Stockholms universitet

Research Report

Department of Statistics

No. 2014:1

Optimal design problems for the bivariate Emax model

Bergrún Tinna Magnúsdóttir

Department of Statistics, Stockholm University, SE-106 91 Stockholm, Sweden

Optimal design problems for the bivariate Emax model

Bergrún Tinna Magnúsdóttir

Abstract

Finding a suitable dose is among the most difficult tasks during clinical development of a new drug. In early phases dose finding studies usually focus on finding a safe dose. Safety variables are thus of main interest. In later phases the focus is shifted towards efficacy. Typically a primary efficacy variable is defined and modeled. Various dose-response models have been suggested. For continuous responses among the most successful ones is the Emax model. Here both efficacy and safety are considered simultaneously and the Emax model is extended to a model with a bivariate response, one response being a primary efficacy variable and one being a primary safety variable. This model is referred to as the bivariate Emax model. The focus is on locally c-optimal designs for the bivariate Emax model and a simplified version of it. More specifically the locally c-optimal designs minimize the asymptotic variance for the estimate of the dose that maximizes the patient's utility. The utility is a function of the efficacy and safety variables and referred to as the Clinical Utility Index (CUI).

1 Introduction

This paper explores optimal design problems for non-linear, multivariate response models when the aim of the experiment is to estimate a function of the model parameters. In the optimal design literature such designs are referred to as c-optimal designs. The focus is on designs for dose finding studies when the aim is to estimate the most desirable dose of the drug under investigation. Effective drugs result in both effects and side-effects and hence the most desirable dose is the one that gives the best possible combination of these. The bivariate Emax model explains how a primary efficacy variable and a primary safety variable depend on the dose of the drug. This model is introduced in the next section. In Section 3 a tool for estimating the most desirable dose, The Clinical Utility Index (CUI), is introduced. The c-optimal design methodology is illustrated in Section 4 and the relevant General Equivalence Theorem for a non-linear, multivariate response model is stated. Section 5 provides the main results; theorems and tables for constructing the locally c-optimal designs. In Section 6 simulation is used to investigate how sensitive the designs are to misspecification of the correlation between the two profiles of the bivariate Emax model. Finally a discussion of results and future work is provided in Section 7.

2 The bivariate Emax model

Understanding of the dose-response relationship is among the most important and challenging problems in drug development. In dose finding studies a primary response variable is usually defined and modeled. Various dose-response models have been suggested. Among the most successful models for modeling a continuous response variable is the Emax model. The model is of the form

$$\text{Response(dose)} = E_0 + E_{max}\frac{\text{dose}}{\text{dose} + ED_{50}} + \epsilon \tag{1}$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$. For a detailed discussion of this model and its mechanistic properties see, e.g., Holford and Sheiner (1981) and Goutelle et al. (2008). A common modification is to assume $E_0 = 0$. This can be done if the placebo effect is known to be zero or if the response variable is placebo adjusted. The model used in this paper is the Emax model including only the two parameters ED_{50} and E_{max}. The problem of deriving optimal designs for the two parameter Emax model has been investigated by several authors see, for example, López-Fidalgo and Wong (2002).

In early phases of the drug development (Phase I) the focus is on safety and the primary response variable is a safety variable. In later phases (Phase II and III) the focus is shifted towards efficacy and the primary response variable is an efficacy variable. In spite of this traditional division into safety and efficacy studies, it is often useful to study efficacy and safety simultaneously. Examples are found in, e.g., Thall and Russell (1998) and Ouellet et al. (2009). Several authors have discussed the construction of optimal designs when simultaneously considering binary efficacy and safety variables see, for example, Li et al. (1995) and Dragalin and Fedorov (2006). Here the focus is on continuous efficacy and safety variables. Dragalin et al. (2008) also discuss the construction of optimal designs when both the efficacy and safety variables are continuous, but both the model and the aim of the study is different from what is assumed here.

The idea in this paper is to consider both efficacy and safety simultaneously and extend the Emax model to two dimensions, one for a primary efficacy variable, Z_1, and one for a primary safety variable, Z_2. High values of Z_1 are here assumed to indicate desirable effect while high values of Z_2 indicate undesirable effect. An example of a primary efficacy variable is the decrease (from baseline) in systolic blood pressure, measured in millimeters of mercury (mmHg). An example of a primary safety variable is increased sleep latency from baseline, measured in minutes. The bivariate Emax model is hereby defined as follows

$$Z_1 = E_{max}\frac{x}{x + ED_{50}} + \varepsilon_1, \qquad Z_2 = S_{max}\frac{x}{x + SD_{50}} + \varepsilon_2 \tag{2}$$

where $(\epsilon_1, \epsilon_2) \sim \mathcal{N}_2(0, \Sigma(\sigma_1, \sigma_2, \rho))$. Further let $\theta = (ED_{50}, E_{max}, SD_{50}, S_{max})$ and $\Omega = (\sigma_1, \sigma_2, \rho)$. In this paper Σ is assumed to be known while θ needs to be estimated. Note that x represents the dose of a drug, so naturally it is assumed that $x \in \chi = [0, \infty[$. E_{max} represents the maximal achievable effect from the drug and S_{max} the maximal realizable side-effect. ED_{50} and SD_{50} represent the doses that give half of the maximal effect and side-effect, respectively. Hence $ED_{50}, SD_{50} > 0$. For explicitness it is also assumed here that $E_{max}, S_{max} > 0$. In this paper a simplified version of the bivariate

Emax model, referred to as the simple bivariate model, is also explored. There it is assumed that the maximal effect and side-effect equal one so that $Z_1 = \frac{x}{x+ED_{50}} + \epsilon_1$, $Z_2 = \frac{x}{x+SD_{50}} + \epsilon_2$ where $(\epsilon_1, \epsilon_2) \sim \mathcal{N}_2(0, \Sigma)$, $\theta = (ED_{50}, SD_{50})$, and $ED_{50}, SD_{50} > 0$ as before.

3 The Clinical Utility Index (CUI)

The bivariate Emax model defined in the previous section models how a primary efficacy variable and a primary safety variable depend on the dose of the drug. The ultimate goal of a dose finding study is to estimate the dose that gives the best possible combination of effects and side-effects. If it were not for side-effects the best dose would be the maximal possible dose because the effect of a drug usually increases monotonically with dose. In practice all drugs have side-effects which also increase with dose. The dose administered to patients needs to be strong enough to give significant effect and mild enough so side-effects will not be severe. In what follows the Clinical Utility Index (CUI) which is a tool for assessing the patient's net benefit from receiving a particular dose of a drug is introduced.

The term and the use of the CUI in dose finding studies is relatively new, but the CUI is an increasingly popular tool for multiattribute decision making in drug development. For a good historical overview of the CUI see, e.g., Carrothers et al. (2011). The CUI combines different aspects regarding the quality of the new drug. After receiving the drug, the patient might experience several different effects and side-effects. These are measured in different scales and are of different importance to the patient. The CUI combines these multidimensional aspects into a single metric. Each possible scenario is given a CUI value for ranking; the higher the CUI the better. A patient taking an anti-diabetic drug might prefer his normal dose to a new, higher, dose that would lead to an 0.5% extra reduction in HbA1c, but with an increased risk for hypoglycemia. The normal dose would then be assigned to a higher CUI value than the new dose. In this paper we only consider one primary efficacy variable, Z_1, and one primary safety variable, Z_2, but it is straightforward to generalize and include more efficacy, safety and possible other variables of importance such as increased cost for a higher dose.

The form and derivation of the CUI should be considered separately for each drug under investigation. The most common approach is however to use a linear combination of the different response variables. This is the approach discussed by Carrothers et al. (2011) and the one that is covered here. For the dose-response models in this paper the CUI is defined as

$$\text{CUI}(x) = k_1 Z_1 - k_2 Z_2. \tag{3}$$

A negative sign is here assigned to the side-effect because high values of Z_2 indicate a negative effect. It should not be forgotten that the response variables are usually measured on different scales. To cope with this Carrothers et al. (2011) explain how all response variables can first be transformed into the same scale, which they call utility, with range from 0 to 1. Then weights, here k_1 and k_2, are assigned to the response variables depending on their relative importance. For a detailed example of how a linear CUI has been used in practice see Ouellet et al. (2009). They use the Sigmoid

Emax model to model how several different efficacy and safety variables depend on dose of drugs for the treatment of insomnia. Then they construct a CUI for decision making. Experts in insomnia disorders defined clinically meaningful differences, c_i, for each response in order to normalize the variables to a common scale. Then a conjoint analysis was used to rank the relative importance of the efficacy and safety variables so weights, v_i could be assigned to the different attributes. The constants in (3) thus represent the weights divided by the clinical meaningful differences, $k_i = \frac{v_i}{c_i}$. The CUI then has the form

$$\text{CUI}(x) = \sum_{\substack{\text{All efficacy} \\ \text{variables}}} \frac{v_i}{c_i} Z_i - \sum_{\substack{\text{All safety} \\ \text{variables}}} \frac{v_i}{c_i} Z_i. \tag{4}$$

Regardless of whether the constants k_1 and k_2 represent weights or weights divided by the clinically meaningful difference, the most desirable dose for a population of patients is the one that maximizes E[CUI]. It is straightforward to show that if such a positive dose exists then, for the bivariate Emax model, it is

$$\begin{aligned} g(\theta) &:= \arg\max_{x>0} \text{E}[\text{CUI}(x)] \\ &= \frac{\sqrt{k_1 ED_{50} E_{max} k_2 SD_{50} S_{max}}(ED_{50} - SD_{50}) - ED_{50} SD_{50}(k_1 E_{max} - k_2 S_{max})}{k_1 ED_{50} E_{max} - k_2 SD_{50} S_{max}}. \end{aligned} \tag{5}$$

For the simple bivariate model this simplifies to

$$g(\theta) := \arg\max_{x>0} \text{E}[\text{CUI}(x)] = \frac{\sqrt{k_1 ED_{50} k_2 SD_{50}}(ED_{50} - SD_{50}) - ED_{50} SD_{50}(k_1 - k_2)}{k_1 ED_{50} - k_2 SD_{50}}. \tag{6}$$

The designs derived in this paper are optimal for estimating these functions i.e. they are optimal for estimating the most desirable dose of a drug for a population of patients, provided that the models and the CUI are reasonable assumptions.

4 Locally c-optimal designs

Optimal design theory deals with designing studies when the response is believed to be related to the explanatory variables through a statistical model. In the case of dose finding studies one needs to decide the number of participants, the study duration, dosing schedules, and so on. Optimal design theory helps in answering the question how big dose each participant in the study should get. A design, ξ, is formally defined as a set of pairs denoted by $\xi = \{x_1, x_2, \ldots, x_n; w_1, w_2, \ldots, w_n\}$. The x_i's are called design points. For dose finding studies they represent the doses i.e. the amount of the drug (in mg) and w_i denotes the proportion of the participants that receive dose x_i. Here n is the number of different study groups and N denotes the number of participants. Finally χ denotes the design space, the set of possible values for the

design points. By constructing a design we mean deciding how many different groups, n, should be included in the study and deciding what the x_i's and the w_i's should be. The design that allocates all observations to one design point, x, is denoted by ξ_x. Designs are said to be optimal if they result in estimates of interest with minimum variance for a given experimental effort, which in the case of dose finding studies is the number of participants, N, in the study. Which design is optimal depends on the statistical model, the design space and the objective of the study. In this paper, the focus is on c-optimal designs which are appropriate when the objective of the study is to estimate some function, $g(\theta)$, of the parameters in the model. Let $\hat{\theta}$ denote the maximum likelihood estimator of the parameters in a non-linear, multivariate response model such as the bivariate Emax model (2). For a linear function $g(\theta) = c^T\theta$, where c is a vector of constants, a design is c-optimal if it minimizes the asymptotic variance of $\sqrt{N}(g(\hat{\theta}) - g(\theta))$ which equals $\Psi = c^T M^{-1}(\xi)c$. Similarly, for a non-linear function, $g(\theta)$, a design is c-optimal if it minimizes

$$\Psi = \nabla g(\theta)^T M(\xi)^{-1} \nabla g(\theta). \tag{7}$$

Ψ is referred to as the criterion function and $M(\xi)$ is the standardized information matrix, $M(\xi) = N^{-1}I(\xi)$. $M(\xi)$ is a symmetric matrix and can be written as

$$M(\xi) = \sum_{i=1}^{n} w_i M(\xi_{x_i}).$$

For the bivariate Emax model

$$M(\xi_x) = \frac{1}{(1-\rho^2)} \begin{pmatrix} \frac{1}{\sigma_1^2}\frac{E_{max}^2 x^2}{(x+ED_{50})^4} & \frac{1}{\sigma_1^2}\frac{-E_{max}x^2}{(x+ED_{50})^3} & \frac{\rho}{\sigma_1\sigma_2}\frac{-E_{max}S_{max}x^2}{(x+ED_{50})^2(x+SD_{50})^2} & \frac{\rho}{\sigma_1\sigma_2}\frac{E_{max}x^2}{(x+ED_{50})^2(x+SD_{50})} \\ \cdot & \frac{1}{\sigma_1^2}\frac{x^2}{(x+ED_{50})^2} & \frac{\rho}{\sigma_1\sigma_2}\frac{S_{max}x^2}{(x+ED_{50})(x+SD_{50})^2} & \frac{\rho}{\sigma_1\sigma_2}\frac{-x^2}{(x+ED_{50})(x+SD_{50})} \\ \cdot & \cdot & \frac{1}{\sigma_2^2}\frac{S_{max}^2 x^2}{(x+SD_{50})^4} & \frac{1}{\sigma_2^2}\frac{-S_{max}x^2}{(x+SD_{50})^3} \\ \cdot & \cdot & \cdot & \frac{1}{\sigma_2^2}\frac{x^2}{(x+SD_{50})^2} \end{pmatrix}$$

and for the simple bivariate model it is

$$M(\xi_x) = \frac{1}{(1-\rho^2)} \begin{pmatrix} \frac{1}{\sigma_1^2}\frac{x^2}{(x+ED_{50})^4} & \frac{\rho}{\sigma_1\sigma_2}\frac{-x^2}{(x+ED_{50})^2(x+SD_{50})^2} \\ \cdot & \frac{1}{\sigma_2^2}\frac{x^2}{(x+SD_{50})^4} \end{pmatrix}.$$

In all that follows it is assumed that M is an invertible matrix.

A general problem with optimal designs for non-linear models is that they depend on the true value of the unknown parameters, θ. The optimal designs constructed in this paper assume prior values for θ. Such designs are called locally optimal designs. For dose finding studies the prior is based on data from preclinical and early clinical trials as well as data from competitor drugs.

The most important tool for verifying that a design is optimal is the General Equivalence Theorem (GET). GET is a synonym for several equivalence theorems used to demonstrate that designs are optimal. Their form depends on the model, the design space and the criterion function. The first GET was discovered for D-optimality by Kiefer & Wolfowitz (1960). The relevant GET for a non-linear, multivariate response model when the interest is on c-optimality is as follows:

Theorem 1. (GET). *Suppose ξ is a design such that $M(\xi)^{-1}$ exists. Then ξ is locally c-optimal with respect to a non-linear function of the model parameters, $g(\theta)$, if and only if,*

$$\nabla g^T M(\xi)^{-1} M(\xi_x) M(\xi)^{-1} \nabla g \;\; \leq \;\; \nabla g^T M(\xi)^{-1} \nabla g, \quad \forall \; x \in \chi. \tag{8}$$

Further, the equal sign holds for $x \in \{x_1, ..., x_n\}$.

The proof for Theorem 1 is given in the Appendix. Note that Ψ in (7) is a special case of the linear criterion, $\Psi = \text{tr}\{AM^{-1}\}$. Fedorov (1972) sets up the framework for optimal designs that minimize the linear criterion in multi-response setting.

5 Designs for estimating the most desirable dose of a drug

The designs derived in this section are optimal for estimating the most desirable dose of a drug, provided that such a positive dose exists and that the models and the CUI, introduced in sections 2 and 3, are reasonable assumptions. More specifically, the designs are locally c-optimal with respect to the functions in (5) and (6) and depend on the parameter vector $(\theta, \Omega, k_1, k_2)$.

5.1 The simple bivariate model

The following theorem is useful for deriving c-optimal designs for the simple bivariate model.

Theorem 2. *Assume that $\xi = \{x_1, \ldots, x_n; w_1, \ldots, w_n\}$ is locally c-optimal for the simple bivariate model with parameters $\theta = (ED_{50}, SD_{50})$, $k = (k_1, k_2)$ and $\Omega = (\sigma_1, \sigma_2, \rho)$ i.e. ξ minimizes $\Psi = \nabla g^T M^{-1} \nabla g$. If $a, b, c \in \mathbf{R} \backslash \{0\}$ then*

1. *ξ is locally c-optimal for the same model, but with one or more of the following changes for the parameters*

   (i) $k = (ak_1, ak_2)$

   (ii) $\Omega = (b\sigma_1, b\sigma_2, \rho)$

2. *$\xi^* = \{cx_1, \ldots, cx_n; w_1, \ldots, w_n\}$ is locally c-optimal for the same model, but with*

   (iii) $\theta = (cED_{50}, cSD_{50})$.

Proof. **(i)** $\nabla g(\theta, ak_1, ak_2) = \nabla g(\theta, k_1, k_2)$ and M does not depend on k. Hence $\Psi(\xi, \theta, ak_1, ak_2) = \Psi(\xi, \theta, k_1, k_2)$.

(ii) $M(\xi, b\sigma_1, b\sigma_2) = \frac{1}{b^2} M(\xi, \sigma_1, \sigma_2)$ and ∇g does not depend on Ω. Hence $\Psi(\xi, \theta, b\sigma_1, b\sigma_2) = \nabla g^T M(\xi, b\sigma_1, b\sigma_2)^{-1} \nabla g = b^2 \nabla g^T M(\xi, \sigma_1, \sigma_2)^{-1} \nabla g = b^2 \Psi(\xi, \theta, \sigma_1, \sigma_2)$.

(iii) $\nabla g(cED_{50}, cSD_{50}) = \nabla g(ED_{50}, SD_{50})$ and $M(\xi^*, cED_{50}, cSD_{50}) = \frac{1}{c^2} M(\xi, ED_{50}, SD_{50})$. Hence $\Psi(\xi^*, cED_{50}, cSD_{50}) = c^2 \Psi(\xi, ED_{50}, SD_{50})$.

□

Theorem 2 implies that without loss of generality 3 out of 7 parameters for the simple bivariate model can be set equal to one. If ξ is for example known to be locally c-optimal for the model when $k = (1, k_2/k_1)$ then Theorem 2 (i) implies that ξ is also locally c-optimal for the model when $k = (k_1, k_2)$. Below is a corollary that can be used to derive locally c-optimal designs for the simple bivariate model when $\rho = 0$, $k_1 = k_2$, $\sigma_1^2 = \sigma_2^2$ and the proportion SD_{50}/ED_{50} is close to one. The proof is provided in the Appendix. Note that if $k_1 = k_2$ then the restriction $SD_{50} > ED_{50}$ is necessary, else the side-effects would always outweigh the effects.

Corollary 1. *Let $\sigma_1^2 = \sigma_2^2$, $k_1 = k_2$ and $\rho = 0$ then, for the simple bivariate model,*

(i) $\xi = \{\sqrt{ED_{50}SD_{50}}; 1\}$ *is locally c-optimal when* $\frac{SD_{50}}{ED_{50}} \in \left]1, \frac{5+\sqrt{21}}{2}\right]$.

(ii) $\xi = \{ED_{50}, SD_{50}; 0.5, 0.5\}$ *is locally c-optimal when* $\frac{SD_{50}}{ED_{50}} \to \infty$.

Some more locally c-optimal designs for the simple bivariate model are provided in Table 1 where both the correlation, ρ, and the proportion SD_{50}/ED_{50} are allowed to vary. Corollary 1 and Table 1 indicate that when SD_{50}/ED_{50} is close to one, then the one point design $\xi = \{\sqrt{ED_{50}SD_{50}}; 1\}$ is locally c-optimal. Else two design points, x_1 and x_2 are needed. These two points have equal weights and on logarithmic scale they are symmetrically located around $\sqrt{ED_{50}SD_{50}}$. Table 1 further indicates that $x_1x_2 = ED_{50}SD_{50}$ and Corollary 1 states that as $SD_{50}/ED_{50} \to \infty$ these points tend to ED_{50} and SD_{50}. Table 1 also shows that ρ influences how close to one SD_{50}/ED_{50} needs to be in order to have a one point locally c-optimal design. Finally note that each design point is associated with two responses, one for efficacy and one for safety. Since the simple bivariate model only has two model parameters this means that $M(\xi)^{-1}$ exists for all designs, ξ.

Table 1: Locally c-optimal designs $\xi = \{x_1, x_2; w_1, 1 - w_1\}$ for the simple bivariate model with respect to $g(\theta)$ in (6). Here $\chi = [0, \infty]$, $ED_{50} = 1$, $k_1 = k_2$ and $\sigma_1^2 = \sigma_2^2$. If the design point x_2 is not specified then x_1 is the only design point.

SD_{50}		ρ -0.9	-0.5	-0.1	0	0.1	0.5	0.9
6	x_1	$\sqrt{6}$	$\sqrt{6}$	$\sqrt{6}$	$\sqrt{6}$	$\sqrt{6}$	0.9863	0.6214
	w_1	1	1	1	1	1	0.5	0.5
	x_2						6.0832	9.6552
7	x_1	$\sqrt{7}$	$\sqrt{7}$	$\sqrt{7}$	2.1308	1.6109	0.9349	0.6412
	w_1	1	1	1	0.5	0.5	0.5	0.5
	x_2				3.2851	4.3455	7.4876	10.9169
8	x_1	$\sqrt{8}$	$\sqrt{8}$	1.9212	1.5879	1.3763	0.9093	0.6594
	w_1	1	1	0.5	0.5	0.5	0.5	0.5
	x_2			4.1640	5.0382	5.8126	8.7978	12.1318
9	x_1	$\sqrt{9}$	$\sqrt{9}$	1.5933	1.4071	1.2641	0.8958	0.6760
	w_1	1	1	0.5	0.5	0.5	0.5	0.5
	x_2			5.6488	6.3960	7.1199	10.0473	13.3137
10	x_1	$\sqrt{10}$	$\sqrt{10}$	1.4420	1.3077	1.1972	0.8885	0.6910
	w_1	1	1	0.5	0.5	0.5	0.5	0.5
	x_2			6.9346	7.6467	8.3530	11.2543	14.4707

5.2 The bivariate Emax model

The following theorem is useful for deriving c-optimal designs for the bivariate Emax model.

Theorem 3. *Assume that* $\xi = \{x_1, \ldots, x_n; w_1, \ldots, w_n\}$ *is locally c-optimal for the bivariate Emax model with parameters* $\theta = (ED_{50}, E_{max}, SD_{50}, S_{max})$ *,* $k = (k_1, k_2)$ *and* $\Omega = (\sigma_1, \sigma_2, \rho)$ *i.e.* ξ *minimizes* $\Psi = \nabla g^T M^{-1} \nabla g$*. If* $a, b, c, d \in \mathbf{R} \backslash \{0\}$ *then*

1. ξ *is locally c-optimal for the same model, but with one or more of the following changes for the parameters*

   (i) $k = (ak_1, ak_2)$

   (ii) $\Omega = (b\sigma_1, b\sigma_2, \rho)$

   (iii) $\theta = (ED_{50}, cE_{max}, SD_{50}, cS_{max})$

2. $\xi^* = \{dx_1, \ldots, dx_n; w_1, \ldots, w_n\}$ *is locally c-optimal for the same model, but with*

   (iv) $\theta = (dED_{50}, E_{max}, dSD_{50}, S_{max})$.

The proof of (i) and (ii) is identical to the proof of Theorem 2 (i) and (ii). The proof of (iii) and (iv) is provided in the Appendix. Theorem 3 implies that without loss of generality 4 out of 9 parameters for the bivariate Emax model can be set equal to one. If ξ is for example known to be locally c-optimal for the model with parameters $\theta = (ED_{50}, 1, SD_{50}, S_{max}/E_{max})$ then Theorem 3 (iii) gives that ξ is also locally c-optimal for the model with parameters $\theta = (ED_{50}, E_{max}, SD_{50}, S_{max})$. For the bivariate Emax model the design space, χ, needs to be restricted or some design weight is assigned to an infinitely high dose. Some locally c-optimal designs for the design space $\chi = [0, 500]$ are provided in Tables 2 and 3. These designs are derived by informative guessing and numerical minimization of Ψ in (7). It is easy to use Theorem 1 to verify that the designs in Table 2 and 3 are indeed locally c-optimal.

Table 2: Locally c-optimal designs $\xi = \{x_1, 500; w_1, 1 - w_1\}$ for the bivariate Emax model with respect to $g(\theta)$ in (5). Here $\chi = [0, 500]$, $ED_{50} = 1$, $SD_{50} = 2$, $\sigma_1^2 = \sigma_2^2$ and $\rho = 0$.

$\frac{S_{max}}{E_{max}}$ $(\frac{k_2}{k_1} = 1)$	1	0.9	0.8	0.7	0.6
x_1	1.1078	0.7436	0.8234	0.9701	1.1030
w_1	0.3944	0.4074	0.4713	0.5285	0.5720
$\frac{k_2}{k_1}$ $(\frac{S_{max}}{E_{max}} = 1)$	1	0.9	0.8	0.7	0.6
x_1	1.1078	0.7374	0.7871	0.8719	0.9423
w_1	0.3944	0.4157	0.4845	0.5387	0.5774

Table 3: Locally c-optimal designs $\xi = \{x_1, x_2, 500; w_1, w_2, 1-w_1-w_2\}$ for the bivariate Emax model with respect to $g(\theta)$ in (5). Here $\chi = [0, 500]$, $ED_{50} = 1$, $k_1 = k_2$ and $E_{max} = S_{max}$. If the design point x_2 is not specified then x_1 and 500 are the only design points.

SD_{50}		ρ ($\frac{\sigma_2^2}{\sigma_1^2} = 1$)					$\frac{\sigma_2^2}{\sigma_1^2}$ ($\rho = 0$)		
		-0.9	-0.5	0	0.5	0.9	0.5	1.5	3
2	x_1	1.0188	1.0793	1.1078	1.0972	0.4985	0.8437	1.3875	2.1699
	w_1	0.1381	0.2800	0.3944	0.5118	0.4883	0.3702	0.4160	0.4670
	x_2					4.9811			
	w_2					0.3639			
3	x_1	1.2978	1.2704	1.2833	0.8407	0.4626	0.9356	1.6057	1.7091
	w_1	0.1987	0.3723	0.4903	0.4736	0.5111	0.4555	0.5197	0.3929
	x_2				6.0010	5.6854			6.3070
	w_2				0.2220	0.4213			0.2684
4	x_1	1.5396	1.4553	1.1358	0.7026	0.4815	1.0312	1.1128	1.0766
	w_1	0.2327	0.4188	0.4582	0.4704	0.5197	0.4956	0.4124	0.3433
	x_2			6.5538	7.6651	6.3468		7.2480	8.9210
	w_2			0.1282	0.3411	0.4305		0.2700	0.5557
5	x_1	1.4473	1.3371	0.9347	0.6662	0.5040	0.9375	0.9327	0.9287
	w_1	0.2211	0.3933	0.4353	0.4767	0.5232	0.4831	0.4016	0.3420
	x_2	6.9420	7.2388	8.7514	8.6498	7.0101	8.5842	8.8849	9.1612
	w_2	0.0538	0.0889	0.2548	0.3853	0.4343	0.0786	0.3833	0.6213

Table 3 indicates that two design points are sufficient as long as $1 < SD_{50}/ED_{50} < c$, where c is some constant. Else, if $SD_{50}/ED_{50} > c$ then an additional design point is needed. Table 3 also indicates how the variance-covariance parameters influence the designs. The larger ρ and σ_2^2/σ_1^2 become the smaller is c for which the statement above holds. A study where the two responses are negatively correlated (for a fixed dose) needs, in this setting, equal or fewer design points than if the two responses are positively correlated. Moreover a study where the variance for the primary safety variable is large compared to the variance for the primary efficacy variable needs, in this setting, equal or more design points than if this were the other way around. Finally note that the designs provided in Table 3 have at least two design points. Each design point is associated with two responses and thus the designs in Table 3 result in nonsingular $M(\xi)$.

6 Covariance structure

In the previous section we saw how ρ influences the locally c-optimal designs. In this section simulation is used to investigate how sensitive the designs are to misspecification of the parameter ρ. The simulation is carried out as follows. Assume that there are $N = 10000$ patients participating in a study for a new investigational drug. Assume further that the bivariate Emax model with parameters $\theta = (ED_{50}, E_{max}, SD_{50}, S_{max}) = (1, 1, 3, 1)$, $\sigma_1 = \sigma_2 = 0.1$ and $\rho = 0$ is appropriate. Data is simulated from this model by using the locally c-optimal design provided in Table 3 that allocates 4930 patients to 1.2833 ml of the drug and the rest to the maximal dose, 500 ml. Figure 1 depicts the simulated data for the first 20 patients in each group. The Gauss-Newton algorithm is then used to fit the simulated data to the model with undefined parameters. Note

that the efficacy and safety profiles are estimated independent of each other. The result is used to estimate the most desirable dose, $g(\theta)$ defined in (5). By repeating this procedure 1000 times a simulated estimate for the standard deviation of $g(\hat{\theta})$ is obtained. The result can be seen in Table 4. The table also gives estimates derived analogously by using the design explained above when the true ρ differs and estimates derived by using the correct c-optimal designs from Table 3 for several different values of ρ. A graphical representation can be seen in Figure 2. The impact from using the locally c-optimal design based on the prior $\rho = 0$ rather than the correct ρ can be seen by comparing the confidence intervals. As expected using the correct locally c-optimal design results in more exact estimates than when using designs based on misspecified priors for ρ. The sizes of the confidence intervals are on the other hand not very different so in practical applications it might be reasonable to ignore the correlation when constructing the locally c-optimal designs.

Figure 1: The mean response curves for the bivariate Emax model with parameters $\theta = (ED_{50}, E_{max}, SD_{50}, S_{max}) = (1, 1, 3, 1)$, $\sigma_1 = \sigma_2 = 0.1$ and $\rho = 0$ and the relevant CUI curve with $k_1 = k_2 = 1$. The locally c-optimal design for this model was used to simulate observations for 40 patients (each having response both for efficacy and safety). The figure to the left shows observations for group 1 while the figure to the right shows observations for group 2.

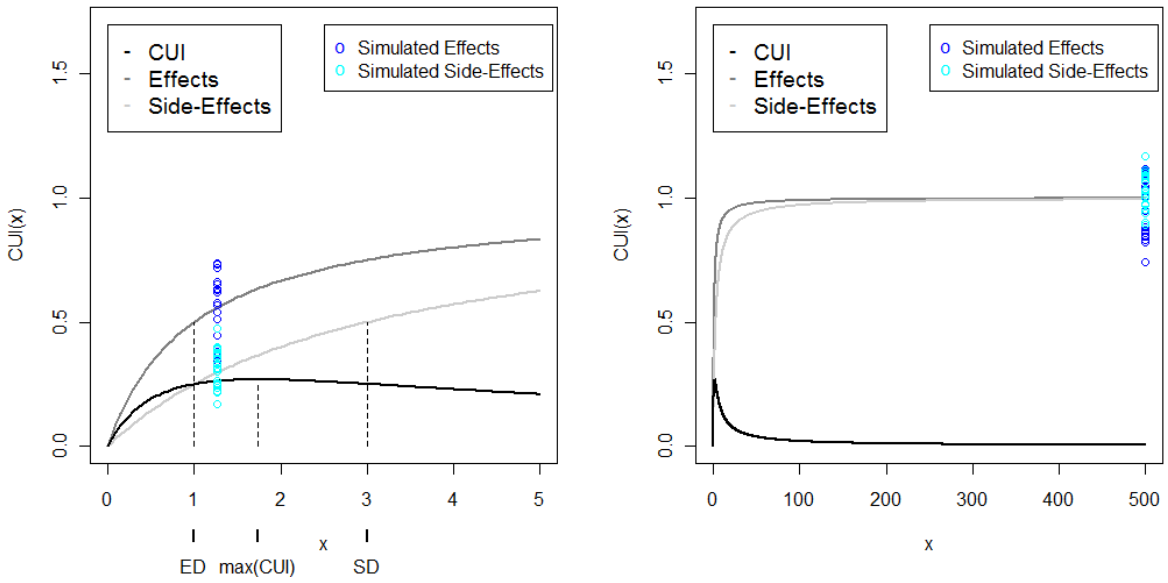
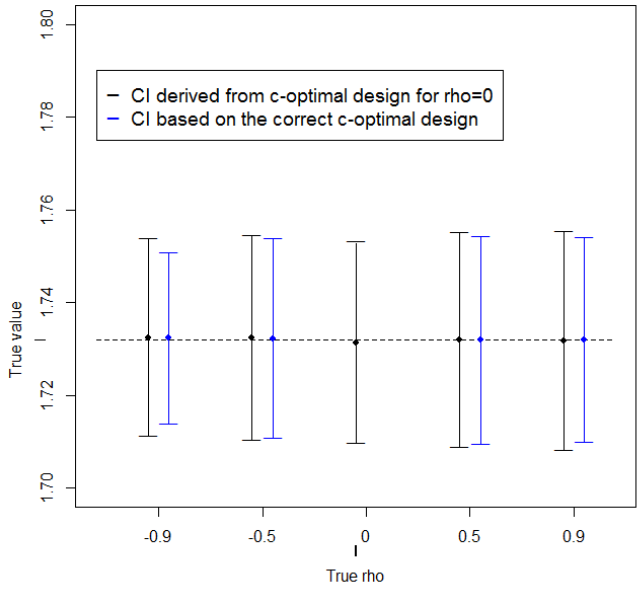


Table 4: Simulated estimate of the most desirable dose, $g(\theta)$, in (5) and it's confidence interval. Here it is assumed that $\sigma_1 = \sigma_2 = 0.1$ and the true model parameters are assumed to be $\theta = (ED_{50}, E_{max}, SD_{50}, S_{max}) = (1, 1, 3, 1)$ and hence $g_{\text{true}}(\theta) = 1.7321$. The parameter ρ is varied and the impact of using the locally c-optimal design based on the prior $\rho = 0$ rather than the correct one is investigated. The designs used for the simulation are the locally c-optimal designs from Table 3.

True $\boldsymbol{\rho}$	**Design used for simulation**	$g(\hat{\boldsymbol{\theta}})$	CI$(g(\hat{\boldsymbol{\theta}}))$
-0.9	c-opt when $\rho = -0.9$ (2 point)	1.7319	$[1.7142, 1.7495]$
-0.9	c-opt when $\rho = 0$ (2 point)	1.7325	$[1.7115, 1.7535]$
-0.5	c-opt when $\rho = -0.5$ (2 point)	1.7316	$[1.7108, 1.7523]$
-0.5	c-opt when $\rho = 0$ (2 point)	1.7321	$[1.7102, 1.7540]$
0	c-opt when $\rho = 0$ (2 point)	1.7322	$[1.7106, 1.7538]$
0.5	c-opt when $\rho = 0.5$ (3 point)	1.7322	$[1.7145, 1.7498]$
0.5	c-opt when $\rho = 0$ (2 point)	1.7320	$[1.7092, 1.7548]$
0.9	c-opt when $\rho = 0.9$ (3 point)	1.7328	$[1.7151, 1.7505]$
0.9	c-opt when $\rho = 0$ (2 point)	1.7322	$[1.7085, 1.7559]$

Figure 2: The simulated estimates of the best dose, $g(\theta)$ and their confidence intervals from Table 4.



7 Discussion

The present paper explores the construction of locally c-optimal designs for nonlinear, bivariate models. The relevant GET is provided with a proof. Particular interest is on the bivariate Emax model and the simple bivariate model and the search for the most desirable dose of a drug. Numerical minimization and the GET are used to derive some locally c-optimal designs for these models. When the correlation ρ and the proportions σ_2^2/σ_1^2, SD_{50}/ED_{50} are sufficiently small one/two point designs are locally c-optimal for the simple bivariate/bivariate Emax model model. Else an extra design point is needed. The simulation carried out in Section 6 indicates on the other hand that the locally c-optimal designs need not be sensitive to misspecification of ρ. Note that in this paper the variance-covariance parameters are assumed to be known. It is of interest to analyze further the impact of these parameters and the uncertainty that arises when they are unknown.

Appendix

Lemma 1. *Let $\phi(x,\xi)$ stand for the derivative of Ψ in the direction ξ_x and let Ψ be a general criterion function to be minimized. A design, ξ, is locally optimal with respect to Ψ if and only if $\phi(x,\xi) \geq 0 \ \forall x \in \chi$. This further implies that $\phi(x,\xi) = 0$ for $x \in \{x_1, ..., x_n\}$.*

Proof. See for example, Pázman (1986). □

Theorem 1. (GET).

Proof. A design is locally c-optimal with respect to a non-linear function, $g(\theta)$, if it minimizes Ψ as defined in (7). First note that the directional derivative can be written of the form $\phi(x,\xi) = tr\Big(\frac{\partial\Psi}{\partial M(\xi)}(M(\xi_x) - M(\xi))\Big)$ (see Pázman 1986). Now

$$\begin{aligned}\frac{\partial\Psi}{\partial M(\xi)} &= \frac{\partial}{\partial M(\xi)}(\nabla g(\theta)^T M(\xi)^{-1}\nabla g(\theta)) \\ &= -(M(\xi)^T)^{-1}\nabla g(\theta)\nabla g(\theta)^T(M(\xi)^T)^{-1} \\ &= -M(\xi)^{-1}\nabla g(\theta)\nabla g(\theta)^T M(\xi)^{-1}.\end{aligned}$$

From the above we get

$$\begin{aligned}\phi(x,\xi) &= tr(-M(\xi)^{-1}\nabla g(\theta)\nabla g(\theta)^T M(\xi)^{-1}(M(\xi_x) - M(\xi))) \\ &= tr(M(\xi)^{-1}\nabla g(\theta)\nabla g(\theta)^T) - tr(M(\xi)^{-1}\nabla g(\theta)\nabla g(\theta)^T M(\xi)^{-1}M(\xi_x)) \\ &= tr(\nabla g(\theta)^T M(\xi)^{-1}\nabla g(\theta)) - tr(\nabla g(\theta)^T M(\xi)^{-1}M(\xi_x)M(\xi)^{-1}\nabla g(\theta)) \\ &= \nabla g(\theta)^T M(\xi)^{-1}\nabla g(\theta) - \nabla g(\theta)^T M(\xi)^{-1}M(\xi_x)M(\xi)^{-1}\nabla g(\theta).\end{aligned}$$

Now

$$\phi(x,\xi) \geq 0 \ \Leftrightarrow \ \nabla g(\theta)^T M(\xi)^{-1}M(\xi_x)M(\xi)^{-1}\nabla g(\theta) \leq \nabla g(\theta)^T M(\xi)^{-1}\nabla g(\theta).$$

□

Corollary 1.

Proof. (i) Let $s := \frac{SD_{50}}{ED_{50}} > 1$, $\theta_s = (1, s)$ and $\xi_s = \{\sqrt{s}; 1\}$. By working out both sides of (8) Theorem 1 gives that ξ_2 is locally c-optimal if and only if $f(x) := x^2\Big(\frac{1}{s(x+1)^4} + \frac{s}{(x+s)^4}\Big) \leq \frac{2}{(\sqrt{s}+1)^4} \ \forall x \in \chi$. The equal sign holds when $x = \sqrt{s}$ and one can show that this point is a local maximum given that $s \in \Big]1, \frac{7+3\sqrt{5}}{2}\Big[$. ($f'(\sqrt{s}) = 0 \ \forall s \in \chi$ and $f''(\sqrt{s}) < 0 \Leftrightarrow s \in \Big]\frac{7-3\sqrt{5}}{2}, \frac{7+3\sqrt{5}}{2}\Big[$). Now $f'(x) = 2x(x - \sqrt{s})(x + \sqrt{s})g(x)/((x + s)^5(x + 1)^5)$ where g(x) is a polynomial of degree 4. If $s \in \Big]1, \frac{5+\sqrt{21}}{2}\Big]$ then all coefficients of $g(x)$ are negative and hence $x = \sqrt{s}$ is a global maximum. This gives that ξ_s is locally c-optimal for this interval and thus $\xi = \{ED_{50}\sqrt{s}; 1\} = \{\sqrt{ED_{50}SD_{50}}; 1\}$ is locally c-optimal for the model with $\theta = ED_{50}\theta_s{=}(ED_{50}, SD_{50})$ given that $\frac{SD_{50}}{ED_{50}} \in \Big]1, \frac{5+\sqrt{21}}{2}\Big]$. (ii) Let $s := \frac{SD_{50}}{ED_{50}}$, $\theta_s = (1, s)$ and $\xi_s = \{1, s; 0.5, 0.5\}$. By working out both sides of (8) Theorem 1 gives that ξ_2 is locally c-optimal if and only if

$$\Big(\frac{32(s+1)^4}{16s^2 + (s+1)^4}\Big)^2\frac{sx^2}{4}\Big(\frac{1}{(x+1)^4} + \frac{s^2}{(x+s)^4}\Big) \leq \frac{16s(s+1)^4}{16s^2 + (s+1)^4} \quad \forall \ x \in \chi,$$

which is equivalent to

$$\frac{16(s+1)^4}{16s^2+(s+1)^4}x^2\Big(\frac{1}{(x+1)^4}+\frac{s^2}{(x+s)^4}\Big)\leq 1 \quad \forall\ x\in\chi. \tag{9}$$

The left hand side of (9) goes to $16x^2\frac{1}{(x+1)^4}$ as $s\to\infty$. Finally $16x^2\frac{1}{(x+1)^4}\leq 1\forall x\in\chi$. This gives that ξ_s is locally c-optimal and thus $\xi=\{ED_{50},ED_{50}s;0.5,0.5\}=\{ED_{50},SD_{50};0.5,0.5\}$ is locally c-optimal for the model with $\theta=(ED_{50},ED_{50}s)=(ED_{50},SD_{50})$ as $\frac{SD_{50}}{ED_{50}}\to\infty$.

□

Theorem 3.

Proof. We start by introducing a shorter notation for the standardized information matrix for the bivariate Emax model provided in Section 4. Let

$$M(\xi)=\begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14}\\ . & b_{22} & b_{23} & b_{24}\\ . & . & b_{33} & b_{34}\\ . & . & . & b_{44}\end{pmatrix}.$$

(iii) First note that $\nabla g(ED_{50},cE_{max},SD_{50},cS_{max})^T=(g_1,\frac{1}{c}g_2,g_3,\frac{1}{c}g_4)$ and hence

$$\begin{aligned}
\Psi(\xi,ED_{50},cE_{max},SD_{50},cS_{max}) &= \begin{pmatrix} g_1 & \frac{1}{c}g_2 & g_3 & \frac{1}{c}g_4\end{pmatrix}\begin{pmatrix} c^2b_{11} & cb_{12} & c^2b_{13} & cb_{14}\\ . & b_{22} & cb_{23} & b_{24}\\ . & . & c^2b_{33} & cb_{34}\\ . & . & . & b_{44}\end{pmatrix}^{-1}\begin{pmatrix} g_1\\ \frac{1}{c}g_2\\ g_3\\ \frac{1}{c}g_4\end{pmatrix}\\
&= \begin{pmatrix} g_1 & \frac{1}{c}g_2 & g_3 & \frac{1}{c}g_4\end{pmatrix}\Bigg(\begin{pmatrix} c & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & c & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14}\\ . & b_{22} & b_{23} & b_{24}\\ . & . & b_{33} & b_{34}\\ . & . & . & b_{44}\end{pmatrix}\begin{pmatrix} c & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & c & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\Bigg)^{-1}\begin{pmatrix} g_1\\ \frac{1}{c}g_2\\ g_3\\ \frac{1}{c}g_4\end{pmatrix}\\
&= \begin{pmatrix} g_1 & \frac{1}{c}g_2 & g_3 & \frac{1}{c}g_4\end{pmatrix}\begin{pmatrix} 1/c & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1/c & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14}\\ . & b_{22} & b_{23} & b_{24}\\ . & . & b_{33} & b_{34}\\ . & . & . & b_{44}\end{pmatrix}^{-1}\begin{pmatrix} 1/c & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1/c & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} g_1\\ \frac{1}{c}g_2\\ g_3\\ \frac{1}{c}g_4\end{pmatrix}\\
&= \frac{1}{c^2}\Psi(\xi,ED_{50},E_{max},SD_{50},S_{max}).
\end{aligned}$$

(iv) First note that $\nabla g(ED_{50},cE_{max},SD_{50},cS_{max})^T=(g_1,dg_2,g_3,dg_4)$ and hence

$$\begin{aligned}
\Psi(\xi^*,dED_{50},E_{max},dSD_{50},S_{max}) &= \begin{pmatrix} g_1 & dg_2 & g_3 & dg_4\end{pmatrix}\begin{pmatrix} b_{11}/d^2 & b_{12}/d & b_{13}/d^2 & b_{14}/d\\ . & b_{22} & b_{23}/d & b_{24}\\ . & . & b_{33}/d^2 & b_{34}/d\\ . & . & . & b_{44}\end{pmatrix}^{-1}\begin{pmatrix} g_1\\ dg_2\\ g_3\\ dg_4\end{pmatrix}\\
&= \begin{pmatrix} g_1 & dg_2 & g_3 & dg_4\end{pmatrix}\Bigg(\begin{pmatrix} 1/d & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1/d & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14}\\ . & b_{22} & b_{23} & b_{24}\\ . & . & b_{33} & b_{34}\\ . & . & . & b_{44}\end{pmatrix}\begin{pmatrix} 1/d & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1/d & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\Bigg)^{-1}\begin{pmatrix} g_1\\ dg_2\\ g_3\\ dg_4\end{pmatrix}\\
&= \begin{pmatrix} g_1 & dg_2 & g_3 & dg_4\end{pmatrix}\begin{pmatrix} d & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & d & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14}\\ . & b_{22} & b_{23} & b_{24}\\ . & . & b_{33} & b_{34}\\ . & . & . & b_{44}\end{pmatrix}^{-1}\begin{pmatrix} d & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & d & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix} g_1\\ dg_2\\ g_3\\ dg_4\end{pmatrix}\\
&= d^2\Psi(\xi,ED_{50},E_{max},SD_{50},S_{max}).
\end{aligned}$$

□

References

[1] Carrothers, T.J., Hodge, F.L., Korsan, R.J., Poland, W.B., and Dykstra, K.H.: Decision-making in drug development: Application of a Clinical Utility Index. *American Association of Pharmaceutical Scientists* 2011; **1**: 85-107.

[2] Dragalin, V. and Fedorov, V.: Adaptive designs for dose-finding based on efficacy-toxicity response. *Journal of Statistical Planning and Inference* 2006; **136**: 1800-1823.

[3] Dragalin, V., Fedorov, V., and Wu, Y.: Two-stage design for dose-finding that accounts for both efficacy and safety. *Statistics in Medicine* 2008; **27**: 5156-5176.

[4] Fedorov, V.: *Theory of Optimal Experiments*. Academic Press, New York 1972.

[5] Goutelle, S., Maurin, M., Rougier, F., Barbaut, X., Bourguignon, L., Ducher, M., and Maire, P.: The Hill equation: a review of its capabilities in pharmacological modelling. *Fundamental & Clinical Pharmacology* 2008; **22**: 633-648.

[6] Holford, N.H.G. and Sheiner, L.B.: Understanding the dose-effect relationship: Clinical application of pharmacokinetic-pharmacodynamic models. *Clinical Pharmacokinetics* 1981; **6**: 429-453.

[7] Kiefer, J. and Wolfowitz, J.: The equivalence of two extremum problems. *Canadian Journal of Mathematics* 1960; **12**: 363-366

[8] Li, Z., Durham, S.D., and Flournoy, N.: An Adaptive Design for Maximization of a Contingent Binary Response. *Adaptive Designs. IMS Lecture Notes-Monograph Series*. Flournoy, N., Rosenberger, W.F. (eds.). Hayward, CA: Institute of Mathematical Statistics 1995; **25**: 179-196.

[9] López-Fidalgo, J. and Wong, W.K.: Design Issues for the Michaelis-Menten Model. *Journal of Theoretical Biology* 2002; **215**: 1-11.

[10] Ouellet, D., Werth, J., Parekh, N., Feltner, D., McCarthy, B., and Lalonde, R.L.: The use of a Clinical Utility Index to Compare Insomnia Compounds: A quantitative Basis for Benefit-Risk Assessment. *Clinical pharmacology & Therapeutics* 2009; **85**: 277-282.

[11] Pázman, A.: *Foundations of Optimum Experimental Design*. D. Reidel Publishing Company, Dodrecht 1986.

[12] Thall, P.F. and Russell, K.E.: A Strategy for Dose-Finding and Safety Monitoring Based on Efficacy and Adverse Outcomes in Phase I/II Clinical Trials. *Biemetrics* 1998; **54**: 251-264.