



## Optimal designs for response functions with a downturn

Seung Won Hyun<sup>a,\*</sup>, Min Yang<sup>b</sup>, Nancy Flournoy<sup>b</sup>

<sup>a</sup> Department of Statistics, North Dakota State University, 203E1A Waldron Hall, Fargo, ND 58102, USA

<sup>b</sup> Department of Statistics, University of Missouri, Columbia, MO 65211, USA

### ARTICLE INFO

#### Article history:

Received 5 March 2009

Received in revised form

30 June 2010

Accepted 5 July 2010

Available online 16 July 2010

#### Keywords:

Toxicological assays

Peak dose

$EC_{50}$

A-optimality

D-optimality

$D_A$ -optimality

Successive mean differences

Optimal weights

Dose–response

Experimental design

Nonlinear response function

Laboratory studies

Endocrine-disrupting chemicals

### ABSTRACT

In many toxicological assays, interactions between primary and secondary effects may cause a downturn in mean responses at high doses. In this situation, the typical monotonicity assumption is invalid and may be quite misleading. Prior literature addresses the analysis of response functions with a downturn, but so far as we know, this paper initiates the study of experimental design for this situation. A growth model is combined with a death model to allow for the downturn in mean doses. Several different objective functions are studied. When the number of treatments equals the number of parameters, Fisher information is found to be independent of the model of the treatment means and on the magnitudes of the treatments. In general, A- and  $D_A$ -optimal weights for estimating adjacent mean differences are found analytically for a simple model and numerically for a biologically motivated model. Results on c-optimality are also obtained for estimating the peak dose and the  $EC_{50}$  (the treatment with response half way between the control and the peak response on the increasing portion of the response function). Finally, when interest lies only in the increasing portion of the response function, we propose composite D-optimal designs.

© 2010 Elsevier B.V. All rights reserved.

### 1. Introduction

A monotone response often is expected as dose increases in toxicological assays. However, strict monotonicity is not always realistic. Interactions between primary and secondary effects may cause a downturn at high doses. In this paper, we obtain optimal designs for such response functions.

Fan and Chaloner (2001), Rabie and Flournoy (2004), Han and Chaloner (2004), Fedorov and Wu (2007a, 2007b), Dragalin et al. (2008) and others have studied optimal designs for experiments in which tertiary or bivariate binary outcomes for toxicity and efficacy were observed for each subject and used to provide information on the probability of efficacy without toxicity. Typically, this probability increases and then decreases.

In this paper, we study designs for experiments in which one continuous outcome is observed for each subject and used to provide information on a response function with a downturn. Models for this later situation have been described by Margolin et al. (1981), Welshons et al. (2003) and others.

\* Corresponding author. Tel.: +1 701 231 8178.

E-mail addresses: [Seung.W.Hyun@ndsu.edu](mailto:Seung.W.Hyun@ndsu.edu), [swhr7@mizzou.edu](mailto:swhr7@mizzou.edu) (S.W. Hyun), [YangMi@Missouri.edu](mailto:YangMi@Missouri.edu) (M. Yang), [FlournoyN@Missouri.edu](mailto:FlournoyN@Missouri.edu) (N. Flournoy).

Margolin et al. (1981) studied the analysis of data from the Ames *Salmonella*/microsome test. Each microbe is placed on a plate and the number of revertant colonies is observed from each plate after chemical dose. One outcome, the number of revertants, was observed from each plate and used to estimate the dose response function. They found that the number of revertants per plate has the nonmonotonic relationship with the test chemical dose. A downturn is exhibited that is commonly ascribed to microbial toxicity. A biologically motivated family of mathematical models was constructed to relate plate counts, on average, to doses.

Welshons et al. (2003) presented an overview of the mechanisms of hormone action that provides the basis for understanding how endocrine-disrupting chemicals with estrogenic activity (EEDCs) can be biologically active at low, environmentally relevant doses. It is traditional to linearly extrapolate responses from high test doses to much lower environmentally relevant doses. Their data calls this practice into question. They point out that this traditional testing method has led to misinterpretations because receptor-mediated responses can first increase and then decrease as dose increases, contradicting traditional model assumptions. Responses to hormones, including estrogens, decrease when receptor occupancy becomes saturated, causing the response function to increase and then decrease with dose. To exemplify this inverted-U pattern, endocrine doses of MCF-7 (human breast cancer) cells were obtained over a wide range of doses (see Fig. 1).

These two experimental situations are similar in that they both exhibit a downturn in the response function. However, Margolin et al. deal with a discrete outcome, i.e., the number of revertants, while Welshons et al. deal with a continuous outcome, i.e., the percentage of cell growth. We adapt Margolin et al.'s model for a continuous response.

There are several interesting features to study when the response functions have a downturn. We obtain optimal designs for estimating three different features in this paper. The first is successive mean differences, motivated by Brez and Hothorn (2003). The other two are the peak dose and the  $EC_{50}$ . The peak dose is the dose producing the maximum response and the  $EC_{50}$  is the dose producing a response that is half way between the minimum response and the maximum response on the increasing part of the response function.

In this paper, doses are taken to be fixed for estimating adjacent mean differences and the optimal allocation of experimental units to these doses are obtained. Optimal designs for estimating adjacent mean differences are to minimize the variances of these differences. For example, optimal designs for estimating  $\mu_1 - \mu_2$  is to minimize  $\text{Var}(\mu_1 - \mu_2)$ . In this case, the variance is minimized when  $\mu_2 = \mu_1 + \delta$ ,  $\delta \rightarrow 0$ . Thus, the optimal design specifies design points as close as possible to the dose producing  $\mu_1$ . Placing all observations at one point provides no information about the mean function at other points. So, this design is useless even though it minimizes the variance. To avoid this impractical design problem, we adopt this restriction that doses are fixed. Although we do not have a rigorous proof, our intuition tells us that is true.

Optimal weights, i.e., the proportional allocation of experimental units to doses, are obtained (1) for estimating adjacent mean differences under  $D_A$ - and  $A$ -optimality criteria (see Atkinson and Donev, 1992), and (2) for estimating the peak dose and the  $EC_{50}$  under  $c$ -optimality criterion, i.e., to minimize the variance of estimating the combination of the parameters. To estimate all the parameters in the model, a composite  $D$ -optimal design is proposed which specifies the design points and weights simultaneously.

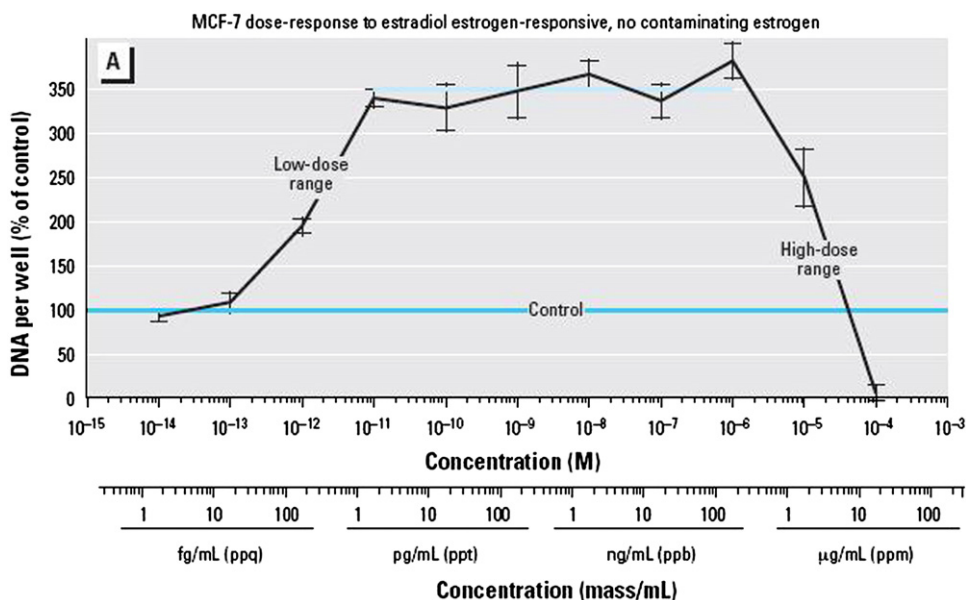


Fig. 1. The source is Welshons et al. (2003).

In Section 2, we introduce notation and establish a frame work for more complex situations,  $D_A$ - and A-optimal weights for estimating adjacent mean differences under a very simple model are obtained. In Section 3, optimal weights for a biologically motivated model are derived. First, optimal designs are obtained for an interesting special case. Next, optimal weights are obtained for three interesting features: adjacent mean differences, the peak dose, and the  $EC_{50}$ . A composite D-optimal design is proposed in Section 4. Finally, this paper concludes with a brief summary.

## 2. Designs for a simple model

Initially suppose that

$$y_{ij} = \mu_i + \varepsilon_{ij}, \quad (1)$$

where the  $\varepsilon_{ij}$  are independent and  $N(0, \sigma^2)$ ,  $j=1, 2, \dots, n_i$ ,  $i=0, 1, \dots, K$ .  $n_1 + n_2 + \dots + n_K = n$ ;  $\mu_i$  is the mean response at design point  $x_i$ ;  $\mu_i < \infty$ ,  $\forall i$  and  $\sigma^2 < \infty$  is assumed unknown. In general, the  $\{x_i\}$  are control variables. In keeping with the motivating biology, we call the  $\{x_i\}$  doses. The likelihood function is

$$\mathcal{L} = \prod_{i=1}^K \prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y_{ij}-\mu_i)^2\right\}.$$

Let  $\xi = \{(x_i, w_i)\}_1^K$  denote a design, where  $x_i$  is the  $i$ th dose and  $w_i = n_i/n$  is the associated design weight (i.e., the proportion of subjects allocated to dose  $x_i$ ).

Then the normalized (i.e., per-subject) Fisher information matrix for maximum likelihood estimates  $\hat{\mu}_i$  and  $\hat{\sigma}^2$  of  $\mu_i$  and  $\sigma^2$ ,  $i=1, \dots, K$  is

$$M_{\mu, \sigma^2} = \frac{1}{\sigma^2} \text{diag}\left(w_1, w_2, \dots, w_K, \frac{1}{2\sigma^2}\right),$$

where  $\sigma^2$  is an unknown constant. Because  $\{\hat{\mu}_i\}_1^K$  are independent of  $\hat{\sigma}^2$  and we are not interested in  $\sigma^2$ , it is sufficient to consider the reduced information for  $\{\mu_i\}_1^K$ :

$$M_{\mu} = \frac{1}{\sigma^2} \text{diag}(w_1, w_2, \dots, w_K). \quad (2)$$

### 2.1. $D_A$ -optimal design for adjacent mean differences

Brez and Hothorn (2003) study response functions with a downturn by estimating the magnitude of each mean difference. The information matrix for  $\{\hat{\mu}_{i+1} - \hat{\mu}_i\}_1^{K-1}$  is  $[AM_{\mu}^{-1}A^T]^{-1}$ , where  $A$  is the  $(K-1) \times K$  matrix:

$$A = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}.$$

The  $D_A$ -optimal design for estimating  $\{\mu_{i+1} - \mu_i\}_1^{K-1}$  given  $K$  doses at  $\{x_1, \dots, x_K\}$  under model (1) is

$$\xi^* = \text{argmindet}[AM_{\mu}^{-1}A^T] = \text{argmin}[w_1 w_2 \cdots w_K]^{-1},$$

which is

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_K \\ \frac{1}{K} & \frac{1}{K} & \cdots & \frac{1}{K} \end{pmatrix}.$$

### 2.2. A-optimal design for adjacent mean differences

A-optimality minimizes the sum of variances for the estimators of interest. A-optimality for adjacent mean differences minimizes

$$\begin{aligned} \Psi &= \text{tr}[AM_{\mu}^{-1}A^T] = \sum_{i=1}^{K-1} \text{Var}(\hat{\mu}_{i+1} - \hat{\mu}_i) = \sum_{i=1}^{K-1} \sigma^2 \left(\frac{1}{w_i} + \frac{1}{w_{i+1}}\right) \\ &= \sigma^2 \left(\frac{1}{w_1} + \frac{2}{w_2} + \frac{2}{w_3} + \cdots + \frac{2}{w_{K-1}} + \frac{1}{w_K}\right), \end{aligned} \quad (3)$$

with the restriction that  $\sum_{i=1}^K w_i = 1$ .

**Theorem 1.** The A-optimal design for estimating  $\{\mu_{i+1}-\mu_i\}_1^{K-1}$  given  $K$  doses  $\{x_1, \dots, x_K\}$  under model (1) is

$$\xi^* = \begin{pmatrix} x_1 & x_2 & \cdots & x_K \\ w_1 & w_2 & \cdots & w_K \end{pmatrix},$$

$$w_2 = w_3 = \cdots = w_{K-1} = \frac{\sqrt{2}}{2-2\sqrt{2}+K\sqrt{2}},$$

$$w_1 = w_K = \frac{1}{2-2\sqrt{2}+K\sqrt{2}}.$$

**Proof.** Let  $w_1+w_K=u$ . Then  $w_2+\cdots+w_{K-1}=1-u$ . Note that (3) is minimized when  $w_1 = w_K = u/2$  and  $w_i = (1-u)/(K-2)$ ,  $i=2, \dots, K-1$ . So the problem reduces to finding

$$\operatorname{argmin} \left( \frac{4}{u} + \frac{2(K-2)^2}{1-u} \right) \tag{4}$$

which is

$$w_1 = w_K = \frac{1}{2-2\sqrt{2}+K\sqrt{2}}.$$

Now one can see that (3) is minimized when (i)  $n_1=n_K$ , (ii)  $n_2=n_3=\cdots=n_{K-1}$  and (iii)  $\sum_{i=1}^K n_i=n$ . The number of replications obtained from Theorem 1 is not typically an integer because  $n_i=w_i n$ . Thus, the number of replications are obtained as  $n_i = \lfloor w_i n \rfloor$ , where  $\lfloor w_i n \rfloor$  denotes the maximum integer that does not exceed  $\lfloor w_i n \rfloor$ . This might lead  $\sum_{i=1}^K n_i \neq n$ . For fixed  $n$ , the number of replications is adjusted as  $n_1=n_K$  or  $n_K \pm 1$  and  $n_i=n_j$  or  $n_j \pm 1$  where  $i, j \in (2, 3, \dots, K-1)$ .  $\square$

### 3. Designs for a biologically motivated model

In our motivating applications, responses at different doses are structurally dependent. To develop a biologically motivated relationship between the mean responses at different doses, let  $F(\theta, x)$  and  $G(\theta, x)$  be strictly monotone increasing continuous and nonnegative functions of dose ranging from zero to one. Define  $\Theta = (\theta_1, \theta_2)$ , where  $\theta_1 = (\theta_{11}, \theta_{12}, \dots, \theta_{1p_1})$  and  $\theta_2 = (\theta_{21}, \theta_{22}, \dots, \theta_{2p_2})$  and let  $\{x_1 < x_2 < \cdots < x_K\}$  be the set of design doses. At the doses, define  $F_i = F(\theta_1, x_i)$ ,  $G_i = G(\theta_2, x_i)$ . Now suppose

$$y_{ij} \sim N(\mu_i(\Theta), \Sigma), \tag{5}$$

where

$$\mu_i(\Theta) = F_i \bar{G}_i, \quad \bar{G}_i = 1 - G_i, \quad \Sigma = \sigma^2 I_{K \times K}.$$

Assume  $\sigma^2$  is unknown.

#### 3.1. Fisher information

Because all  $\mu_i$  are functions of the same parameters, the response densities at different doses are dependent on each other. Set  $\mu(\Theta) = (\mu_1(\Theta), \mu_2(\Theta), \dots, \mu_K(\Theta))$ . Since there are  $n_i$  observations at each  $x_i$ , the normalized Fisher information matrix for  $\hat{\Theta}$  over the entire experiment can be written as

$$M(\xi; \Theta) = \mu'(\Theta)^T M_\mu \mu'(\Theta),$$

where

$$\mu'(\Theta) = \begin{pmatrix} \frac{\partial \mu(\Theta)^T}{\partial \theta_{11}} & \frac{\partial \mu(\Theta)^T}{\partial \theta_{12}} & \cdots & \frac{\partial \mu(\Theta)^T}{\partial \theta_{1p_1}} & \frac{\partial \mu(\Theta)^T}{\partial \theta_{21}} & \frac{\partial \mu(\Theta)^T}{\partial \theta_{22}} & \cdots & \frac{\partial \mu(\Theta)^T}{\partial \theta_{2p_2}} \end{pmatrix}.$$

By Taylor series expansion, the covariance matrix for estimating  $\mu(\Theta)$  is approximately  $\Sigma_{\mu(\Theta)} = \mu'(\Theta) M^{-1}(\xi; \Theta) [\mu'(\Theta)]^T$ , where  $M^{-1}(\xi; \Theta)$  denotes the generalized inverse of  $M(\xi; \Theta)$ , and the normalized Fisher information matrix for  $\hat{\mu}(\Theta)$  is

$$M(\xi; \mu(\Theta)) = [\mu'(\Theta) M^{-1}(\xi; \Theta) [\mu'(\Theta)]^T]^{-1}.$$

**Example.** As an illustration, we adapt Margolin et al.'s (1981) model. Assuming model (5), set

$$F_i = 1 - \exp[-(\alpha + \beta x_i)], \quad \bar{G}_i = \exp[-\gamma x_i]$$

and so the mean response function is

$$\mu_i = F_i \bar{G}_i = \{1 - \exp[-(\alpha + \beta x_i)]\} \exp[-\gamma x_i], \tag{6}$$

where  $\alpha > 0, \beta \geq 0, \gamma \geq 0$ . Since the mean response function is parameterized by  $\Theta, \Theta = (\theta_1, \theta_2)$ , where  $\theta_1 = (\alpha, \beta)$  and  $\theta_2 = (\gamma)$ , the Fisher information matrix is much more complicated than for the simple model (1).

Let  $\varphi_i = e^{-(\alpha + \beta x_i + \gamma x_i)}$ . Then the normalized information matrix for  $\hat{\Theta} = \{\hat{\alpha}, \hat{\beta}, \hat{\gamma}\}$  can be written as

$$M(\xi; \Theta) = \frac{1}{\sigma^2} \sum_{i=1}^K w_i \begin{pmatrix} \varphi_i^2 & x_i \varphi_i^2 & -x_i \mu_i \varphi_i \\ x_i \varphi_i^2 & x_i^2 \varphi_i^2 & -x_i^2 \mu_i \varphi_i \\ -x_i \mu_i \varphi_i & -x_i^2 \mu_i \varphi_i & [x_i \mu_i]^2 \end{pmatrix},$$

which yields the information matrix for  $\mu(\hat{\Theta})$ :

$$M(\xi; \mu(\Theta)) = [\mu'(\Theta)M^{-1}(\xi; \Theta)[\mu'(\Theta)]^T]^{-1},$$

where

$$\mu'(\Theta) = \begin{pmatrix} \varphi_1 & x_1 \varphi_1 & -x_1 \mu_1 \\ \varphi_2 & x_2 \varphi_2 & -x_2 \mu_2 \\ \vdots & \vdots & \vdots \\ \varphi_K & x_K \varphi_K & -x_K \mu_K \end{pmatrix}.$$

### 3.2. Optimal designs when $K=p_1+p_2$

This section provides a very general optimality result for the case when the number of doses is the same as the number of parameters.

**Theorem 2.** For any optimality criterion that is a concave function of the elements on  $M(\xi; \mu(\Theta))$ , assuming  $M(\xi; \mu(\Theta))$  is of full rank and that the number of doses is the same as the number of parameters in the model of  $\mu(\Theta)$ , the optimal design is independent of the magnitude of the doses and of the parameters.

**Proof.** There are  $K$  different doses and  $p_1+p_2$  parameters in model (5). Set  $K=p_1+p_2$ . The Fisher information for  $\hat{\Theta}$  can be expressed as

$$M(\xi; \Theta) = \frac{1}{\sigma^2} \sum_{i=1}^K w_i \begin{pmatrix} \frac{\partial \mu(\Theta)}{\partial \theta_1} & \frac{\partial \mu(\Theta)}{\partial \theta_2} & \dots & \frac{\partial \mu(\Theta)}{\partial \theta_K} \end{pmatrix}^T \begin{pmatrix} \frac{\partial \mu(\Theta)}{\partial \theta_1} & \frac{\partial \mu(\Theta)}{\partial \theta_2} & \dots & \frac{\partial \mu(\Theta)}{\partial \theta_K} \end{pmatrix} = \mu'(\Theta)^T M_\mu \mu'(\Theta), \tag{7}$$

where  $M_\mu$  is given by (2). The covariance matrix for  $\mu(\hat{\Theta})$  is

$$\begin{aligned} \mu'(\Theta)M(\xi; \Theta)^{-1}\mu'(\Theta)^T &= \mu'(\Theta)[\mu'(\Theta)^T M_\mu \mu'(\Theta)]^{-1}\mu'(\Theta)^T \\ &= \mu'(\Theta)[\mu'(\Theta)]^{-1}M_\mu^{-1}[\mu'(\Theta)^T]^{-1}\mu'(\Theta)^T = M_\mu^{-1}. \quad \square \end{aligned}$$

**Remark.** The covariance matrix in this case is the same as for independent treatments. The optimal weights do not depend on the model of the treatment means nor on the magnitudes of the doses.

### 3.3. Designs for estimating adjacent mean differences

This section provides an approach for deriving the optimal weights for given doses.

#### 3.3.1. Under $D_A$ -optimality

The  $D_A$ -optimality criterion is the determinant of the covariance matrix of  $\{\hat{\mu}_{i+1}(\Theta) - \hat{\mu}_i(\Theta)\}_1^{K-1}$ :

$$\Psi = \det[AM^{-1}(\xi; \mu(\Theta))A^T] = \det[BM^{-1}(\xi; \Theta)B^T], \tag{8}$$

where  $B = A\mu'(\Theta)$ . Let  $w = (w_1, \dots, w_{K-1})$  because  $w_K = 1 - \sum_{i=1}^{K-1} w_i$ . The minimum is obtained among all designs where the adjacent mean differences are estimable. The generalized inverse prevents the singularity of the Fisher information matrix for  $\{\hat{\mu}_{i+1}(\Theta) - \hat{\mu}_i(\Theta)\}_1^{K-1}$ . The singularity implies  $D_A$ -admissible designs for estimating the adjacent mean differences are restricted to  $p_1+p_2 \leq K \leq p_1+p_2+1$ . However, this restriction does not apply to  $A$ -optimal designs which are the main focus of this paper.

**Theorem 3.** The nonnegative solutions of  $(\partial/\partial w)[\det[BM^{-1}(\xi; \Theta)B^T]] = 0$  are the  $D_A$ -optimal weights.

**Proof.** See A.1 in the appendix.

In contrast to assuming the simple model (1), it is hard to obtain an analytical solution for optimal weights assuming (5) and (6). However, the optimal weights usually can be found numerically using a Newton–Raphson algorithm. If no nonnegative solution is found, we recommend initializing the smallest negative weight as zero and searching again.

Fig. 2 shows model (5) with mean response function (6);  $\alpha = 0.11$ ,  $\beta = 1$  and  $\gamma = 2$ . Optimal weights for selected sets of design doses are shown in Table 1.

Note that the replications are distributed uniformly over three doses without regard to the value of the doses or the values of the parameters. When there are less than three design points,  $\Theta$  is not estimable, but  $\mu(\Theta)$  and the adjacent mean differences can be estimated through the sample means. This is equivalent to using the simple model, which has been addressed in Section 2. When there are more than four design points, the variance–covariance matrix of adjacent mean differences is singular regardless of the design. It is not meaningful to consider  $D_A$ -optimal designs here. However, we can consider the A-optimality criterion. Designs 1 and 2 have the same criterion value even though they have different doses. This illustrates Theorem 2. The designs with any three fixed doses have optimal weights that are uniformly put on all doses and have the same criterion disregarding the levels of doses. Designs 1 and 3 both have three design points, but they have different criterion values. To see why, define  $\mu(x)$  to be the mean response at dose  $x$ . Then Design 1 is for estimating  $\{\mu(1.00) - \mu(0.30), \mu(0.30) - \mu(0.00)\}$  and Design 3 is for estimating  $\{\mu(1.00) - \mu(0.70), \mu(0.70) - \mu(0.30), \mu(0.30) - \mu(0.00)\}$ . They are optimal for different sets of contrasts. The criterion values for these design differ because they have different interesting estimators.

In contrast, Table 2 shows  $D_A$ -optimal design and criterion for selected sets of design doses under simple model (1).

Under the simple model (1), Fisher information matrices for  $\{\hat{\mu}_{i+1} - \hat{\mu}_i\}_1^{K-1}$  with more than four points are non-singular. Since the mean function is not parameterized under the simple model, the Fisher information matrix is always full rank and  $D_A$ -optimal designs put equal weight on all design points. Note that there is no difference between the simple model and the biologically motivated model when the number of design points are three.

3.3.2. Under A-optimality

The criterion function for the A-optimal design is

$$\begin{aligned} \Psi &= \sum_{i=1}^{K-1} \text{Var}(\hat{\mu}_{i+1}(\Theta) - \hat{\mu}_i(\Theta)) \\ &= \text{tr}[A\mu'(\Theta)M(\xi; \Theta)^{-1}\mu'(\Theta)^T A^T] \\ &= [\text{tr}(M(\xi, \Theta)^{-1}V)], \end{aligned}$$

where  $V = \sum_{i=1}^{K-1} [B_i]^T B_i$ ,  $B_i = C_i \mu'(\Theta)$  and  $C_i = i_{th}$  row of the matrix A. A similar theorem to Theorem 4 under D-optimality is given by

**Theorem 4.** The nonnegative solutions of  $(\partial/\partial w)[\text{tr}(M(\xi; \Theta)^{-1}V)] = 0$  are the A-optimal design weights.

**Proof.** See A.2 in the appendix.

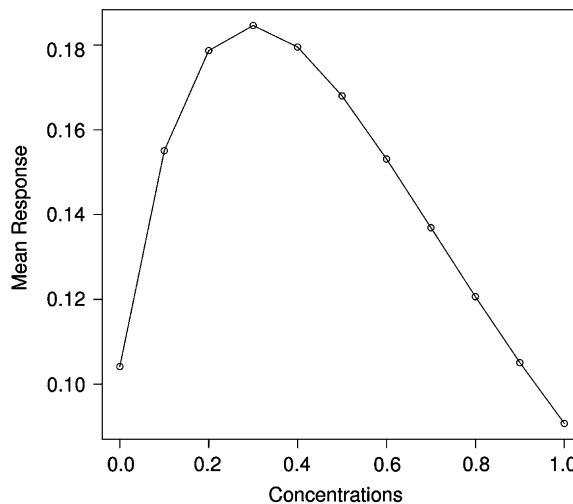


Fig. 2.  $\alpha = 0.11$ ,  $\beta = 1$  and  $\gamma = 2$ .

**Table 1**D<sub>A</sub>-optimal weights for estimating adjacent mean differences given selected sets of design doses for the biologically motivated model.

| Design | Design doses           | Optimal weights        | $\Psi$ |
|--------|------------------------|------------------------|--------|
| 1      | 0.00, 0.30, 1.00       | 0.333, 0.333, 0.333    | 27.054 |
| 2      | 0.00, 0.50, 1.00       | 0.333, 0.333, 0.333    | 27.054 |
| 3      | 0.00, 0.30, 0.70, 1.00 | 0.333, 0.333, 0, 0.333 | 0.463  |

**Table 2**D<sub>A</sub>-optimal weights for estimating adjacent mean differences given selected sets of design doses for the simple model.

| Design | Design doses                 | Optimal weights                   | $\Psi$  |
|--------|------------------------------|-----------------------------------|---------|
| 1      | 0.00, 0.30, 1.00             | 0.333, 0.333, 0.333               | 27.054  |
| 2      | 0.00, 0.50, 1.00             | 0.333, 0.333, 0.333               | 27.054  |
| 3      | 0.00, 0.30, 0.70, 1.00       | 0.250, 0.250, 0.250, 0.250        | 256.410 |
| 4      | 0.00, 0.25, 0.50, 0.75, 1.00 | 0.200, 0.200, 0.200, 0.200, 0.200 | 3125    |

**Table 3**

A-optimal weights for estimating adjacent mean differences given selected sets of design doses for the biologically motivated model.

| Design | Design doses                 | Optimal weights           | $\Psi$ |
|--------|------------------------------|---------------------------|--------|
| 1      | 0.00, 0.30, 1.00             | 0.292, 0.414, 0.292       | 11.680 |
| 2      | 0.00, 0.50, 1.00             | 0.292, 0.414, 0.292       | 11.680 |
| 3      | 0.00, 0.30, 0.70, 1.00       | 0.324, 0.400, 0, 0.275    | 9.594  |
| 4      | 0.00, 0.25, 0.50, 0.75, 1.00 | 0.356, 0.407, 0, 0, 0.236 | 8.123  |

**Table 4**

A-optimal weights for estimating adjacent mean differences given selected sets of design doses for the simple model.

| Design | Design doses                 | Optimal weights                   | $\Psi$  |
|--------|------------------------------|-----------------------------------|---------|
| 1      | 0.00, 0.30, 1.00             | 0.292, 0.414, 0.292               | 11.680  |
| 2      | 0.00, 0.50, 1.00             | 0.292, 0.414, 0.292               | 11.680  |
| 3      | 0.00, 0.30, 0.70, 1.00       | 0.207, 0.292, 0.292, 0.207        | 277.77  |
| 4      | 0.00, 0.25, 0.50, 0.75, 1.00 | 0.160, 0.226, 0.226, 0.226, 0.160 | 3448.27 |

There is no analytical solution that satisfies Theorem 4. But the solution can be found numerically by a Newton–Raphson algorithm. Optimal designs for given sets of doses under the biologically motivated model are shown in Table 3 with  $\alpha = 0.11$ ,  $\beta = 1$ ,  $\gamma = 2$ .

The A-optimal designs shown in Table 3 show how replications need to be distributed over given sets of doses to efficiently estimate adjacent mean differences. Designs 1 and 2 have different doses but have the same weights which is not surprising because of Theorem 2. Because Designs 1 and 3 have different interesting estimators, they have the same doses but different weights and criterion values. From the numerical results in Table 3, it seems that A-optimal designs maybe based on only three points. Additional numerical results (not shown) support this conjecture.

In contrast, Table 4 shows optimal designs given a set of doses under the simple model (1).

Note that the criterion for the biologically motivated model is much less than one for the simple model when there are more than three design points. This shows that parameterizing the mean function greatly increase information.

### 3.4. c-Optimal design for estimating the peak dose

In this section, c-optimal designs are obtained for estimating the peak dose most precisely. The peak dose is the dose producing maximum mean response. Under mean dose model (6), the peak dose is the solution of the following equation:

$$\frac{\partial \mu}{\partial x} = (\beta + \gamma) \exp[-(\alpha + \beta x + \gamma x)] - \gamma \exp[-\gamma x] = 0.$$

In this model, the peak dose can be expressed in an explicit form:

$$x^* = -\frac{1}{\beta} \left\{ \log \left( \frac{\gamma}{\beta + \gamma} \right) + \alpha \right\}. \quad (9)$$

The c-optimality criterion for estimating the peak dose is to minimize

$$\text{Var}(\hat{x}^*) = x^{*\prime} M(\xi; \Theta)^{-1} [x^{*\prime}]^\top = \text{tr}[M(\xi; \Theta)^{-1} Q],$$

where

$$x^{*\prime} = \left( \frac{\partial x^*}{\partial \alpha} \quad \frac{\partial x^*}{\partial \beta} \quad \frac{\partial x^*}{\partial \gamma} \right) = \left( -\frac{1}{\beta} \quad \frac{1}{\beta^2} \left\{ \log\left(\frac{\gamma}{\beta+\gamma}\right) + \alpha \right\} + \frac{1}{\beta(\beta+\gamma)} \quad -\frac{1}{\gamma(\beta+\gamma)} \right)$$

and  $Q = [x^{*\prime}]^\top x^{*\prime}$ . We have the following theorem.

**Theorem 5.** *The nonnegative solutions of  $(\partial/\partial w) \text{tr}[M(\xi; \Theta)^{-1} Q] = 0$  are the A-optimal design weights.*

**Proof.** See A.2 in the appendix. Note  $Q$  has the same property as  $V$  in Theorem 4 because both  $Q$  and  $V$  are symmetric nonnegative definite matrices and independent of  $w$ . □

The weights can be obtained through the Newton–Raphson method. Assuming  $\alpha = 0.11$ ,  $\beta = 1$ ,  $\gamma = 2$ , c-optimal weights are given in Table 5. This table suggests that optimal designs for estimating the peak dose under the c-optimality criterion have three points: one point from lower bound of the range of possible design points, one from the upper bound and one from somewhere in the interior of the range.

### 3.5. c-Optimal design for estimating the $EC_{50}$

There is an interest in estimating  $EC_{50}$  under c-optimal criterion assuming mean response model (6). The peak dose is given by (9) and the minimum dose is  $x=0$ . The mean response at the peak and minimum doses, respectively, are

$$\mu(x^*) = \left( \frac{\beta}{\beta+\gamma} \right) \exp \left[ \frac{\gamma}{\beta} \left\{ \log\left(\frac{\gamma}{\beta+\gamma}\right) + \alpha \right\} \right], \quad \mu(0) = 1 - \exp[-\alpha]. \tag{10}$$

By definition, the  $EC_{50}$  satisfies  $EC_{50} \leq x^*$  and

$$\{1 - \exp[-(\alpha + \beta EC_{50})]\} \exp[-\gamma EC_{50}] = \frac{1}{2} \left\{ \left( \frac{\beta}{\beta+\gamma} \right) \exp \left[ \frac{\gamma}{\beta} \left\{ \log\left(\frac{\gamma}{\beta+\gamma}\right) + \alpha \right\} \right] + 1 - \exp[-\alpha] \right\}. \tag{11}$$

Eq. (11) cannot be solved analytically, but the solution  $EC_{50}$  can be obtained numerically.

The c-optimality criterion is to minimize the variance of the estimate of  $EC_{50}$ :  $\Psi = \text{Var}(\hat{EC}_{50}) = EC'_{50} M(\xi; \Theta)^{-1} [EC'_{50}]^\top = \text{tr}[M(\xi; \Theta)^{-1} Q^*]$ , where  $Q^* = [EC'_{50}]^\top EC'_{50}$  and  $EC'_{50} = (\partial EC_{50} / \partial \alpha \quad \partial EC_{50} / \partial \beta \quad \partial EC_{50} / \partial \gamma)$ . The challenge is to obtain an expression for  $EC'_{50}$  without having an explicit expression for the  $EC_{50}$  which is a function of the unknown parameters of  $(\alpha, \beta, \gamma)$ . The method we use has two parts: (i) take first derivatives to (11) with respect to  $\alpha, \beta$ , and  $\gamma$ ; (ii) solve for the partial derivatives of  $EC_{50}$ . The application of this method to the problem at hand is given in Appendix A.3.

**Theorem 6.** *The nonnegative solutions of  $(\partial/\partial w) \text{tr}[M(\xi; \Theta)^{-1} Q^*] = 0$  are the A-optimal design weights.*

**Proof.** See Appendix A.2. Note  $Q^*$  has the same property as  $V$  in Theorem 4 because both  $Q^*$  and  $V$  are symmetric nonnegative definite matrices and independent of  $w$ . □

Again set  $\alpha = 0.11$ ,  $\beta = 1$ ,  $\gamma = 2$ . Insulating  $\mu(x^*) = 0.184$  and  $\mu(0) = 0.104$  into Eq. (11) yields  $EC_{50} = 0.073$  from which it follows that  $EC'_{50} = [-0.202, 0.004, 0.087]$ . Given this information, numerically derived c-optimal weights for estimating  $EC_{50}$  are shown in Table 6.

This table also suggests that the c-optimal designs for estimating the  $EC_{50}$  consists of three points including the lower bound of the range of possible design points, the upper bound and one point somewhere in the interior.

## 4. A composite D-optimal design for $(\alpha, \beta)$

In Section 3, the doses are fixed. This restriction is sensible when interest is in estimating the adjacent mean differences. For the peak dose and the  $EC_{50}$  cases, this restriction is made due to the complexity of the problem. But what about optimal designs that determine both doses and weights? In this section, we consider this challenge. As is often the case in practice,

**Table 5**  
c-Optimal weights for estimating the peak dose for selected sets of design doses.

| Design | Design doses                       | Optimal weights              | $\Psi$ |
|--------|------------------------------------|------------------------------|--------|
| 1      | 0.00, 0.50, 1.00                   | 0.246, 0.172, 0.581          | 18.99  |
| 2      | 0.00, 0.30, 0.70, 1.00             | 0.271, 0.145, 0, 0.584       | 13.74  |
| 3      | 0.00, 0.25, 0.50, 0.75, 1.00       | 0.269, 0.147, 0, 0, 0.583    | 12.992 |
| 4      | 0.00, 0.20, 0.40, 0.60, 0.80, 1.00 | 0.261, 0.155, 0, 0, 0, 0.583 | 12.364 |



**Table 6**  
c-Optimal weights for estimating the  $EC_{50}$  for selected sets of design doses.

| Design | Design doses                       | Optimal weights              | $\Psi$ |
|--------|------------------------------------|------------------------------|--------|
| 1      | 0.00, 0.50, 1.00                   | 0.112, 0.185, 0.602          | 8.184  |
| 2      | 0.00, 0.30, 0.70, 1.00             | 0.172, 0.252, 0, 0.574       | 5.308  |
| 3      | 0.00, 0.25, 0.50, 0.75, 1.00       | 0.191, 0.253, 0, 0, 0.555    | 5.142  |
| 4      | 0.00, 0.20, 0.40, 0.60, 0.80, 1.00 | 0.214, 0.259, 0, 0, 0, 0.525 | 5.220  |

consider the case in which interest only lies in the two parameters  $\{\alpha, \beta\}$ . The parameter  $\gamma$  is nuisance parameter, which we deal with by considering a collection of fixed values of  $\gamma$ .

So now take  $\gamma = \gamma_0$  to be fixed and known. The model for  $\mu_i$  becomes

$$\mu_i = \{1 - \exp[-(\alpha + \beta x_i)]\} \exp[-\gamma_0 x_i].$$

To efficiently estimate  $\{\alpha, \beta\}$ , one needs only to consider the Fisher information matrix for  $\{\alpha, \beta\}$ :

$$M(\xi; \alpha, \beta) = \frac{1}{\sigma^2} \sum_{i=1}^K w_i \begin{pmatrix} \varphi_1(c_i) & x_i \varphi_1(c_i) \\ x_i \varphi_1(c_i) & x_i^2 \varphi_1(c_i) \end{pmatrix},$$

where  $\varphi_1(c_i) = e^{-2c_i}$  and  $c_i = \alpha + \beta x_i + \gamma_0 x_i$ . Some routine algebra establishes the following result.

**Lemma 7.**

$$M(\xi; \alpha, \beta) = A^T(\alpha, \beta) C(\xi; \alpha, \beta) A(\alpha, \beta),$$

where

$$A^T(\alpha, \beta) = \frac{1}{\sigma} \begin{pmatrix} 1 & 0 \\ \alpha & 1 \\ -\beta + \gamma_0 & \beta + \gamma_0 \end{pmatrix},$$

$$C(\xi; \alpha, \beta) = \begin{pmatrix} \sum_{i=1}^K w_i \varphi_1(c_i) & \sum_{i=1}^K w_i \varphi_2(c_i) \\ \sum_{i=1}^K w_i \varphi_2(c_i) & \sum_{i=1}^K w_i \varphi_3(c_i) \end{pmatrix},$$

$$\varphi_2(c_i) = c_i e^{-2c_i}, \text{ and } \varphi_3(c_i) = c_i^2 e^{-2(c_i)}.$$

Loewner ordering is adopted here. The Loewner order states that  $X \leq Y$  if  $Y - X$  is positive semidefinite matrix. If an information matrix  $M(\xi; \alpha, \beta) \leq M(\xi_s; \alpha, \beta)$  (Loewner ordering), then design  $\xi_s$  is not inferior to design  $\xi$  under commonly used optimality criteria (such as A- and D-optimality) and  $\xi_s$  is said to dominate  $\xi$ . In the locally optimal context, from Lemma 7, it follows that  $C(\xi; \alpha, \beta) \leq C(\xi_s; \alpha, \beta)$  implies that  $M(\xi; \alpha, \beta) \leq M(\xi_s; \alpha, \beta)$ .

We say a design for a two parameter model has a simple format if it has two support points. If for any design  $\xi$ , there exists a design  $\xi_s$  with a simple format such that  $C(\xi; \alpha, \beta) \leq C(\xi_s; \alpha, \beta)$ , then we can study the simple design  $\xi_s$  instead of  $\xi$ . Designs with a simple format greatly simplify the optimality problem.

Next, we show that there exists a design  $\xi_s$  with a simple format such that  $C(\xi; \alpha, \beta) \leq C(\xi_s; \alpha, \beta)$ . If the sample space is normalized such that  $x_i \in (0, 1)$ , then by definition  $c_i \in (\alpha, \alpha + \beta + \gamma_0]$ . Yang and Stufken (2009) show that if  $\Psi_1(c)$ ,  $\Psi_2(c)$  and  $\Psi_3(c)$  in  $C(\xi; \alpha, \beta)$  satisfy the following conditions, then all designs are dominated by a design with only two points, with  $c_1 = \alpha$  and  $c_2 \in (\alpha, \alpha + \beta + \gamma_0]$ :

- (i)  $\Psi_1(c)$ ,  $\Psi_2(c)$ , and  $\Psi_3(c)$  are continuous functions on  $[A, B]$  that are three times differentiable on  $(A, B)$ ;
- (ii)  $\Psi_1'(c)(\Psi_2'(c)/\Psi_1'(c))' / (\Psi_3'(c)/\Psi_1'(c))' / (\Psi_2'(c)/\Psi_1'(c))' < 0$  for  $c \in (A, B)$ ; and
- (iii)  $\lim_{c \rightarrow A} \Psi_2'(c)/\Psi_1'(c)(\Psi_1(A) - \Psi_1(c)) = 0$ .

Easily it can be shown that  $\Psi_1(c)$ ,  $\Psi_2(c)$  and  $\Psi_3(c)$  satisfy these three conditions. Thus for any design  $\xi$ , there exists a design  $\xi_s$  with two points,  $x_1 = 0$  and  $x_2 \in (0, 1)$ , such that  $C(\xi; \alpha, \beta) \leq C(\xi_s; \alpha, \beta)$ . Thus, we can focus on the simple format designs to derive some specific optimal designs.

Consider the D-optimal design for estimating  $(\alpha\beta)$  using Fisher information matrix  $C(\xi_s; \alpha, \beta)$ . Since  $c_i \in (\alpha, \alpha + \beta + \gamma_0]$ , only the two point designs  $\{(0, w_1); (x_2, w_2)\}$  need to be considered, where  $x_2 \in (0, 1]$ . Because

$$\det(C(\xi_s, \alpha, \beta)) = w_1 w_2 [x_2 \varphi_1(c_2)]^2,$$

the D-optimal design obtained is  $\{(0, 1/2); (x_2^*, 1/2)\}$ , where  $x_2^* = 1/(\beta + \gamma_0)$  if  $\beta + \gamma_0 \geq 1$  or 1 if  $\beta + \gamma_0 < 1$ . Although it is an analytical solution, the design depends on the value of  $\gamma$ . Because the exact value of  $\gamma$  is unknown, we propose a composite design that is constructed from a collection of D-optimal designs with fixed  $\gamma$ , letting  $\gamma$  take on a range of

values:  $\gamma \in \{\gamma_1, \gamma_2, \dots, \gamma_s\}$ . The composite designs are constructed from the simple optimal designs

$$\zeta_i^* = \begin{pmatrix} 0 & \frac{1}{\beta + \gamma_i} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad \gamma = \gamma_i, \quad i = 1, 2, \dots, s.$$

Suppose equal weights are given to each individual design yielding the composite design

$$\zeta_c^* = \begin{pmatrix} 0 & \frac{1}{\beta + \gamma_1} & \frac{1}{\beta + \gamma_2} & \dots & \frac{1}{\beta + \gamma_s} \\ \frac{1}{2} & \frac{1}{2s} & \frac{1}{2s} & \dots & \frac{1}{2s} \end{pmatrix}. \tag{12}$$

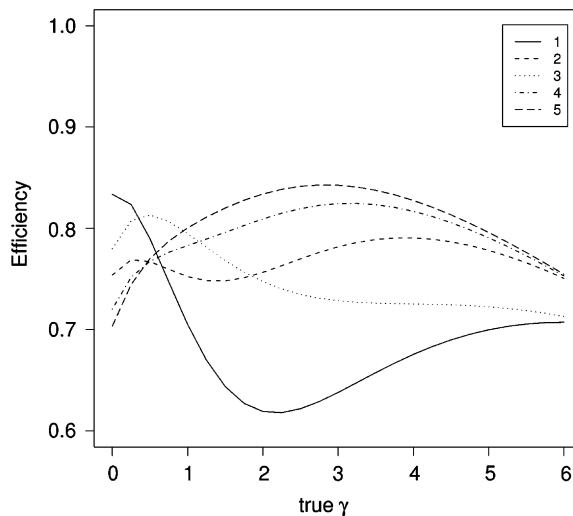
The composite design protects against the uncertainty in  $\gamma$  by using a range of  $\gamma$  values. The efficiencies of composite design compared to a D-optimal design at  $\gamma = \gamma_i$  are

$$Eff_i = \left\{ \frac{|M_{\zeta_c^*}|}{|M_{\zeta_i^*}|} \right\}_{\gamma = \gamma_i}^{1/2}.$$

Then the composite optimal design is quite good compared to D-optimal design is shown in the following example.

**Table 7**  
A composite D-optimal designs for different ranges of  $\gamma$ .

| Design | Values of $\gamma$ used to construct the composite design | $\zeta_c^*$   |
|--------|---|---|
| 1      | 0.00, 6.00  | $\begin{pmatrix} 0.00 & 0.14 & 1.00 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$   |
| 2      | 0.00, 3.00, 6.00  | $\begin{pmatrix} 0.00 & 0.14 & 0.25 & 1.00 \\ \frac{1}{2} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$  |
| 3      | 0.00, 0.80, 5.20, 6.00                                    | $\begin{pmatrix} 0.00 & 0.14 & 0.16 & 0.55 & 1.00 \\ \frac{1}{2} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}$                           |
| 4      | 0.00, 2.00, 4.00, 6.00                                    | $\begin{pmatrix} 0.00 & 0.14 & 0.20 & 0.33 & 1.00 \\ \frac{1}{2} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}$                           |
| 5      | 0.00, 1.50, 3.00, 4.50, 6.00                              | $\begin{pmatrix} 0.00 & 0.14 & 0.18 & 0.25 & 0.40 & 1.00 \\ \frac{1}{2} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} \end{pmatrix}$ |



**Fig. 3.** Efficiencies of composite D-optimal designs as a function of the true value of  $\gamma$ .

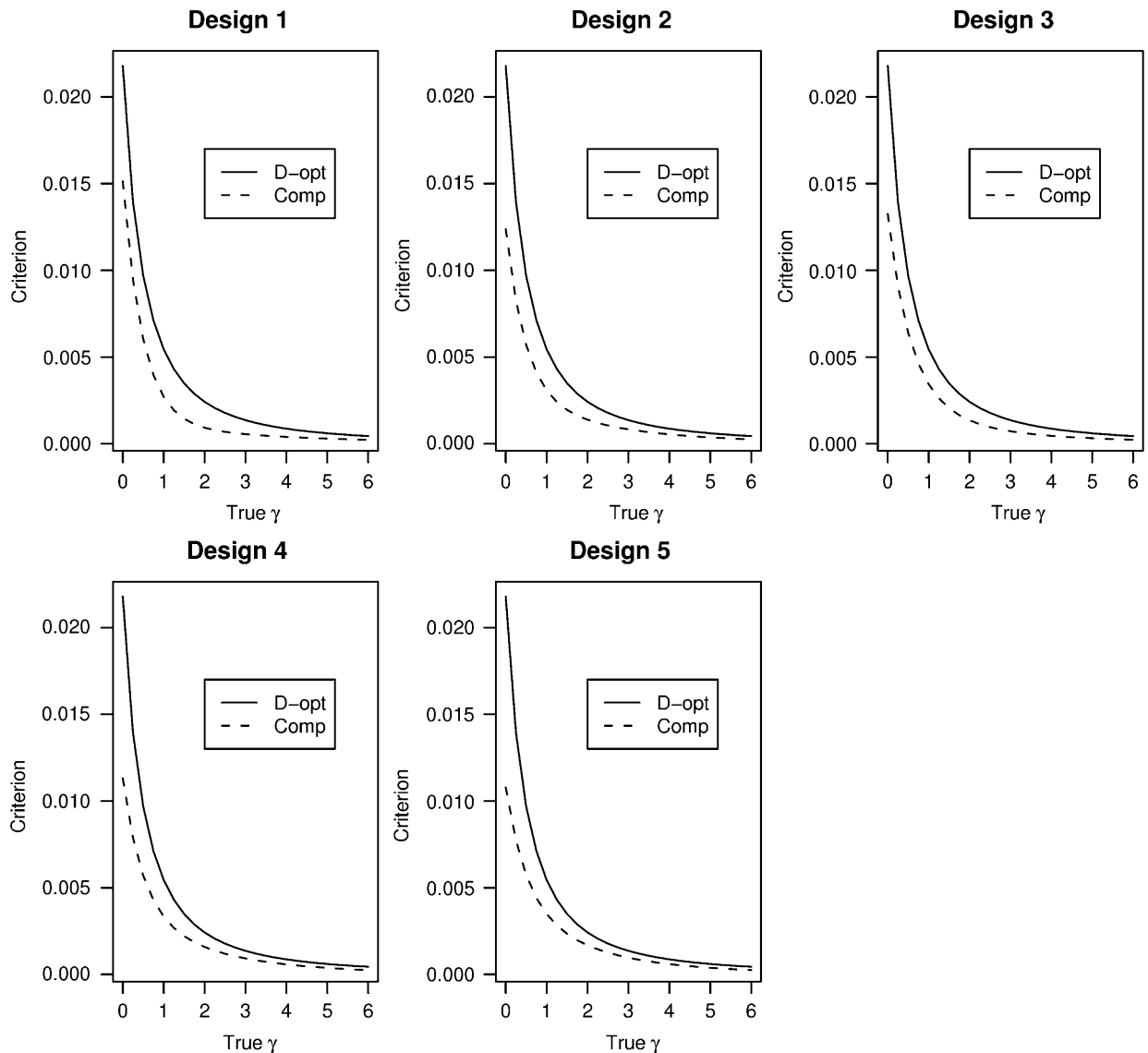


Fig. 4. Information of D-optimal design against composite D-optimal designs.

**Example.** Suppose  $\alpha = 0.11$ ,  $\beta = 1.00$ . Table 7 shows the composite designs for different ranges of values of  $\gamma$ . The efficiencies of the composite designs relative to a D-optimal design at  $\gamma = \gamma_i$  are shown in Fig. 3, where  $\gamma \in \{0.00, 0.10, 0.20, \dots, 5.90, 6.00\}$ .

Fig. 3 shows that all efficiencies from composite designs are greater than 0.6. Design 1 does not have a good efficiency with two values of  $\gamma$ . However, the efficiency improves a great deal when one or two more values of  $\gamma$  are added, as is the case for Designs 2, 3 and 4 shown in Fig. 3. Design 5 has five values of  $\gamma$  but the efficiency does not improve much. Fig. 4 shows that how much information the composite designs have compared to D-optimal design at true value of  $\gamma$ . When true value of  $\gamma$  is greater than 2, the D-optimality criterion goes to zero indicating that the designs have little information to lose. Thus, it is more meaningful to consider efficiencies at the lower values of true  $\gamma$ . Figs. 3 and 4 show that the composite design with three or four values of  $\gamma$  works very well compared to the D-optimal design at fixed and known  $\gamma$ . In this example, the composite design 3 is recommended as the best choice because it has pretty good efficiency at the lower true value of  $\gamma$ .

We can not guarantee that the composite design weights are the best. However, the proposed weights are better than uniformly distributed weights. Figs. 5 and 6 show that the efficiency of the composite designs is always to be greater when using uniformly distributed weights at the same doses.

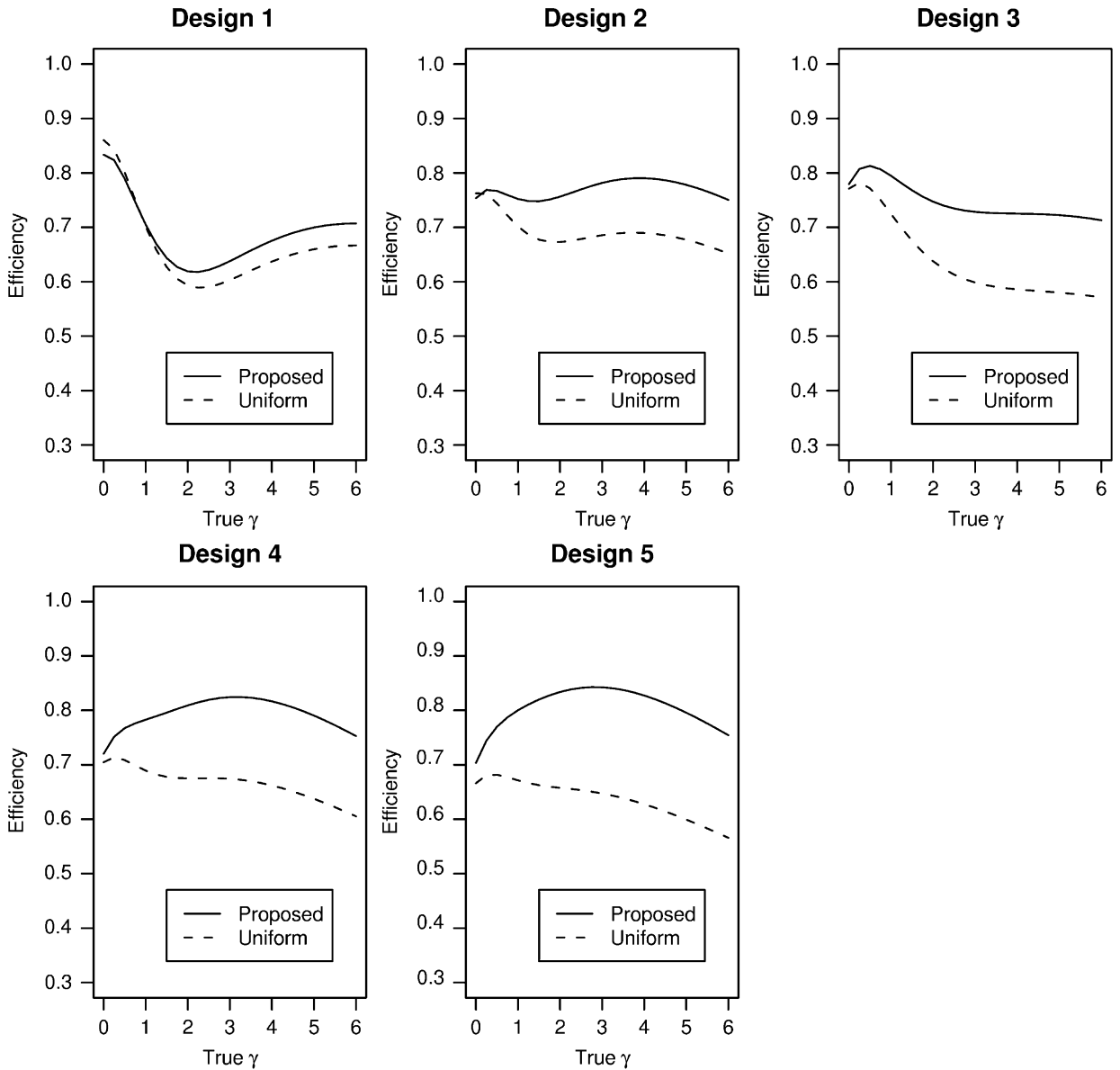


Fig. 5. Efficiency of the proposed composite design against uniformly distributed weights at the same doses.

The A-optimality criteria also can be used to construct composite designs. In this paper, the D-optimal design is used rather than A-optimal design because D-optimal design for estimating the parameter  $\{\alpha, \beta\}$  has the explicit form.

## 5. Conclusion

We obtained  $D_A$ -optimal and A-optimal weights. In the case of a simple model, the analytical solutions for both  $D_A$ - and A-optimality are given. For a model inspired by Margolin et al. (1981), numerical methods are developed to determine how many subjects to assign to each dose for several different objective functions.

When interest is in estimating adjacent mean differences, doses are fixed because the variance of mean differences is minimized by putting all treatments at one point, which is not productive. c-Optimal weights for the peak dose and the  $EC_{50}$  are also obtained. But there is the limitation that the  $D_A$ -, A- and c-optimal weights can be obtained only numerically if the number of doses is not the same as the number of parameters. A procedure to approximate a D-optimal design in an explicit form is developed in Section 4. In this procedure, D-optimal designs for fixed  $\gamma$  are utilized. Notice that we assign  $\gamma$  a value from a certain range rather than estimating it. This reduces the sensitivity of optimal designs from wrong estimate of  $\gamma$ .

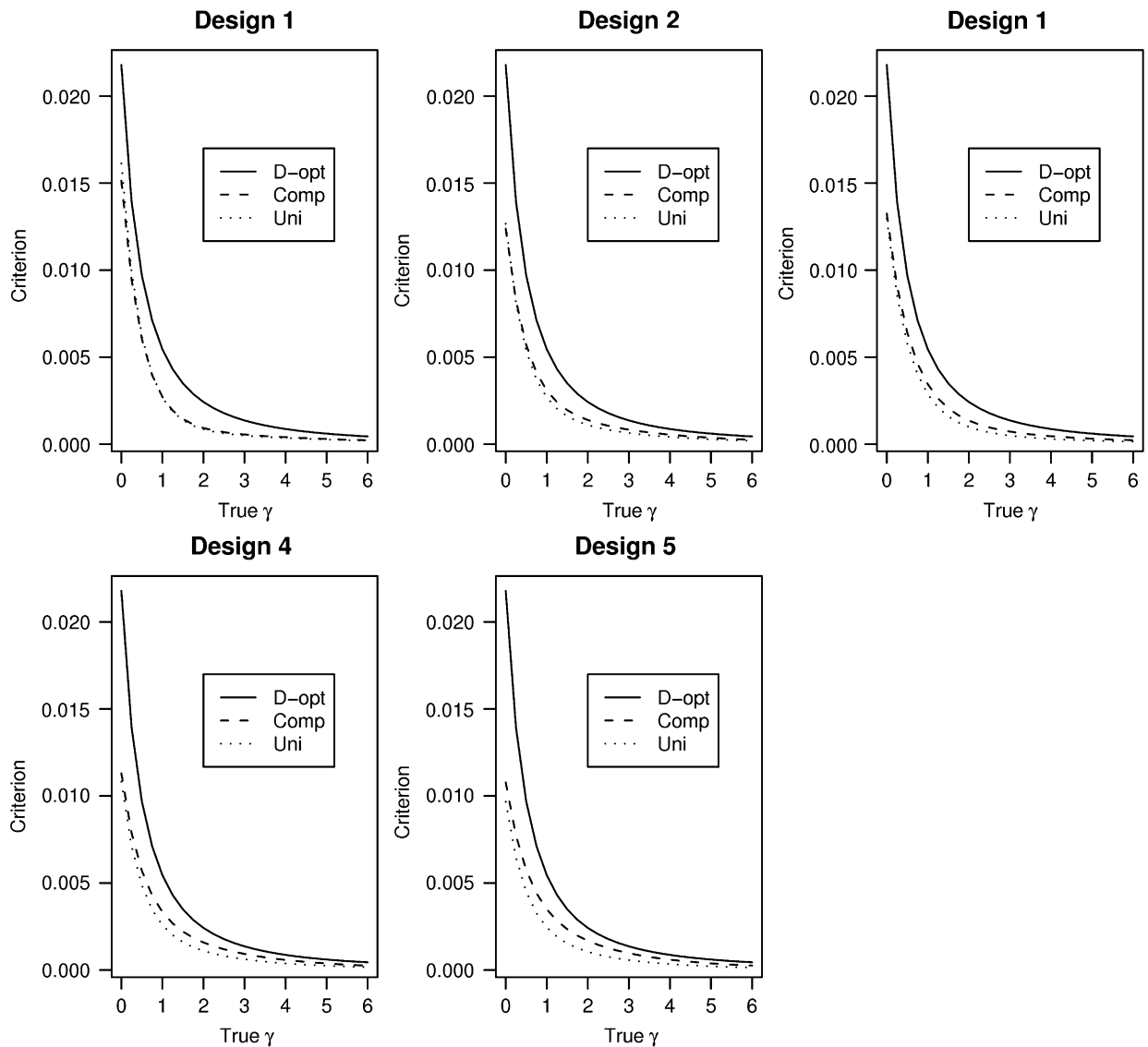


Fig. 6. Information of D-optimal design vs. the proposed composite design vs. uniformly distributed weights.

Optimal designs give the most information but sometimes researchers want to have designs that have more design points but still are very close to optimal. The composite design can handle this problem. It is easier for the researcher to use our method which is a partially analytical search rather than a fully numerical search.

Although this paper presents an analytical solution for the composite D-optimal design, there still are unsolved questions: how best to select the values of  $\gamma$  and how to assign weights for each  $\gamma$ . It is a good idea to get the values of  $\gamma$  using prior knowledge from a biologist. In this paper, weights are uniformly distributed. The reason is that our biology consultants had no idea. We know this is not entirely satisfactory, but it is a place to start and it seems to work well. We hope this paper stimulates the development of other methods that bring better results. Possibilities include a minimax solution or Bayesian formulation.

We model mean responses by combining an increasing response function with a decreasing response function. Most results are based on exponential functions as inspired by Margolin et al. (1981). Analysis of the data in Welshons et al. (2003) suggests an increasing logistic function paired with a decreasing exponential function would provide a better fit to their data. Thus, optimality results for this combination, or better, more general results are a future goal.

Also, in this paper identical variances are assumed at each dose. Allowing the variance to increase with the mean response is another important generalization.

Discussions with biologists suggest many features of the model are of interest. Although peak dose is of interest, designing for it is not easy. Eq. (9) shows the expression for the peak dose involves ratios of parameters. Hence, estimates

involve the ratio of asymptotically normal random variables. Thus extremely large sample sizes may be needed for the information matrices for estimating the peak dose and the  $EC_{50}$  to be a good approximation to the true covariance matrices. We use the classical Taylor series linear approximation to obtain estimates of the variance of features such as the peak dose and the  $EC_{50}$ . For maximum likelihood estimates of nonlinear functions of the parameters, alternative transformations proposed by [Duty and Flournoy \(2007, 2009\)](#) produce confidence intervals of equal width but better coverage. Their ideas may be useful in the context of the motivating example for this paper.

The current paper initiates the development of design for response functions with a downturn and we anticipate that it will lead to the development of alternative ways of learning about the features of interest.

**Acknowledgements**

We are grateful to the associate and executive editors as well as the referees for their helpful comments and suggestions, which led to a significant improvement of the quality of this paper.

**Appendix A**

*A.1. Proof of Theorem 3*

The minima must satisfy two conditions:

Condition (1). The first derivatives of (8) with respect to  $w_i$  are zero:

$$\frac{\partial}{\partial w_i} [\det[BM^-(\xi; \Theta)B^T]] = 0.$$

Condition (2). T is nonnegative definite matrix and T has elements

$$T_{ij} = \frac{\partial^2}{\partial w_i \partial w_j} [\det[BM^-(\xi; \Theta)B^T]], \quad i, j \in \{1, 2, \dots, K-1\}.$$

To derive  $T_{ij}$ , we use theorems in [Harville \(1997\)](#). Set  $C = [BM^-(\xi; \Theta)B^T]$ ;  $F_i = B(M^-(\xi; \Theta)(\partial M(\xi; \Theta)/\partial w_i)M^-(\xi; \Theta))B^T$  and  $L_{ij} = [L_{ji}]^T = M^-(\xi; \Theta)(\partial M(\xi; \Theta)/\partial w_i)M^-(\xi; \Theta)(\partial M(\xi; \Theta)/\partial w_j)M^-(\xi; \Theta) = l_i l_j^T$ , where  $l_i = M^-(\xi; \Theta)(\partial M(\xi; \Theta)/\partial w_i)P_1 D_1^{1/2}$ .

Let  $T^*$  be a  $(K-1) \times (K-1)$  matrix with elements

$$\begin{aligned} T_{ij}^* &= \frac{\partial^2}{\partial w_i \partial w_j} [\log \det[BM^-(\xi; \Theta)B^T]] \\ &= \text{tr}(C^{-1}B(L_{ij} + [L_{ij}]^T)B^T) - \text{tr}(C^{-1}F_i C^{-1}F_j) \\ &= \underbrace{\text{tr}(C^{-1}BL_{ij}B^T)}_{T_{ij}^{(1)}} + \underbrace{\text{tr}(C^{-1}BL_{ji}B^T) - \text{tr}(C^{-1}F_i C^{-1}F_j)}_{T_{ij}^{(2)}}. \end{aligned}$$

Let  $T \geq 0$  mean that T is nonnegative definite (NND) matrix. Since  $\det[C] > 0$ ,  $T \geq 0$  if and only if  $T^* \geq 0$ , we need to show that  $T^*$  is NND matrix.  $L_{ij}$  is a symmetric NND matrix because  $M^-(\xi; \Theta)$  is a full rank NND matrix. So, there exists nonsingular matrix  $P_1$  and diagonal matrix  $D_1$  such that  $M^-(\xi; \Theta) = P_1 D_1 P_1^T$ .

Now  $T^* = T^{(1)} + T^{(2)}$ , where  $T^{(1)} = (T_{ij}^{(1)})$  and  $T^{(2)} = (T_{ij}^{(2)})$  are  $(K-1) \times (K-1)$  matrices with elements given above.  $T^*$  is NND matrix if and only if  $T^{(1)}$  and  $T^{(2)}$  are NND matrices. Consider the elements of  $T^{(1)}$  first. Now there exists nonsingular matrix  $P_2$  and diagonal matrix  $D_2$  such that  $P_2 D_2 P_2^T = C^{-1}$  because  $C^{-1}$  is symmetric NND matrix. Define  $U_i = [Bl_i]^T P_2 D_2^{1/2}$  and denote the first, second, ..., sth rows of  $U_i$  by  $u_1, u_2, \dots, u_s$ . Let  $G_i = \text{vec}(U_i) = (u_1 u_2 \dots u_s)$ .

$$\begin{aligned} T_{ij}^{(1)} &= \text{tr}(C^{-1}BL_{ij}B^T) = \text{tr}(C^{-1}BL_{ji}B^T) = \text{tr}(C^{-1}Bl_j[Bl_i]^T) \\ &= \text{tr}([Bl_i]^T C^{-1}Bl_j) = \text{tr}(U_i U_j^T) = \text{vec}(U_i)[\text{vec}(U_j)]^T \\ &= G_i [G_j]^T. \end{aligned}$$

Define  $G = (G_1 \ G_2 \ \dots \ G_{k-1})^T$ . Then  $T^{(1)} = GG^T$ . Therefore  $T^{(1)}$  is NND.

Now let  $\otimes$  represents the Kronecker product and consider the elements of  $T^{(2)}$ :

$$\begin{aligned} T_{ij}^{(2)} &= \text{tr}(C^{-1}BL_{ij}B^T) - \text{tr}(C^{-1}F_i C^{-1}F_j) \\ &= \text{tr}([Bl_i]^T C^{-1}Bl_j) - \text{tr}(C^{-1}F_j C^{-1}F_i) \\ &= \text{tr}(l_i^T B^T C^{-1}Bl_j) - \text{tr}(C^{-1}F_j C^{-1}F_i) \\ &= \text{tr}(l_i^T B^T C^{-1}Bl_j) - \text{tr}(C^{-1}BP_1 D_1^{1/2} l_j^T B^T C^{-1}Bl_i D_1^{1/2} P_1^T B^T) \\ &= \text{vec}(l_i^T) \text{diag}(B^T C^{-1}, B^T C^{-1}, B^T C^{-1}) \text{vec}(l_j^T)^T - \text{tr}(l_i^T B^T C^{-1}Bl_j D_1^{1/2} P_1^T B^T C^{-1}BP_1 D_1^{1/2}) \end{aligned}$$

$$\begin{aligned}
 &= \text{vec}(l_i^\top) [I \otimes (B^\top C^{-1} B)] \text{vec}(l_j^\top)^\top - \text{tr}(l_i^\top B^\top C^{-1} B l_j D_1^{1/2} P_1^\top B^\top C^{-1} B P_1 D_1^{1/2}) \\
 &= \text{vec}(l_i^\top) [I \otimes (B^\top C^{-1} B)] \text{vec}(l_j^\top)^\top - \text{vec}(l_i^\top) [(B^\top C^{-1} B) \otimes (D_1^{1/2} P_1^\top B^\top C^{-1} B P_1 D_1^{1/2})] \text{vec}(l_j^\top)^\top \\
 &= \text{vec}(l_i^\top) [(B^\top C^{-1} B) \otimes (I - D_1^{1/2} P_1^\top B^\top C^{-1} B P_1 D_1^{1/2})] \text{vec}(l_j^\top)^\top.
 \end{aligned}$$

Both  $B^\top C^{-1} B$  and  $I - D_1^{1/2} P_1^\top B^\top C^{-1} B P_1 D_1^{1/2}$  are symmetric NND matrices because  $C^{-1}$  is symmetric NND matrix and  $I - D_1^{1/2} P_1^\top B^\top C^{-1} B P_1 D_1^{1/2}$  is idempotent matrix. Then the Kronecker product  $\otimes$  of two symmetric NND matrix is also symmetric NND matrix. Thus, there exists nonsingular matrix  $P_3$  and diagonal matrix  $D_3$  such that  $(B^\top C^{-1} B) \otimes (I - D_1^{1/2} P_1^\top B^\top C^{-1} B P_1 D_1^{1/2}) = P_3 D_3 P_3^\top$ . Then

$$\begin{aligned}
 T_{ij}^{(2)} &= \text{vec}(l_i^\top) P_3 D_3 P_3^\top \text{vec}(l_j^\top)^\top \\
 &= \text{vec}(l_i^\top) P_3 D_3^{1/2} D_3^{1/2} P_3^\top \text{vec}(l_j^\top)^\top \\
 &= F_i [F_j]^\top,
 \end{aligned}$$

where  $F_i = \text{vec}(l_i^\top) P_3 D_3^{1/2}$ . Set  $F = (F_1 \ F_2 \ \dots \ F_{k-1})^\top$ . Then  $T^{(2)} = FF^\top$ . Therefore  $T^{(2)}$  is NND matrix.

We have proved that  $T$  is NND matrix because  $T^{(1)}$  and  $T^{(2)}$  are NND matrix. Condition (2) is always true since  $T$  is NND matrix. Therefore, the criterion (8) is minimized if Condition (1) is satisfied.

### A.2. Proof of Theorem 4

The criterion  $\Psi$  is minimized when the first derivatives of (9) with respect to  $w_i$  are zero and the second derivatives are nonnegative. To derive the second derivatives, theorems in Harville (1997) are used again. The A-optimal design must satisfy Condition (1)

$$\frac{\partial}{\partial w_i} [\text{tr}(M(\xi; \Theta)^{-1} V)] = 0.$$

Condition (2).  $H$  is NND, where  $H$  is a  $(K-1) \times (K-1)$  matrix with elements

$$\begin{aligned}
 H_{ij} &= \frac{\partial^2}{\partial w_i \partial w_j} [\text{tr}(M(\xi; \Theta)^{-1} V)] \\
 &= \text{tr} \left\{ V \left[ \left( M(\xi; \Theta)^{-1} \frac{\partial M(\xi; \Theta)}{\partial w_i} M(\xi; \Theta)^{-1} \frac{\partial M(\xi; \Theta)}{\partial w_j} M(\xi; \Theta)^{-1} \right) + \left( M(\xi; \Theta)^{-1} \frac{\partial M(\xi; \Theta)}{\partial w_j} M(\xi; \Theta)^{-1} \frac{\partial M(\xi; \Theta)}{\partial w_i} M(\xi; \Theta)^{-1} \right) \right] \right\}.
 \end{aligned}$$

Again, let  $L_{ij} = l_i l_j^\top$ , where  $l_i = M(\xi; \Theta)^{-1} (\partial M(\xi; \Theta) / \partial w_i) P_1 D_1^{1/2}$ . Applying the chain rule,  $H_{ij}$  can be rewritten as

$$\begin{aligned}
 H_{ij} &= \text{tr}\{V(L_{ij} + L_{ji})\} = \text{tr}(V L_{ij}) + \text{tr}(V L_{ji}) \\
 &= 2 \text{tr}(V L_{ji}) = 2 \text{tr}(V l_j l_i^\top) \\
 &= 2 \text{tr}(l_i^\top V l_j) = 2 \text{tr}(Q_i Q_j^\top) \\
 &= 2 \text{vec}(Q_i) [\text{vec}(Q_j)]^\top = 2 R_i [R_j]^\top,
 \end{aligned}$$

where  $Q_i = l_i^\top P_4 D_4^{1/2}$  and  $R_i = \text{vec}(Q_i)$ .

There exists a nonsingular matrix  $P_4$  and diagonal matrix  $D_4$  such that  $P_4 D_4 P_4^\top = V$  because  $V$  is symmetric NND. Set  $R = (R_1 \ R_2 \ \dots \ R_{k-1})^\top$ . Then  $H = 2RR^\top$ . Therefore,  $H$  is NND. Again, Condition (2) is always satisfied because  $H$  is NND and the A-optimal weights are obtained if Condition (1) is satisfied.

### A.3. Obtaining $EC_{50}$ without an explicit expression for $EC_{50}$

1. Take first derivatives to (11) with respect to  $\alpha, \beta$ , and  $\gamma$ :

$$\begin{aligned}
 \frac{\partial(12)}{\partial \alpha} &= -\gamma \frac{EC_{50}}{\partial \alpha} \exp[-\gamma EC_{50}] + \exp[-(\alpha + \beta EC_{50} + \gamma EC_{50})] + \beta \frac{\partial EC_{50}}{\partial \alpha} \exp[-(\alpha + \beta EC_{50} + \gamma EC_{50})] \\
 &\quad + \gamma \frac{\partial EC_{50}}{\partial \alpha} \exp[-(\alpha + \beta EC_{50} + \gamma EC_{50})] - \frac{1}{2} \left\{ \left( \frac{\beta}{\beta + \gamma} \right) \frac{\gamma}{\beta} \exp \left[ \frac{\gamma}{\beta} \left\{ \log \left( \frac{\gamma}{\beta + \gamma} \right) + \alpha \right\} \right] + \exp[-\alpha] \right\} = 0,
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial(12)}{\partial \beta} &= -\gamma \frac{EC_{50}}{\partial \beta} \exp[-\gamma EC_{50}] + \left( EC_{50} + \beta \frac{\partial EC_{50}}{\partial \beta} \right) \exp[-(\alpha + \beta EC_{50} + \gamma EC_{50})] + \gamma \frac{\partial EC_{50}}{\partial \beta} \exp[-(\alpha + \beta EC_{50} + \gamma EC_{50})] \\
 &\quad - \frac{\gamma}{2(\beta + \gamma)^2} \exp \left[ \frac{\gamma}{\beta} \log \left( \frac{\gamma}{\beta + \gamma} \right) + \frac{\gamma \alpha}{\beta} \right] + \frac{1}{2} \left( \frac{\beta}{\beta + \gamma} \right) \left[ \frac{\gamma \alpha}{\beta^2} + \frac{\gamma}{\beta^2} \log \left( \frac{\gamma}{\beta + \gamma} \right) + \frac{\gamma}{\beta(\beta + \gamma)} \right] \exp \left[ \frac{\gamma}{\beta} \log \left( \frac{\gamma}{\beta + \gamma} \right) + \frac{\gamma \alpha}{\beta} \right] = 0,
 \end{aligned}$$

$$\begin{aligned} \frac{\partial(12)}{\partial\gamma} &= \left(-EC_{50}-\gamma\frac{EC_{50}}{\partial\gamma}\right)\exp[-\gamma EC_{50}] + \left(EC_{50}+\gamma\frac{\partial EC_{50}}{\partial\gamma}\right)\exp[-(\alpha+\beta EC_{50}+\gamma EC_{50})] \\ &+ \beta\frac{\partial EC_{50}}{\partial\gamma}\exp[-(\alpha+\beta EC_{50}+\gamma EC_{50})] - \frac{\beta}{2(\beta+\gamma)^2}\exp\left[\frac{\gamma}{\beta}\log\left(\frac{\gamma}{\beta+\gamma}\right)+\frac{\gamma\alpha}{\beta}\right] \\ &+ \frac{1}{2}\left(\frac{\beta}{\beta+\gamma}\right)\left[\frac{\alpha}{\beta}+\frac{1}{\beta}\log\left(\frac{\gamma}{\beta+\gamma}\right)+\frac{1}{(\beta+\gamma)}\right]\exp\left[\frac{\gamma}{\beta}\log\left(\frac{\gamma}{\beta+\gamma}\right)+\frac{\gamma\alpha}{\beta}\right] = 0. \end{aligned}$$

2. Solve for the partial derivatives of  $EC_{50}$ :

$$\begin{aligned} \frac{\partial EC_{50}}{\partial\alpha} &= \frac{1}{2}\left\{\left(\frac{\beta}{\beta+\gamma}\right)\frac{\gamma}{\beta}\exp\left[\frac{\gamma}{\beta}\log\left(\frac{\gamma}{\beta+\gamma}\right)+\frac{\gamma\alpha}{\beta}\right] + \exp[-\alpha]\right\} - \exp[-(\alpha+\beta EC_{50}+\gamma EC_{50})] \\ &\quad \frac{1}{(\beta+\gamma)\exp[-(\alpha+\beta EC_{50}+\gamma EC_{50})]-\gamma\exp[-\gamma EC_{50}]}, \\ \frac{\partial EC_{50}}{\partial\beta} &= \frac{1}{2}\left\{\frac{\gamma}{(\beta+\gamma)^2}\exp\left[\frac{\gamma}{\beta}\log\left(\frac{\gamma}{\beta+\gamma}\right)+\frac{\gamma\alpha}{\beta}\right] - \left(\frac{\beta}{\beta+\gamma}\right)\left\{\frac{\gamma\alpha}{\beta^2}+\frac{\gamma}{\beta^2}\log\left(\frac{\gamma}{\beta+\gamma}\right)+\frac{\gamma}{\beta(\beta+\gamma)}\right\}\exp\left[\frac{\gamma}{\beta}\log\left(\frac{\gamma}{\beta+\gamma}\right)+\frac{\gamma\alpha}{\beta}\right]\right\} \\ &\quad \frac{1}{(\beta+\gamma)\exp[-(\alpha+\beta EC_{50}+\gamma EC_{50})]-\gamma\exp[-\gamma EC_{50}]} \\ &\quad - \frac{EC_{50}\exp[-(\alpha+\beta EC_{50}+\gamma EC_{50})]}{(\beta+\gamma)\exp[-(\alpha+\beta EC_{50}+\gamma EC_{50})]-\gamma\exp[-\gamma EC_{50}]}, \\ \frac{\partial EC_{50}}{\partial\gamma} &= \frac{1}{2}\left\{\frac{\beta}{(\beta+\gamma)^2}\exp\left[\frac{\gamma}{\beta}\log\left(\frac{\gamma}{\beta+\gamma}\right)+\frac{\gamma\alpha}{\beta}\right] - \left(\frac{\beta}{\beta+\gamma}\right)\left\{\frac{\alpha}{\beta}+\frac{1}{\beta}\log\left(\frac{\gamma}{\beta+\gamma}\right)+\frac{1}{(\beta+\gamma)}\right\}\exp\left[\frac{\gamma}{\beta}\log\left(\frac{\gamma}{\beta+\gamma}\right)+\frac{\gamma\alpha}{\beta}\right]\right\} \\ &\quad \frac{1}{(\beta+\gamma)\exp[-(\alpha+\beta EC_{50}+\gamma EC_{50})]-\gamma\exp[-\gamma EC_{50}]} \\ &\quad - \frac{EC_{50}\exp[-(\alpha+\beta EC_{50}+\gamma EC_{50})]-EC_{50}\exp[-\gamma EC_{50}]}{(\beta+\gamma)\exp[-(\alpha+\beta EC_{50}+\gamma EC_{50})]-\gamma\exp[-\gamma EC_{50}]}. \end{aligned}$$

## References

- Atkinson, A.C., Donev, A.N., 1992. Optimum Experimental Designs. Oxford University Press.
- Brez, F., Hothorn, L.A., 2003. Statistical analysis of monotone or non-monotone dose-response data from in vitro toxicological assays. *Alternatives to Laboratory Animals* 31, 81–96.
- Duty, P., Flournoy, N., 2007. Practical considerations for constructing confidence intervals for nonlinear functions. In: Joint Statistical Meetings 2004 Proceeding.
- Duty, P., Flournoy, N., 2009. The composition method: a simple strategy for improving coverage of confidence intervals. TR-MU-STAT-2009-10-08, Department of Statistics, University of Missouri.
- Dragalin, V., Fedorov, V., Wu, Y., 2008. Adaptive designs for selecting drug combinations based on efficacy-toxicity response. *Journal of Statistical Planning and Inference* 138, 352–373.
- Fan, S.K., Chaloner, K., 2001. Optimal designs for a continuation-ratio model, mODa6 *Advances in Model-Oriented Design and Analysis*, Physica-Verlag, Heidelberg, pp. 77–86.
- Fedorov, V., Wu, Y., 2007a. Dose finding designs for continuous responses and binary utility. *Journal of Biopharmaceutical Statistics* 17, 1085–1096.
- Fedorov, V., Wu, Y., 2007b. Generalized probit model in design of dose finding experiments, mODa8 *Advances in Model-Oriented Design and Analysis*. Physica-Verlag, Heidelberg, pp. 67–73.
- Han, C., Chaloner, K., 2004. A note on optimal designs for two or more treatment groups. *Statistics Probability Letters* 69, 81–89.
- Harville, D.A., 1997. *Matrix Algebra from a Statistics Perspective*. Springer.
- Margolin, B.H., Kaplan, N., Zeiger, E., 1981. Statistical analysis of the Ames Salmonella/microsome test. *Proceedings of the National Academy of Sciences of the United States of America* 78, 3779–3783.
- Rabie, H.S., Flournoy, N., 2004. Optimal designs for contingent response models, mODa7 *Advances in Model-Oriented Design and Analysis*. Physica-Verlag, Heidelberg, pp. 133–142.
- Welshons, W.V., Thayer, K.A., Judy, B.M., Taylor, J.A., Curran, E.M., Saal, F.S., 2003. Large effects from small exposures. I. Mechanisms for endocrine-disrupting chemicals with estrogenic activity. *Environmental Health Perspectives* 111, 994–1006.
- Yang, M., Stufken, J., 2009. Support points of locally optimal designs for non-linear models with two parameters. *The Annals of Statistics* 37, 518–541.

## Further reading

- Bechhofer, R.E., Tamhane, A.C., 1981. Incomplete block designs for comparing treatments with a control: general theory. *Technometrics* 23, 45–57.
- Biedermann, S., Dette, H., Zhu, W., 2006. Optimal designs for dose-response models with restricted design spaces. *Journal of the American Statistical Association* 101, 747–759.
- Brez, F., Hothorn, L.A., 2001. Testing dose-response relationships with a priori unknown, possibly nonmonotone shapes. *Journal of Biopharmaceutical Statistics* 11, 193–207.
- Cheng, C.-S., Wu, C.-F., 1980. Balanced repeated measurements designs. *Annals of Statistics* 8, 1272–1283.
- Combes, R.D., 1997. Statistical analysis of dose-response data from in vitro assays: an illustration using Salmonella mutagenicity data. *Toxicology in Vitro* 11, 683–687.
- Hedayat, A.S., Majumdar, D., 1985. Families of A-optimal block designs for comparing test treatments with a control. *Annals of Statistics* 13, 757–767.
- Hedayat, A.S., Jacroux, M., Majumdar, D., 1988. Optimal designs for comparing test treatments with controls. *Statistical Science* 3, 462–476.



- Hedayat, A.S., Yan, B., Pezzuto, J.M., 1997. Modeling and identifying optimum designs for fitting dose-response curves based on raw optical density data. *Journals of the American Statistical Association* 92, 1132–1140.
- Hothorn, L.A., 2006. Multiple comparisons and multiple contrasts in randomized dose-response trials—confidence interval oriented approaches. *Journal of Biopharmaceutical Statistics* 16, 711–731.
- Kiefer, J., 1958. On the nonrandomized optimality and randomized nonoptimality of symmetrical designs. *Annals of Mathematical Statistics* 29, 675–699.
- Majumdar, D., Notz, W.I., 1983. Optimal incomplete block designs for comparing treatments with a control. *Annals of Statistics* 11, 258–266.
- Marcus, R., Genzi, A., 1994. Simultaneous confidence intervals for umbrella contrasts of normal means. *Computational Statistics & Data Analysis* 17, 393–407.
- Pukelsheim, F., 2006. *Optimal Design of Experiment*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.
- Simpson, D.G., Margolin, B.H., 1990. Nonparametric testing for dose-response curves subject to downturns: asymptotic power considerations. *Annals of Statistics* 18, 373–390.
- Stufken, J., 1987. A-optimal block designs for comparing test treatments with a control. *Annals of Statistics* 15, 1629–1638.