison of hazard probabilities across time periods, it is important that the time periods are of equal length.

The probability of event occurrence may vary across time periods and subjects. To explain part of the between-subject variability in hazard probability, we model it as a function of between-subject and within-subject predictor variables by formulating a generalized linear model for each time period and combining these models in one single equation. To fix ideas, we restrict ourselves to models with treatment assignment as the sole explanatory variable but, of course, such models can be readily extended to include some covariates from the subject. Subjects are randomly allocated to either the placebo or one out of q active treatments, and all observations are assumed to be independent. The model for time period $k = 1, \dots, p$ for subject $j = 1, \dots n_i$ in treatment $i = 1, \dots, q$ is given by

$$g\left[h(t_{ijk})\right] = \eta_{ijk} = \sum_{k=1}^{p} \alpha_k D_{jk} + \sum_{i=1}^{q} \beta_i Z_{ij}. \tag{1}$$

Here, g is a link function and η_{ijk} is the linear predictor. The dummy variable D_{jk} is equal to 1 if the observation is in time period k and 0 otherwise. The regression coefficient β_i is the mean difference in the linear predictor η_{ijk} between subjects in the placebo group and those in treatment condition i. Our model has two simplifying assumptions: (i) the effect of each treatment does not vary across time periods, and so β_i does not have a subscript k; and (ii) there are no cross-over treatments because the time period indicator k does not appear in the subscript of Z_{ij}, and so treatment group membership is the same across time. The dummy variable Z_{ij} takes on the value 1 if subject j is randomized to receive treatment i and equals 0 otherwise. Subjects for whom all dummy variables Z_{ij} are equal to 0 are in the placebo group. This implies that the baseline hazard probability $g^{-1}(\alpha_k)$ in time period k is the hazard probability for subjects allocated to the placebo. The collection of all the values of $g^{-1}(\alpha_k)$, $k = 1, 2, \ldots, p$, is the set of baseline hazard probabilities.

To estimate the model parameters, the data should be presented in the person–period format. Each subject provides data for all time periods until and including the time period during which he or she experiences the event or drops out from the study for reasons other than event occurrence or because the study has reached its termination date. The event indicator y_{ijk} is binary and takes on the value 1 if the event has occurred and 0 otherwise. For fitting model (1), standard software for fitting generalized linear models can be employed. This is carried out by means of iteratively weighted least squares [27] after rewriting model (1) in matrix notation with $g\,[(h(t)] = X\theta$, where $h(t)$ is a vector of probabilities of length $\sum_{k=1}^{p} \sum_{i=0}^{q} n_{ik}$ with n_{ik} the number of subjects in treatment condition i who enter period k, with the placebo indicated $i = 0$. The link function g is applied to each element of the vector $h(t)$. The sample size n_{ik} is calculated from $n_{ik} = n_{i0} S_i(t_{k-1})$ with n_{i0} the number of subjects in condition i at the beginning of the study and $S_i(t_k)$ the survival function in condition i at the end of time period k. The total number of subjects required for the trial is predetermined and equal to $n = \sum_{i=0}^{q} n_{i0}$.

The $(p + q) \times 1$ vector of unknown parameters in the model is $\theta = (\alpha_1, \ldots \alpha_p, \beta_1, \ldots, \beta_q)'$. Its estimator has asymptotic covariance matrix $var(\hat{\theta}) = (X'WX)^{-1}$, where W is a diagonal matrix of dimension $\sum_{k=1}^{p} \sum_{i=0}^{q} n_{ik}$ with weights w_{ijk} that depend on the hazard probabilities $h(t_{ijk})$ and the link function g. For example, if we have the logit link $g\left[h(t_{ijk})\right] = ln\left\{h(t_{ijk})/[1 - h(t_{ijk})]\right\}$, the weights are $w_{ijk} = h(t_{ijk})[1 - h(t_{ijk})]$. All subjects at the same time period in the same treatment arm have the same weights because we do not consider covariates of the subjects other than the treatment group the subject is assigned to. The design matrix X has $\sum_{k=1}^{p} \sum_{i=0}^{q} n_{ik}$ rows and $(p + q)$ columns. Each row represents an observation from a subject at a particular period, and the columns are such that the first p columns represent indicators for the time period and the next q columns correspond to treatment group indicators. Specifically, each row in the X matrix has length $p + q$ with a value 1 at positions k and $p + i$ if the observation is from period k and the subject is in the i^{th} treatment group. If the observation comes from period k and the subject is in the placebo group, the row has a value 1 at position k and a value 0 elsewhere. We observe that the covariance matrix depends on the baseline sample sizes n_{i0} in the placebo and treatment conditions, the survival functions $S_i(t_{k-1})$ and the hazard probabilities through the weights w_{ik}.

3. Optimal allocations and multiple-objective optimal designs

In this section, we determine a multiple-objective optimal design for assigning subjects to various groups using the relationship between constrained optimal designs and compound optimal designs described in Cook and Wong [28]. The approach is in general a graphical one using the idea of an efficiency plot. The optimal designs developed in this paper are suitable for multiple comparison problems with treatment group as a classification factor and a pre-determined number of treatment groups and total sample size. The solution to the optimization problem determines the number of subjects to be assigned to each group. This is in contrast to finding optimal designs for dose response models defined on a pre-specified continuous dose interval, when the number of doses, the optimal doses, and the probability mass at each dose have to be additionally determined.

Specifically, we are concerned with approximate designs, which means that instead of finding the optimal number of subjects to be assigned for each treatment, we determine the optimal proportion of subjects to be assigned for each treatment. Such designs are characterized by the proportion π_i of subjects assigned to the i^{th} treatment group $i = 1, \ldots, q$ and the placebo group with $i = 0$. These proportions are called design weights and satisfy $0 \leqslant \pi_i \leqslant 1$ for all i and $\sum_{i=0}^{q} \pi_i = 1$.

An obvious objective in the design of a treatment-placebo study is to find a design ξ that provides the most accurate estimate of the effect of each of the q treatments relative to the placebo group. The differential effect for the i^{th} treatment group compared with the control group is estimated by $\hat{\beta}_i$ in model (1), and the quality of the estimate is measured by its variance. A suitable design criterion is to find a design that minimizes this or equivalently, one seeks a design to maximize the design efficiency given by $E_i(\xi) = var\left(\hat{\beta}_i;\xi_i^*\right)/var\left(\hat{\beta}_i;\xi\right)$, where $var\left(\hat{\beta}_i;\xi_i^*\right)$ and $var\left(\hat{\beta}_i;\xi\right)$ are the variances obtained from the optimal design ξ_i^* and the design ξ, respectively. This ratio compares the merit of the design ξ relative to the optimal design for estimating the effect of the i^{th} treatment and clearly is a number between 0 and 1. If the ratio is 1/2, this means that the design has to be replicated twice to do as well as the optimal design for estimating β_i. In practice, we always want to implement designs with high efficiencies. When there are multiple comparisons, one seeks a design to minimize a weighted combination of such differences with weights representing the importance of each of the differences, or equivalently, finds a design that maximizes a linear weighted combination of efficiencies.

A placebo-treatment study usually has several objectives of different importance. In our example, some of the q treatments to be compared with the placebo may be first line drugs, and the rest are second line drugs. If there is greater interest in the former group, this suggests that we want a design to provide more accurate estimates of the effects of the first line drugs than second line drugs. This leads to a dual-objective optimal design problem, where the first objective is to compare first line drugs with the placebo group, and the second objective is to compare second line drugs with the placebo with user-specified efficiency for the first set of comparisons. Our objectives are to minimize the variances of the estimators for the various differential effects by choice of a design or equivalently maximize the efficiencies of the estimators. Because these variances can be on different scales, we recommend working with standardized optimality criteria defined in terms of weighted linear combinations of the $-(E_i(\xi))^{-1}$'s.

For multiple-objective optimal design problems, we assume that the i^{th} objective $\phi_i(\xi)$ can be formulated as a concave function of the information matrix and in terms of the efficiency $E_i(\xi)$ of the design $\xi, i = 1, \dots, q$. As an example, if our goal is to minimize the variance of the estimate of the i^{th} treatment effect, we set $\phi_i(\xi) = -(E_i(\xi))^{-1}$; see Example 1 in [28]. Further, we assume that these objectives can be ordered in terms of their importance so that the multiple-objective optimal design can provide higher efficiencies for the more important objectives. More specifically, let e_i be the user-specified efficiency required of the design for the i^{th} objective and let $e_i \geqslant e_{i+1}, i = 1, 2, \dots, q-1$. Our constrained optimization problem is then to find a constrained optimal design that maximizes $E_q(\xi)$ among all designs ξ that have efficiencies at least e_i for the i^{th} objective, $i = 1, 2, \dots, q-1$. Of course, if the efficiencies sought are too high, the multiple-objective optimal design may not exist. Such constrained optimal design problems are relatively easy to formulate but difficult to solve. However, they can be found indirectly from compound optimal designs defined as follows. For a fixed vector $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_q)'$ with $0 \leqslant \lambda_i \leqslant 1$ and $\sum_{i=1}^{q} \lambda_i = 1$, the compound optimal design ξ_λ maximizes the function

$$\Phi(\xi|\lambda) = \sum_{i=1}^{q} \lambda_i \phi_i(\xi). \tag{2}$$

Compound optimal designs are easier to find than the constrained optimal designs because, for fixed λ, a convex combination of concave criteria is still concave, and so standard algorithms can be directly used to find optimal designs for the single concave criterion. For each λ, the compound optimal design is generated, and the desired constrained optimal design is then found indirectly from reviewing all the compound optimal designs and determining which one satisfies the constraints. Cook and Wong [28] provide details.

When $q = 2$, we have a dual-objective optimal design problem, which is our focus here. In this case, λ is a scalar, and we may let $\lambda_1 = \lambda$ and $\lambda_2 = 1 - \lambda$. If $\lambda = 0$, the dual-objective design criterion reduces to $\Phi(\xi|\lambda) = \phi_2(\xi)$, which measures the accuracy of the inference for comparing the treatment 2 group with the placebo group provided by the design ξ. The proportion of subjects assigned to the treatment 1 group is zero in this case but will increase with increasing values of λ. Conversely, the proportion of subjects in the treatment 2 group decreases to zero as the value of λ increases to 1. Our formulation ensures that the efficiency of the compound optimal design under the first objective increases with increasing values of λ but decreases under the second objective with increasing values of λ.

For dual-objective optimal design problems, the relationship between compound optimal designs and constrained optimal designs can be found analytically or graphically. In the latter case, this is carried out by first constructing an efficiency plot, where the two types of efficiencies are graphed against the values of λ between 0 and 1. We then draw a horizontal line where $E_1(\xi) = e_1$ is and read off the corresponding λ. The sought design is then ξ_λ with $E_1(\xi_\lambda) = e_1$. The graphical method relies on an efficiency plot, and an example of such a plot is given in the next section for a two-objective optimal design.

In general, compound optimal designs have to be found numerically. For each fixed λ, we used the function `constrOptim.nl` in the `R` package Alabama [29] to maximize the function $\Phi(\xi|\lambda)$ in (2) subject to the inequality constraints $0 \leqslant \pi_i \leqslant 1$ for all i and the equality constraint $\sum_{i=0}^{q} \pi_i = 1$. Our search procedure uses the adaptive barrier method, which we briefly describe in the supplement.

4. Results for the Weibull survival function

The covariance matrix $var(\hat{\theta})$ depends on the design proportions π_i and on the survival function $S(t)$. There are many survival functions, and our optimal design methodology described in Section 3 is applicable to all of them. For illustration, we use the popular Weibull survival function, which is very flexible as it allows for increasing, decreasing, and constant hazard rates across time. The continuous time survival function is given by $S(t) = \exp(-\gamma t^{\tau})$ and the hazard rate by $h(t) = \gamma \tau t^{\tau-1}$. The Weibull survival function has two parameters: $\gamma \in [0, +\infty)$ is a scale parameter, and $\tau \in [0, +\infty)$ is a shape parameter. For $\tau > 1$ the hazard rate increases across time, for $\tau < 1$ it decreases and for $\tau = 1$ it is constant across time. The time variable t is the amount of time that has elapsed in the study. Time is rescaled in such a way that the minimal value $t = 0$ corresponds to the beginning of the study, and the maximal value $t = 1$ corresponds to the end of the study. For convenience, the scale parameter γ is replaced by $-\log(1-\omega)$, and ω is the proportion of subjects who have experienced the event by the end of the study (i.e., by $t = 1$). Under this reparametrization, the survival function is $S(t) = (1-\omega)^{t^{\tau}}$, and the hazard rate is $h(t) = -\tau t^{\tau-1} \log(1-\omega)$. Survival in discrete time is calculated by evaluating $S(t)$ at each discrete time point at $t = t_k = k/p$, the value of the time variable t at the end of period k. The hazard probability is $h(t_k) = \left\{S(t_{k-1}) - S(t_k)\right\} / S(t_{k-1})$.

We now derive optimal designs for a trial with one placebo and two treatment conditions. We assumed the survival function in the placebo group has a Weibull distribution with three possible values for each of the parameters τ and ω : $\tau = 0.5$, 1, and 2 and $\omega = 0.25$, 0.5 and 0.75. We construct optimal designs for a 12-month trial with 12 time periods and compare them with designs with fewer time periods. For instance, we study the effect of having 6 or 12 time periods on the survival and hazard probabilities in the study. The logit link function $g[h(t_{ijk})] = \log\left\{h(t_{ijk})/[1-h(t_{ijk})]\right\}$ is used to relate the hazard probability to treatment condition. The hazard probability function in the placebo condition is assumed to have a Weibull distribution, and the two treatment effects β_1 and β_2 give the differences between the hazard probabilities in the placebo and the two treatment groups on the logit scale. We considered two possible sets of treatment effects: $(\beta_1, \beta_2) = (-0.5, -1)$ and $(\beta_1, \beta_2) = (0.5, 1)$. The two negative treatment effects in the first set imply that both treatments decrease the probability of event occurrence. An increased risk is represented by the second set. For both sets, the effect of the first treatment (β_1) is smaller in absolute value than that of the second (β_2). Even though the two sets only differ in the direction of the effects, they are not the same because the hazard probabilities in the two treatments do not only depend on β_1 and β_2 but also on the baseline hazard probabilities.

There are two objectives in this problem, that is, find a design to maximize the efficiencies of the two treatment effect estimators $\hat{\beta}_1$ and $\hat{\beta}_2$. Optimal designs are found when the comparison of the first or the second treatment is of primary importance. Accordingly, the two single objectives are $\phi_1(\xi) = -(E_1(\xi))^{-1}$ and $\phi_2(\xi) = -(E_2(\xi))^{-1}$, and the compound optimality criterion is $\Phi(\xi|\lambda) = \lambda\phi_1(\xi) + (1-\lambda)\phi_2(\xi)$ with $0 \leqslant \lambda \leqslant 1$. We note that the two locally dual-objective optimal designs for the two cases are not the same unless they have the same nominal values with $\beta_1 = \beta_2$. To find them, we make the following assumptions: (i) the two treatment effects have nominal values $(\beta_1, \beta_2) = (-0.5, -1)$; (ii) comparing treatment 1 group with the placebo group is of primary interest; (iii) there are $k = 12$ periods; (iv) $\tau = 1$; (v) $\omega = 0.5$; and (vi) select a grid with a step size of 0.001 to discretize the range of values for λ in [0, 1].

Figure 1 shows features of the dual-objective optimal design. The optimal design weights π_i $(i = 0, 1, 2)$ as a function of λ are shown in the left panel. For this example, the proportion of subjects in the placebo group depends only slightly on the value of λ, and the design is unbalanced with varying proportions π_i across the three treatment groups. The right panel of Figure 1 shows the efficiency plot, where efficiencies of the compound optimal designs ξ_λ for the two objectives are plotted against the values of λ.

The figure also shows that the two efficiency graphs intersect near $\lambda = 0.5$ suggesting that the compound optimal design with $\lambda = 0.5$ is about equally efficient under both criteria. This approximate value of $\lambda = 0.5$ was also found for other values of k, ω, τ, β_1, and β_2 that we studied. It is, however, not always

Statistics in Medicine

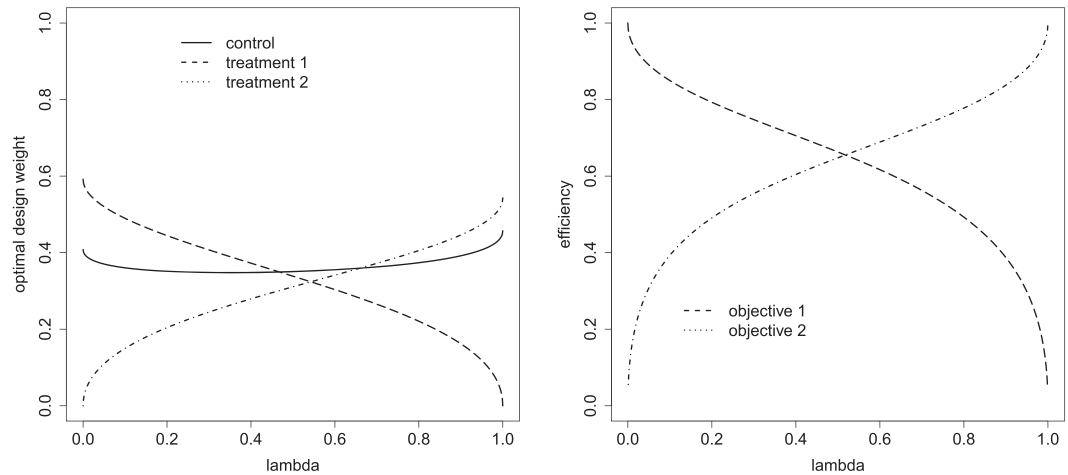

Figure 1. Optimal proportions and efficiency plot for two-objective optimal designs with $k = 12$ periods, $\tau = 1$, $\omega = 0.5$, and the treatment effects are $(\beta_1, \beta_2) = (-0.5, -1)$.

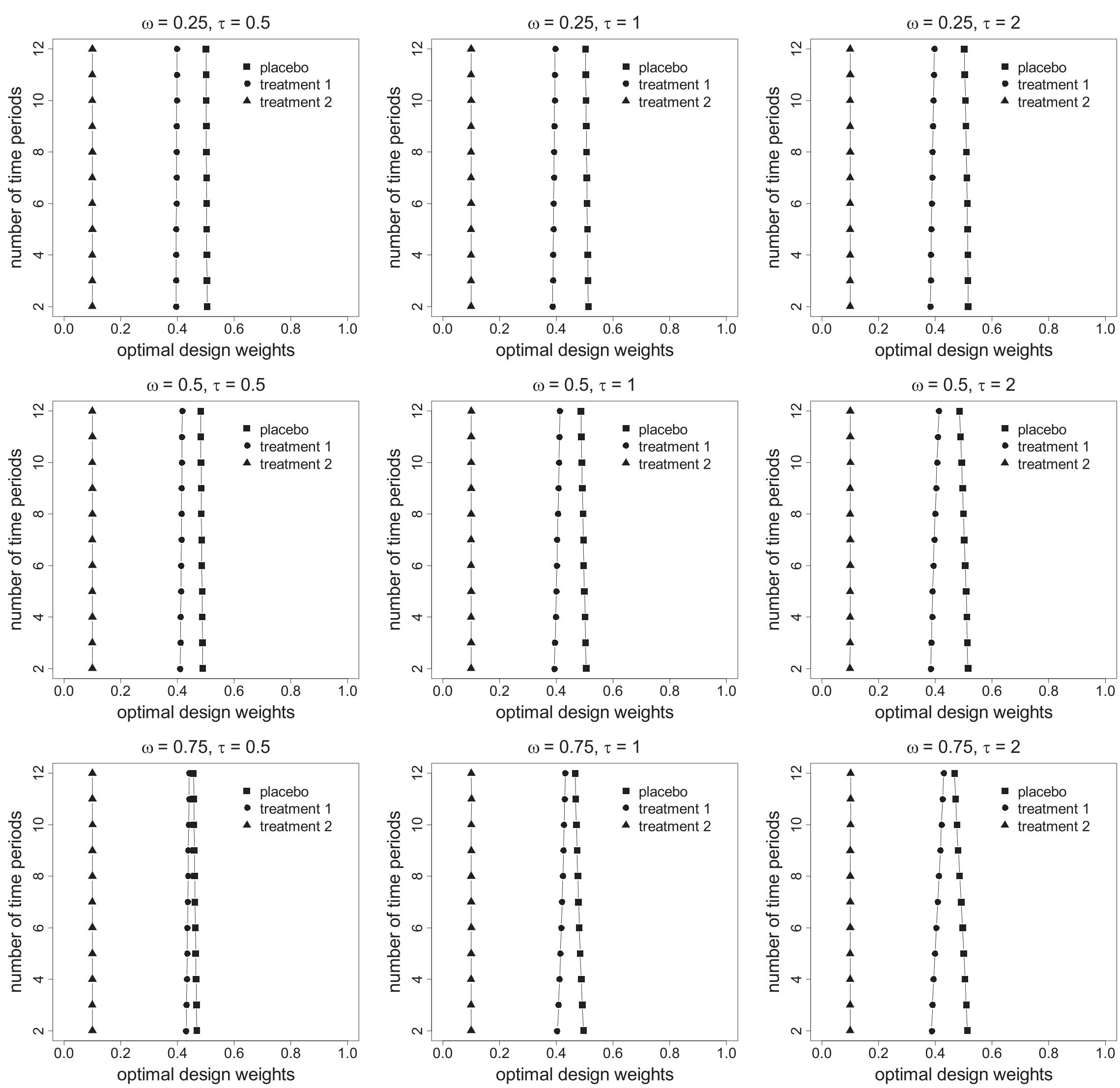

Figure 2. Optimal design weights as a function of the number of time periods p, the proportion of event occurrence ω, and the shape parameter τ for the case where $(\beta_1, \beta_2) = (0.5, 1)$. Primary objective is the comparison of treatment 1 with the placebo, and the related efficiency is 0.9.

found in dual-objective optimal design problems. We note in this problem, the two objectives are highly competitive because large efficiencies cannot be achieved for both of them simultaneously. In this particular example, if we want a design that has an efficiency of 0.9 for the first objective, we have from the efficiency plot, $\lambda = 0.966$, and note that the corresponding efficiency of the compound optimal design for the second objective is low and approximately equal to 0.26. More generally, the two objectives were found to be competitive for all values of k, ω, τ, β_1, and β_2 that we studied.

Figures 2 and 3 display how optimal design weights for positive treatment effects $(\beta_1, \beta_2) = (0.5, 1)$ change when the model assumptions change. Figure 2 assumes that the primary objective is to compare treatment 1 with the placebo, and Figure 3 assumes the primary objective is to compare treatment 2 with the placebo; in either case, the sought efficiency of the generated design for the primary criterion is 0.9. In each figure, the number of time points is plotted versus the optimal design weights for treatments 1 and 2 and the placebo group. The nine subfigures in each figure show how optimal design weights change when a specific parameter in the design problem is changed. Effects of the changes in the optimal weights because of changes in ω are displayed row-wise, and the corresponding changes because of changes in τ are displayed column-wise. There are no upper and lower bounds other than 0 and 1 for these treatment weights. Because the two treatment effects β_1 and β_2 are different, the optimal design weights in Figures 2

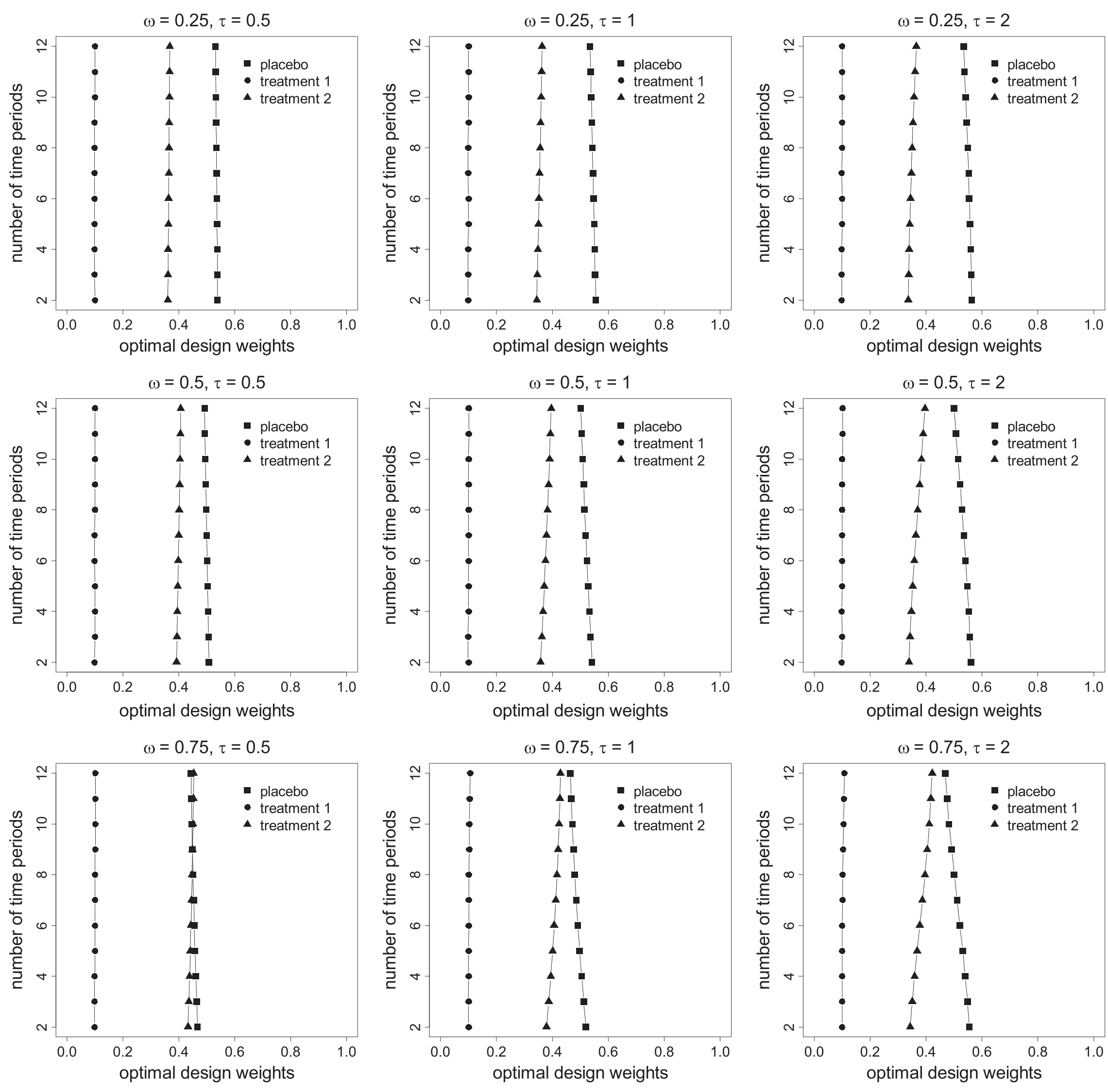

Figure 3. Optimal design weights as a function of the number of time periods p, the proportion of event occurrence ω, and the shape parameter τ for the case where $(\beta_1, \beta_2) = (0.5, 1)$. Primary objective is the comparison of treatment 2 with the placebo, and the related efficiency is 0.9.

and 3 are also different and are far from being equal, suggesting that very different proportions of subjects are to be assigned to the various groups.

A weight near 0.1 is assigned to the treatment group whose comparison to the placebo is of least importance, and this weight is hardly influenced by the number of time periods k, the shape parameter τ, and the proportion of event occurrence ω. This weight was also near 0.1 when both treatment effects were negative with $(\beta_1, \beta_2) = (-0.5, -1)$. As expected, a larger weight is assigned to the treatment whose comparison with the placebo is more important, and this weight is almost always smaller than the weight for the placebo group. The weight assigned to the placebo group seems to be always near 0.5, and in almost all cases, more subjects are assigned to the placebo group. Such an optimal design may appear counterintuitive to clinicians, but assigning more subjects to the placebo group makes sense because the placebo group is used in all comparisons, and so having more subjects in the placebo group provides a more accurate estimate of its effect. More specifically, if we have a total of n subjects for a placebo-treatment study and we are equally interested to compare the k treatment group means with the placebo group mean, allocating $n/(1+\sqrt{k})$ number of subjects to the placebo group and the rest equally to the k treatment groups minimizes the sum of the variances of the estimated treatment effects (see page 116 of [30]). This implies that more subjects are assigned to the placebo group. For our design problems, it is therefore not surprising that about 50% of subjects are assigned to the placebo group. The exact percentage depends on the relative interest in the various objectives and the nominal values of the model parameters. For instance, in our problem when $\beta_1 = -0.5$ and $\beta_2 = 1$, our calculation shows about 40% of subjects are optimally allocated to the placebo group.

The aforementioned figures show that overall, the weights of the optimal design are hardly influenced by the number of time periods k when the shape parameter τ and/or the proportion event occurrence ω are small. This implies that changing the number of time periods during the course of the trial hardly influences the optimal design in these cases. However, when we have larger values of τ and/or ω, we observe larger changes in the optimal design weights as the number of time periods increases. Further, we note that for any value ω, the difference in weights between the placebo and the treatment of primary interest increases when τ increases. On the other hand, for any value τ, the difference in weights between the placebo and the treatment of primary interest group decreases when ω increases. Special attention should be paid to the case where $\omega = 0.75, \tau = 0.5$, and the comparison of the second treatment with the placebo is of primary importance (lower left panel in Figure 3). In this case, the difference in weights between the placebo and the treatment 2 group is small. When there is a small number of time periods, we observe the optimal strategy is to assign more weight to the placebo group, and when there is a large number of time periods, more weight is assigned to treatment 2.

The results change somewhat when another level of efficiency is required for the primary objective. When this efficiency is 0.8, then the weight for the treatment whose comparison with the placebo is least important is always near 0.2 (figures not shown). Again, the weights for the other treatment and placebo depend on the underlying survival function and the number of time periods. Similar figures were drawn for negative treatment effects $(\beta_1, \beta_2) = (-0.5, -1)$. It appeared that the number of time periods k, the shape parameter τ, and the proportion of event occurrence ω have smaller or negligible effects on the optimal allocation scheme. For this reason, we do not display these figures.

We compared the efficiencies of equal-weight designs relative to the optimal designs found in Figures 2 and 3. For all scenarios and for both objectives, the efficiencies of the equally weighted design range between 0.7 and 0.75, suggesting that the equally weighted design does not perform well relative to the unequally weighted optimal design.

We close this section by noting that an advantage of the efficiency plot in a dual-objective design problem is that if the roles of the two objectives are interchanged, that is, the primary objective becomes the secondary objective, and vice versa, the new optimal allocation scheme can be readily worked out. One simply plots the same efficiencies against $1-\lambda$ instead of λ and deduce the new constrained optimal design in the same way as before to obtain the sought dual-objective optimal design. This strategy applies for any dual-objective optimal design problem with concave optimality criteria.

5. Effect of the number of time periods

In the previous section, the optimal proportions were presented for 2–12 time periods. Increasing the number of time periods results in smaller variances $var(\hat{\beta}_1)$ and $var(\hat{\beta}_2)$ but also in larger trial cost because subjects are measured more often, and costs are associated with taking repeat measurements. To select the most cost-efficient number of time periods, the costs as a function of the number of time periods should be taken into account. We consider two cost functions in this paper. Cost function 1 assumes all the n subjects are measured at all time periods, even if they experience the event prior to the last time period. Such a cost function is realistic in trials where not only the event is of interest but also secondary measurements. An example is a school-based smoking prevention intervention where not only the target event, smoking initiation, is observed but also secondary measurements such as knowledge on smoking and health. The costs of taking one measurement are denoted c_m. The costs to treat a subject vary across the treatment conditions, and they are denoted by c_0, c_1, and c_2 for the placebo, treatments 1, and 2, respectively. Our cost function 1 is then calculated as

$$C = n(\pi_0 c_0 + \pi_1 c_1 + \pi_2 c_2) + n(p+1)c_m. \tag{3}$$

In the aforementioned equation, we note that the number of measurements per subject is $(p+1)$ because we also have a measurement at baseline. With cost function 2, we assume that subjects leave the study once they have experienced the event, and measurements are not taken after event occurrence. This implies that for the same budget more subjects can be recruited. Cost function 2 is defined as

$$C = n(\pi_0 c_0 + \pi_1 c_1 + \pi_2 c_2) + n\left(\pi_0 \sum_{k=0}^{p} S_0(t_k) + \pi_1 \sum_{k=0}^{p} S_1(t_k) + \pi_2 \sum_{k=0}^{p} S_2(t_k)\right) c_m, \tag{4}$$

where $S_0(t_k)$, $S_1(t_k)$, and $S_2(t_k)$ are, respectively, the survival functions for the placebo, treatment 1, and treatment 2 groups by the end of time period k and $t_0 = 0$ is the baseline.

To determine the most cost-efficient number of time periods, we normalize the optimality criterion $var(\hat{\beta}_1)$ or $var(\hat{\beta}_2)$ by multiplying it by the costs C. As such, the optimality criterion is penalized by the amount of costs that follow from the number of time periods. A design with a large number of time periods produces a smaller $var(\hat{\beta}_1)$ or $var(\hat{\beta}_2)$ than a design with a small number of time periods, but it also has higher costs, and thus, we should impose a larger penalty. By using the normalized variance, a fair comparison on the number of time periods can be made. The optimal number of time periods is the one for which the normalized variance is minimal. All other designs with different numbers of time periods can then be evaluated relative to the design with the optimal number of time periods using the efficiency measure.

Table I shows the optimal number of time periods for multiple-objective optimal designs for which $var(\hat{\beta}_1)$ is the main objective and an efficiency of 0.9 is achieved on this objective. The optimal number of time periods is calculated for $\omega = 0.25$, 0.5, and 0.75; $\tau = 0.5$; 1 and 2; and for six combinations of the costs at the subject level. The costs at the measurement level are standardized to $c_m = 1$. Table II gives results when $var(\hat{\beta}_2)$ is the main objective. We first discuss results for cost function 1. Both tables show that the optimal number of time periods increases as the shape parameter τ increases. So, subjects should be measured more often and hence be followed for a longer amount of time, if the probability of event occurrence is largest at the end of the trial. The optimal number of time periods decreases with increasing ω. Subjects may be measured less often if the proportion of event occurrence increases. The optimal number of time periods also increases with the subject level costs c_0, c_1, and c_2, so more measurements should be taken if it becomes more expensive to include a subject in the trial. The optimal number of time periods is higher when cost function 2 is used, and in most cases, it is equal to 12. It follows that the optimal number of time periods required to follow subjects increases when subjects

Table I. Optimal number of time periods for multiple-objective optimal designs when $var(\hat{\beta}_1)$ is the primary objective and for cost function 1. Numbers in parentheses are the optimal number of time periods for cost function 2.

			$\omega = 0.25$			$\omega = 0.5$			$\omega = 0.75$		
c_0	c_1	c_2	$\tau = 0.5$	$\tau = 1$	$\tau = 2$	$\tau = 0.5$	$\tau = 1$	$\tau = 2$	$\tau = 0.5$	$\tau = 1$	$\tau = 2$
5	5	5	5(7)	12(12)	12(12)	4(10)	11(12)	12(12)	4(12)	7(12)	12(12)
5	10	15	7(11)	12(12)	12(12)	6(12)	12(12)	12(12)	5(12)	9(12)	12(12)
10	10	10	9(12)	12(12)	12(12)	7(12)	12(12)	12(12)	6(12)	9(12)	12(12)
10	20	30	12(12)	12(12)	12(12)	10(12)	12(12)	12(12)	7(12)	11(12)	12(12)
20	20	20	12(12)	12(12)	12(12)	11(12)	12(12)	12(12)	8(12)	12(12)	12(12)
20	40	60	12(12)	12(12)	12(12)	12(12)	12(12)	12(12)	10(12)	12(12)	12(12)

Statistics in Medicine

Table II. Optimal number of time periods for multiple-objective optimal designs when $var(\hat{\beta}_2)$ is the primary objective and for cost function 1. Numbers in parentheses are the optimal number of time periods for cost function 2.

			$\omega = 0.25$			$\omega = 0.5$			$\omega = 0.75$		
c_0	c_1	c_2	$\tau = 0.5$	$\tau = 1$	$\tau = 2$	$\tau = 0.5$	$\tau = 1$	$\tau = 2$	$\tau = 0.5$	$\tau = 1$	$\tau = 2$
5	5	5	5(7)	12(12)	12(12)	4(9)	9(10)	12(12)	3(10)	6(12)	10(12)
5	10	15	8(12)	12(12)	12(12)	6(12)	11(12)	12(12)	5(12)	7(12)	10(12)
10	10	10	8(12)	12(12)	12(12)	6(12)	12(12)	12(12)	5(12)	8(12)	11(12)
10	20	30	12(12)	12(12)	12(12)	9(12)	12(12)	12(12)	7(12)	9(12)	11(12)
20	20	20	12(12)	12(12)	12(12)	10(12)	12(12)	12(12)	7(12)	9(12)	12(12)
20	40	60	12(12)	12(12)	12(12)	12(12)	12(12)	12(12)	9(12)	11(12)	12(12)

leave the trial after event occurrence. It is therefore important to carefully decide whether subjects should be followed after event occurrence because this can appreciably impact the optimal number of time periods.

6. An example: risperidone maintenance treatment in schizophrenia

Prevention of relapse is a crucial task in the maintenance treatment of schizophrenia. For a long time, it has been believed that doses of antipsychotic drugs can be reduced during the maintenance period, but in recent years, it has been suggested that the initial dose of antipsychotics should be maintained in long-term treatment regimes. Wang *et al*. [31] performed a randomized controlled trial to compare the effect of three different dose regimes. In the 4-week group, the initial optimal therapeutic dose of risperidone was given for 4 weeks and then followed by a 50% reduction for the remainder of the trial. In the 26-week group, a 50% reduction was performed at 26 weeks and maintained until the end of the trial. In the no-dose-reduction group, the initial optimal therapeutic dose was given throughout the trial. The authors performed a discrete-time survival analysis using a logistic regression model. For the purpose of analysis, the observations were grouped into 100-day periods; all subjects were observed until 500 days because the probability of event occurrence was negligible at longer study duration.

The baseline logit hazard probabilities in the no-dose-reduction group (i.e., the baseline group) were estimated to be $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) = (-3.654, -3.706, -3.972, -4.363, -5.018)$. The effects of the 4-week group and 26-week group on the logit scale were estimated to be 1.219 and 0.822, respectively. In all three groups, the risk of event occurrence decreases across time. Lowest risk of event occurrence is observed in the no-dose-reduction group, highest risk in the 4-week group. At 500 days, the amount of relapse was 18%, 15%, and 8% in the 4-week group, the 26-week group, and the no-dose-reduction group, respectively. The three logit hazard probability functions are parallel, which implies the effect of treatment does not vary across time periods. That is, these functions are estimated on the basis of a proportional odds model. A non-proportional odds model was also fitted to the data but showed a non-significant decrease in the -2 log-likelihood value as compared with the proportional odds model and is not further considered in the derivation of the optimal design.

Sample sizes at baseline were 125 in the 4-week group, 120 in the 26-week group, and 129 in the no-dose-reduction group. In other words, the design was highly balanced with respect to the sizes of the treatment groups. We use optimal design methodology to study if a balanced design is indeed the best choice for a trial like this. Suppose the trial has two objectives: the comparison of the no-dose-reduction group with the 4-week group and the comparison of the no-dose-reduction group with the 26-week group. Using multiple-objective optimal designs, we derive optimal allocations. We use the estimated model parameter values to derive the optimal designs.

The optimal design weights rounded to two decimals are $\pi_0 = 0.57$, $\pi_1 = 0.33$, and $\pi_2 = 0.10$ in case an efficiency of 0.9 is to be achieved for the objective $var(\hat{\beta}_1)$ and $\pi_0 = 0.54$, $\pi_1 = 0.10$, and $\pi_2 = 0.36$ in case an efficiency of 0.9 is to be achieved for the objective $var(\hat{\beta}_2)$. These values hold for $p = 2$, 3, 4, and 5, so changing the number of time periods during the course of the trial does not affect the optimal design weights. Note that the design is unbalanced: about half of the subjects are allocated to the no-dose-reduction group and about a third to the dose-reduction group whose comparison with the no-dose-reduction group is of primary interest. For each number of time periods and for both objectives, the efficiency of the equally balanced design is between 0.69 and 0.72. This shows balanced designs can perform sub-optimally.

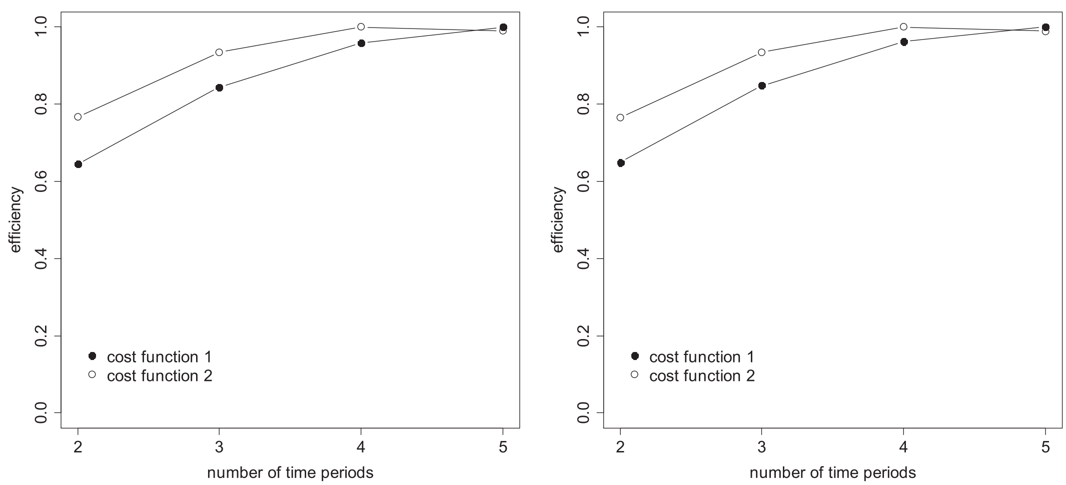

Figure 4. Efficiencies of $var(\hat{\beta}_1)$ (left) and $var(\hat{\beta}_2)$ (right) as functions of the number of time periods in the risperidone maintenance treatment example.

The optimal number of time periods depends on the costs per subject for each of the three treatment groups and the costs to take a measurement. Such costs are not given in [31], and we postulate that the subject level costs for the no-dose-reduction group are higher than those for the two dose-reduction groups, and the costs for the 26-week group are higher than those of the 4-week group. Accordingly, one may choose as an example values $c_0 = 30$, $c_1 = 20$, and $c_2 = 10$ with $c_m = 1$. Figure 4 shows that for both objectives, the optimal number of time periods is $p = 5$ when the cost function 1 is used, and for the cost function 2, the best choice for the number of time periods is $p = 4$ periods. For both cost functions, a loss of efficiency is observed when $p = 3$ and more so when the number of time periods is $p = 2$.

7. Conclusions

Our results showed that equal treatment allocation is not always the best choice when we compare several treated groups with the placebo group. The weight assigned to a treatment group depends on the importance of the comparison of the treatment relative with the placebo, the cost function, and the values of the parameters in the Weibull model. It is therefore important that the ordering of the objectives is carefully justified in the design phase of a trial. Extra care should be exercised when the objectives are competitive because if the efficiency requirements in the constraints are too demanding; we may not be able to find a design that meets the multiple criteria. If testing effects of the various treatment effects is the primary goal, care should be taken to have a large enough sample size to have reasonable statistical power for the tests. Power analysis for trials with discrete-time survival endpoints is discussed in Jóźwiak and Moerbeek [32].

The optimal design weights may depend on the parameters ω and τ for the two-parameter Weibull distribution that we have assumed, and so our optimal designs are locally optimal. It is therefore important that the values of these parameters are carefully specified *a priori* because incorrect values can result in a substantial loss of efficiency. Prior estimates may be obtained from similar studies or subject matter knowledge. If an experimenter cannot provide a single point estimate for each parameter but has a prior distribution for the model parameters, a Bayesian optimal design approach may be used [33]. Such a design strategy can provide some protection against misspecification of the nominal values of the parameters. In practice, robustness properties of an optimal design to other misspecifications such as in the link function or survival model (log-logistic or log-normal, etc.) should be studied before the design is implemented.

For our setup, we also observed that increasing the number of time periods results in a more efficient design but also comes with higher trial cost. A cost function was therefore incorporated in the optimization problem to ascertain the optimal number of time periods. However, actual cost functions may not be known with certainty, and therefore, misspecification in the cost function can be a concern as well. In summary, before a design is implemented, it is important to routinely ascertain robustness properties of the design to all possible misspecifications in the model assumptions.

We do not carry out such an investigation here for space consideration, but it can be readily carried out using our proposed methodology, which is not limited by our illustrative model and the application used in this paper.

Acknowledgements

This research was supported by a grant from the Netherlands Organisation for Scientific Research (NWO) awarded to M. Moerbeek, grant number 452-08-004. The research of W. K. Wong reported in this paper was partially supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R01GM107639. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank the editor and two anonymous reviewers for their constructive comments, which have helped us improve the quality of the paper.

References

1. Dumville JC, Hahn S, Miles JNV, Torgerson DJ. The use of unequal randomisation ratios in clinical trials: a review. *Contemporary Clinical Trials* 2006; **27**(1):1–12.
2. Jadad AR, Enkin MW. *Randomized Controlled Trials: Questions, Answers and Musings*. Blackwell: Malden, MA, 2007.
3. Torgerson DJ, Torgerson CJ. *Designing Randomised Trials in Health, Education and The Social Sciences: An Introduction.* Palgrave Macmillan: Basingstoke, 2008.
4. Allison DB, Allison RL, Faith MS, Paultre F, Pi-Sunyer FX. Power and money: designing statistically powerful studies while minimizing financial costs. *Psychological Methods* 1997; **2**(1):20–33.
5. Dette H. On robust and efficient designs for risk estimation in epidemiological studies. *Scandinavian Journal of Statistics* 2004; **31**(3):319–331.
6. Hsu LM. Unbalanced designs to maximize statistical power in psychotherapy efficacy studies. *Psychotherapy Research* 1994; **4**(2):95–106.
7. Schouten HJA. Sample size formula with a continuous outcome for unequal group sizes and unequal variances. *Statistics in Medicine* 1999; **18**(1):87–91.
8. Liu XF. Statistical power and optimum sample allocation ratio for treatment and control having unequal costs per unit of randomization. *Journal of Educational and Behavioral Statistics* 2003; **28**(3):231–248.
9. Moerbeek M, Wong WK. Sample size formulae for trials comparing group and individual treatments in a multilevel model. *Statistics in Medicine* 2008; **27**(15):2850–2864.
10. Zhu W, Wong WK. Optimal treatment allocation in comparative biomedical studies. *Statistics in Medicine* 2000; **19**(5):639–648.
11. Zhu W, Wong WK. Optimum treatment allocation for dual-objective clinical trials with binary outcomes. *Communications in Statistics, Theory and Methods* 2000; **29**(5–6):957–974.
12. Cox DR. Regression models and life tables. *Journal of the Royal Statistical Society Series B* 1972; **34**(2):187–220.
13. Chow SC, Shao J, Wang H. *Sample Size Calculations in Clinical Research* 2nd edition. Chapman and Hall: New York, 2008.
14. Collett D. *Modelling Survival Data in Medical Research* 2nd edition. Chapman & Hall/CRC: London, UK, 2003.
15. Hosmer DW, Lemeshow S, May S. *Applied Survival Analysis. Regression Modeling of Time-To-Event Data* 2nd edition. Wiley: Hoboken, New Jersey, 2008.
16. Julious SA. *Sample Sizes for Clinical Trials*. Chapman and Hall/CRC Press: Boca Raton, 2010.
17. Machin D, Cheung YB, Parmar M. *Survival Analysis: A Practical Approach* 2nd edition. Wiley: Chichester, 2006.
18. Kalish LA, Harrington DP. Efficiency of balanced treatment allocation for survival analysis. *Biometrics* 1988; **44**(3): 815–821.
19. Sverdlov O, Ryeznik Y, Wong WK. Doubly-adaptive biased coin designs for balancing competing objectives in time-to-event trials. *Statistics and Its Interface* 2012; **5**(4):401–414.
20. Singer JD, Willett JB. *Applied Longitudinal Data Analysis. Modeling Change and Event Occurrence*. Oxford University Press: Oxford, 2003.
21. Singer JD, Willett JB. It's about time: using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational and Behavioral Statistics* 1993; **18**(2):155–195.
22. Joźẃiak K, Moerbeek M. Cost-effective designs for trials with discrete-time survival endpoints. *Computational Statistics and Data Analysis* 2012; **56**(6):2086–2096.
23. Moerbeek M. Sample size issues for cluster randomized trials with discrete-time survival endpoints. *Methodology* 2012; **8**(4):146–158.
24. Moerbeek M, Joźẃiak K. Optimal designs for event history analysis. *Zeitschrift für Psychologie* 2013; **221**(4):160–173.
25. Joźẃiak K, Moerbeek M. Optimal treatment allocation and study duration for trials with discrete-time survival endpoints. *Journal of Statistical Planning and Inference* 2013; **143**(5):971–982.
26. Sverdlov O, Rosenberger WF. On recent advances in optimal allocation designs in clinical trials. *Journal of Statistical Theory and Practice* 2013; **7**(4):753–773.
27. McCullagh P, Nelder JA. *Generalized Linear Models*. Chapman and Hall: London, 1989.
28. Cook RD, Wong WK. On the equivalence of constrained and compound optimal designs. *Journal of the American Statistical Association* 1994; **89**(426):687–692.

29. Varadhan R. *Alabama (R Package Version 2011.9-1)*, 2012. http://cran.r-project.org/web/packages/alabama/index.html [Accessed on 18 February 2013].
30. Fleiss JL. *The Design and Analysis of Clinical Experiments*. Wiley: New York, 1986.
31. Wang CY, Xiang YT, Cai ZJ, Weng YZ, Bo QJ, Zhao JP, Liu TQ, Wang QH, Weng SM, Zhang HY, Chen DF, Tang WK. Risperidone maintenance treatment in schizophrenia: a randomized, controlled trial. *American Journal Of Psychiatry* 2010; **167**(6):676–685.
32. Joźẃiak K, Moerbeek M. Power analysis for trials with discrete-time survival endpoints. *Journal of Educational and Behavioral Statistics* 2012; **37**(5):630–654.
33. Woods DC, Lewis SM, Eccleston JA, Russell KG. Designs for generalized linear models with several variables and model uncertainty. *Technometrics* 2006; **48**(2):284–292.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site.

