Title

Optimal designs for active controlled dose-finding trials with efficacy-toxicity outcomes

Journal

Biometrika, 104(4)

ISSN

0006-3444

Authors

Schorning, K
Dette, H
Kettelhake, K
et al.

Publication Date

2017-12-01

DOI

10.1093/biomet/asx057

Peer reviewed

Biometrika (2017), **104**, 4, *pp.* 1003–1010
Printed in Great Britain
doi: 10.1093/biomet/asx057
Advance Access publication 9 October 2017

Optimal designs for active controlled dose-finding trials with efficacy-toxicity outcomes

By K. SCHORNING, H. DETTE, K. KETTELHAKE

Fakultät für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany
Kirsten.Schorning@rub.de holger.dette@rub.de Katrin.Kettelhake@rub.de

W. K. WONG

Department of Biostatistics, University of California at Los Angeles,
650 Charles E. Young Dr. South, Los Angeles, California 90095, U.S.A.
wkwong@ucla.edu

and F. BRETZ

Statistical Methodology, Novartis Pharma AG, 4002 Basel, Switzerland
frank.bretz@novartis.com

Summary

We derive optimal designs to estimate efficacy and toxicity in active controlled dose-finding trials when the bivariate continuous outcomes are described using nonlinear regression models. We determine upper bounds on the required number of different doses and provide conditions under which the boundary points of the design space are included in the optimal design. We provide an analytical description of minimally supported optimal designs and show that they do not depend on the correlation between the bivariate outcomes.

Some key words: Admissible design; Equivalence theorem; Particle swarm optimization; Tchebycheff system.

1. Introduction

Most literature on optimal design of experiments concerns univariate outcomes. In practice, however, experiments are often conducted to measure multiple outcomes that may be correlated. For instance, pharmaceutical dose-finding trials invariably measure bivariate outcomes involving efficacy and toxicity. Heise & Myers (1996) used the Gumbel bivariate binary quantal response model and Magnusdottir (2013) applied c-optimal designs to the bivariate Emax model to study efficacy and toxicity. Similarly, Fan & Chaloner (2004) proposed using a continuation ratio model for a trinomial outcome, where the outcome of a patient is classified as no reaction when neither toxicity nor efficacy occurs, efficacy for efficacy without toxicity, and adverse reaction for toxicity. Adaptive dose-finding trials incorporating both efficacy and safety have also been investigated; see for example, Dragalin et al. (2008).

Recently, the use of active controls instead of placebos in dose-finding trials has received considerable attention (Temple & Ellenberg, 2000; Splawinski & Kuzniar, 2004; Helms et al., 2015). Dette et al. (2014, 2015) discussed design issues for such trials with univariate outcomes. To the best of our knowledge, however, the problem of determining optimal designs for active controlled trials with bivariate mean outcomes has not yet been investigated. In this paper we address this problem in two steps. In § 3 we provide new results about optimal designs for various models with efficacy and toxicity outcomes without an active control. In particular we derive new upper bounds on the required number of doses and determine

analytically the optimal designs with the minimum number of doses. Secondly, we show in § 4 how to obtain locally optimal designs for active controlled dose-finding trials from the preceding results for uncontrolled trials and demonstrate our approach in an example.

2. OPTIMAL DESIGNS FOR BIVARIATE OUTCOMES

Given a statistical model defined on a dose interval $\mathcal{D} = [L, R] \subset \mathbb{R}_0^+$, the design problem is to determine for a given design criterion the optimal number of doses, k, the dose levels $d_1, \ldots, d_k \in \mathcal{D}$, and the number of patients assigned to each dose. That is, for a given total sample size n_1, the optimal design needs to specify the number of patients n_{1i} at each dose level d_i subject to $\sum_{i=1}^{k} n_{1i} = n_1$. We use the index 1 here in the notation, i.e., $n_1, n_{1i}, \ldots$, since in later sections we use the index 2 for active controlled clinical trials.

Let Y_{ij} be the two-dimensional outcome variable at dose level d_i from subject j and assume that

$$Y_{ij} = (Y_{ij}^e, Y_{ij}^t)^{\mathrm{T}} \sim \mathcal{N}_2\{\eta_1(d_i, \theta_1), \Sigma_1\} \qquad (j = 1, \ldots, n_{1i}; i = 1, \ldots, k). \tag{1}$$

The mean response $\eta_1(d, \theta_1) = \{\eta_1^e(d, \theta_1^e), \eta_1^t(d, \theta_1^t)\}^{\mathrm{T}} \in \mathbb{R}^2$ describes the expected efficacy (η_1^e) and toxicity (η_1^t) at dose level $d \in \mathcal{D}$, where the $(s_1^e + 1)$- and $(s_1^t + 1)$-dimensional vectors θ_1^e and θ_1^t define the parameters in the models η_1^e and η_1^t, respectively. The parameter $\theta_1 = \{(\theta_1^e)^{\mathrm{T}}, (\theta_1^t)^{\mathrm{T}}\}^{\mathrm{T}}$ varies in a compact parameter space, say $\Theta_1 \subset \mathbb{R}^{s_1}$, where $s_1 = s_1^e + s_1^t + 2$. The unknown covariance matrix is

$$\Sigma_1 = \mathrm{cov}(Y) = \begin{pmatrix} \sigma_e^2 & \rho\sigma_e\sigma_t \\ \rho\sigma_e\sigma_t & \sigma_t^2 \end{pmatrix},$$

where $-1 < \rho < 1$ denotes the correlation between the two outcome variables and the variances of the random variables Y_{ij}^e and Y_{ij}^t are given by σ_e^2 and σ_t^2, respectively. The variables $Y_{11}, \ldots, Y_{kn_{1k}}$ are assumed to be independent.

We further assume that η_1 is continuously differentiable with respect to the parameter θ_1 and denote, respectively, the gradients of the two mean responses with respect to θ_1^e and θ_1^t by

$$f_e(d, \theta_1^e) = \frac{\partial}{\partial \theta_1^e} \eta_1^e(d, \theta_1^e) = \{f_0^e(d), \ldots, f_{s_1^e}^e(d)\}^{\mathrm{T}}, \quad f_t(d, \theta_1^t) = \frac{\partial}{\partial \theta_1^t} \eta_1^t(d, \theta_1^t) = \{f_0^t(d), \ldots, f_{s_1^t}^t(d)\}^{\mathrm{T}}.$$

Throughout, let 0_q denote the q-dimensional vector with all entries equal to 0; sometimes the subscript will be omitted for simplicity. The Fisher information matrix is given by the $s_1 \times s_1$ matrix

$$\begin{aligned} \mathcal{I}_1(d, \theta_1) &= \left\{ \frac{\partial}{\partial \theta_1} \eta_1(d, \theta_1) \Big|_{\theta_1 = \theta_1} \right\}^{\mathrm{T}} \Sigma_1^{-1} \left\{ \frac{\partial}{\partial \theta_1} \eta_1(d, \theta_1) \Big|_{\theta_1 = \theta_1} \right\} = \begin{pmatrix} f_e(d) & 0_{s_1^e+1} \\ 0_{s_1^t+1} & f_t(d) \end{pmatrix} \Sigma_1^{-1} \begin{pmatrix} f_e^{\mathrm{T}}(d) & 0_{s_1^t+1}^{\mathrm{T}} \\ 0_{s_1^e+1}^{\mathrm{T}} & f_t^{\mathrm{T}}(d) \end{pmatrix} \\ &= \frac{1}{\sigma_e^2 \sigma_t^2 (1 - \rho^2)} F(d), \end{aligned}$$

where

$$F(d) = \begin{pmatrix} \sigma_t^2 \mathcal{F}_1 & -\rho\sigma_e\sigma_t \mathcal{F}_2 \\ -\rho\sigma_e\sigma_t \mathcal{F}_2^{\mathrm{T}} & \sigma_e^2 \mathcal{F}_3 \end{pmatrix} \tag{2}$$

and the blocks in the matrix are defined by $\mathcal{F}_1 = f_e(d, \theta_1^e) f_e^{\mathrm{T}}(d, \theta_1^e) \in \mathbb{R}^{(s_1^e+1)\times(s_1^e+1)}$, $\mathcal{F}_2 = f_e(d, \theta_1^e) f_t^{\mathrm{T}}(d, \theta_1^t) \in \mathbb{R}^{(s_1^e+1)\times(s_1^t+1)}$ and $\mathcal{F}_3 = f_t(d, \theta_1^t) f_t^{\mathrm{T}}(d, \theta_1^t) \in \mathbb{R}^{(s_1^t+1)\times(s_1^t+1)}$. We have suppressed the dependence of the matrices F, $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$ on the parameter θ_1 in our notation.

We consider approximate designs in the sense of Kiefer (1974), which are defined as probability measures with finite support on the design space $\mathcal{D}$. If an approximate design ξ has k support points, say $d_1, \ldots, d_k$,

with corresponding positive weights $\omega_1, \ldots, \omega_k$ such that $\sum_{i=1}^k \omega_i = 1$, and n_1 observations can be taken, a rounding procedure is applied to obtain integers n_{1i} $(i = 1, \ldots, k)$ from the possibly rational numbers $\omega_i n_1$. The information matrix $M_1(\xi, \theta_1)$ of a design ξ is defined by the $s_1 \times s_1$ matrix

$$M_1(\xi, \theta_1) = \int_{\mathcal{D}} \mathcal{I}_1(d, \theta_1)\, \mathrm{d}\xi(d) = \sum_{i=1}^{k} \frac{\omega_i}{\sigma_e^2 \sigma_t^2 (1-\rho^2)} F(d_i),$$

where the matrix $F(d)$ is defined in (2).

Under standard regularity conditions, the maximum likelihood estimator $\hat{\theta}_1$ is asymptotically normally distributed, that is, ${n_1}^{1/2}(\hat{\theta}_1 - \theta_1)$ tends in distribution to $\mathcal{N}_{s_1}\{0, M_1^{-1}(\xi, \theta_1)\}$ as n_1 tends to infinity. Consequently, we search for designs that maximize the information matrix $M_1(\xi, \theta_1)$ in an appropriate sense. To be precise, let $p \in [-\infty, 1)$ and let $K \in \mathbb{R}^{s_1 \times m}$ be a given matrix of full column rank. A design ξ^* is called locally ϕ_p-optimal for estimating the linear combination $K^{\mathrm{T}}\theta_1$ if it maximizes the concave functional $\phi_p(\xi) = [\mathrm{tr}\{K^{\mathrm{T}} M_1^{-}(\xi, \theta_1) K\}^{-p}/m]^{1/p}$ among all designs ξ satisfying $\mathrm{Range}(K) \subset \mathrm{Range}\{M_1(\xi, \theta_1)\}$, i.e., $K^{\mathrm{T}}\theta_1$ is estimable by the design ξ (Kiefer, 1974). Here, $\mathrm{tr}(A)$ and A^- denote the trace and a generalized inverse of the matrix A, respectively.

One key advantage of working with approximate designs is that convex optimization theory can be applied and a general equivalence theorem is available to verify whether a design is optimal among all designs. Its proof is a direct application of Theorem 7.14 in Pukelsheim (2006).

THEOREM 1. *Let K be a $s_1 \times m$ matrix of full column rank. If $p \in (-\infty, 1)$, a design ξ^* with* $\mathrm{Range}(K) \subset \mathrm{Range}\{M_1(\xi^*, \theta_1)\}$ *is locally ϕ_p-optimal for estimating the linear combination $K^{\mathrm{T}}\theta_1$ if and only if there exists a generalized inverse G of the information matrix $M_1(\xi^*, \theta_1)$ such that*

$$\tau(d, \xi^*) = \mathrm{tr}\left[\mathcal{I}_1(d, \theta_1) G K \{C_K(\xi^*)\}^{p+1} K^{\mathrm{T}} G^{\mathrm{T}}\right] - \mathrm{tr}\{C_K(\xi^*)\}^p \leqslant 0 \tag{3}$$

holds for all $d \in \mathcal{D}$, where $C_K(\xi^*) = \{K^{\mathrm{T}} M_1^{-}(\xi^*, \theta_1) K\}^{-1}$. If $p = -\infty$, a design ξ^* with $\mathrm{Range}(K) \subset \mathrm{Range}\{M_1(\xi^*, \theta_1)\}$ *is locally $\phi_{-\infty}$-optimal for estimating the linear combination $K^{\mathrm{T}}\theta_1$ if and only if there exists a generalized inverse G of the information matrix $M_1(\xi^*, \theta_1)$ and a nonnegative-definite matrix $E \in \mathbb{R}^{m \times m}$ with $\mathrm{tr}(E) = 1$ such that*

$$\mathrm{tr}\left[\mathcal{I}_1(d, \theta_1) G K C_K(\xi^*) E C_K(\xi^*) K^{\mathrm{T}} G^{\mathrm{T}}\right] - \lambda_{\min}\{C_K(\xi^*)\} \leqslant 0$$

holds for all $d \in \mathcal{D}$, where $\lambda_{\min}\{C_K(\xi^*)\}$ denotes the minimum eigenvalue of $C_K(\xi^*)$. Moreover, at the support points of any ϕ_p-optimal design there is equality in the above inequalities.

The sensitivity function on the left-hand side of (3) can be used to provide a lower bound on the ϕ_p-efficiency of any design for $p > -\infty$ (Dette, 1996), that is,

$$\mathrm{eff}_p(\xi) = \frac{\phi_p(\xi)}{\sup_\nu \phi_p(\nu)} \geqslant \frac{\mathrm{tr}\{C_K(\xi)\}^p}{\max_{d \in \mathcal{D}} \mathrm{tr}[\mathcal{I}_1(d, \theta_1) G K \{C_K(\xi)\}^{p+1} K^{\mathrm{T}} G^{\mathrm{T}}]}.$$

Characterizations of the type (3) are also useful for finding optimal designs analytically if the model is not too complicated. However, regression models with a multivariate outcome are complex and in practice optimal designs have to be found numerically (Chang, 1997; Sagnol, 2011). For such calculations, sharp bounds on the number of support points of the optimal designs reduce the complexity of the optimization problem substantially and will be derived in the following section.

3. OPTIMAL DESIGNS FOR DOSE-FINDING TRIALS WITHOUT AN ACTIVE CONTROL

3·1. *Introduction*

A design ξ_1 is called admissible if there does not exist a design ξ_2 such that $M_1(\xi_1, \theta_1) \neq M_1(\xi_2, \theta_1)$ and $M_1(\xi_1, \theta_1) \leqslant_L M_1(\xi_2, \theta_1)$ with respect to the Loewner ordering (Karlin & Studden, 1966). Recently, the

characterization of the number of support points of admissible designs has received considerable attention (Yang & Stufken, 2009; Yang, 2010; Dette & Melas, 2011; Yang & Stufken, 2012; Dette & Schorning, 2013), leading to substantially smaller bounds than provided by the classical approach using Caratheodory's theorem (Pukelsheim, 2006). We show in Theorem A.1 of the Supplementary Material that these results can be proved under weaker assumptions than usually made in the literature using the theory of Tchebycheff systems (Karlin & Studden, 1966). More specifically, in Theorem A.1 we provide a characterization of admissible designs that generalizes Theorem 3.1 in Dette & Melas (2011) and can be used to derive new bounds on the number of support points of admissible designs for the commonly used nonlinear regression models in dose-finding trials without an active control for joint efficacy-toxicity outcomes.

3·2. *Bounds on the number of support points*

The function which maps the matrix M to $(K^{\mathrm{T}}M^{-}K)^{-1}$ is increasing with respect to the Loewner ordering on the set of all $s_1 \times s_1$ matrices satisfying $\mathrm{Range}(K) \subset \mathrm{Range}(M)$ (Pukelsheim, 2006). That is, if $M_1 \geqslant_L M_2$ then $(K^{\mathrm{T}}M_1^{-}K)^{-1} \geqslant_L (K^{\mathrm{T}}M_2^{-}K)^{-1}$ for all matrices M_1 and M_2 satisfying the range inclusion. Because the ϕ_p-criteria are monotone, we have $\phi_p(\xi) \leqslant \phi_p(\xi^*)$ for any designs ξ and ξ^*, where $M_1(\xi,\theta) \leqslant_L M_1(\xi^*,\theta)$. The following results give upper bounds on the number of support points of such designs. For a precise formulation, we define the index $I(\xi)$ of a design ξ on the interval $[L,R]$ as the number of support points, where each interior support point is counted as one and each support point at the boundary of the interval $[L,R]$ is counted as one half. The following results are proved in the Supplementary Material, where one can also find additional results for exponential models. For the following two statements let ξ denote an arbitrary design on the dose interval $\mathcal{D} = [L,R]$.

THEOREM 2. *Assume that the model for efficacy is given by* $\eta_1^e(d,\theta_1^e) = \vartheta_0^e + \vartheta_1^e d + \vartheta_2^e d^2$.

(a) *If* $\eta_1^t(d,\theta_1^t) = \vartheta_0^t + \vartheta_1^t d + \vartheta_2^t d^2$, *there exists a design* ξ^* *with at most three support points such that* $M_1(\xi^*,\theta_1) \geqslant_L M_1(\xi,\theta_1)$. *If* $I(\xi) \geqslant 2$, ξ^* *can be chosen such that its support contains* L *and* R.
(b) *If* $\eta_1^t(d,\theta_1^t) = \vartheta_1^t d(\vartheta_2^t + d)^{-1}$ *or* $\eta_1^t(d,\theta_1^t) = \vartheta_0^t + \vartheta_1^t d(\vartheta_2^t + d)^{-1}$, *there exists a design* ξ^* *with at most five support points such that* $M_1(\xi^*,\theta_1) \geqslant_L M_1(\xi,\theta_1)$. *If* $I(\xi) \geqslant 4$, ξ^* *can be chosen such that its support contains* L *and* R.

THEOREM 3. *Assume that both the model for efficacy and the model for toxicity are given by* $\eta(d,\theta) = \vartheta_0 + \vartheta_1 d(\vartheta_2 + d)^{-1}$ *or* $\eta(d,\theta) = \vartheta_1 d(\vartheta_2 + d)^{-1}$. *Then there exists a design* ξ^* *with at most five support points such that* $M_1(\xi^*,\theta_1) \geqslant_L M_1(\xi,\theta_1)$. *If* $I(\xi) \geqslant 4$, ξ^* *can be chosen such that its support contains* L *and* R.

Remark 1. The remaining cases can be obtained by swapping the roles of η^e and η^t in Theorems 2 and 3. For example, if $\eta_1^e(d,\theta_1) = \vartheta_0^e + \vartheta_1^e d(\vartheta_2^e + d)^{-1}$ and $\eta_1^t(d,\theta_1) = \vartheta_0^t + \vartheta_1^t d + \vartheta_2^t d^2$, then it follows from Theorem 2(b) that for any design ξ there exists a design ξ^* with at most five support points such that $M_1(\xi^*,\theta_1) \geqslant_L M_1(\xi,\theta_1)$. Moreover, if $I(\xi) \geqslant 4$, then ξ^* can be chosen such that its support contains L and R. The other cases are obtained similarly.

3·3. *Minimally supported D-optimal designs*

Let $\#\,\mathrm{supp}(\xi)$ denote the number of support points of a design ξ and let $m^* = \min\{\#\,\mathrm{supp}(\xi) \mid \det\{M_1(\xi,\theta_1)\} > 0,\ \xi \text{ is a design on } \mathcal{D}\}$ be the minimal number of support points of a design with a non-singular information matrix in model (1). A design ξ is called minimally supported if $\det\{M_1(\xi,\theta_1)\} > 0$ and the number of support points is given by m^*. In general, optimal designs have to be determined numerically for complex models, and even then many current algorithms may not work well. However, restricting the search to minimally supported designs can greatly simplify the optimization problem, which may then allow us to determine locally D-optimal designs.

THEOREM 4. *If the number of parameters in the mean function for the efficacy model is the same as for the toxicity model, i.e., $s_1^e = s_1^t$, the minimally supported locally D-optimal design for model (1) is a uniform design. Moreover, its support points do not depend on the entries in the covariance matrix Σ_1.*

The following result provides minimally supported D-optimal designs for some dose-response models. Its proof uses Theorem 4, which reduces the optimization problem to the determination of the support points. Here and elsewhere, we let $a \vee b$ denote the maximum of the two numbers in the set $\{a, b\}$.

THEOREM 5. *Assume the user-selected dose interval is $\mathcal{D} = [L, R]$.*

(a) *Assume that the model for efficacy is given by $\eta_1^e(d, \theta_1^e) = \vartheta_0^e + \vartheta_1^e d + \vartheta_2^e d^2$.*

  (i) *If $\eta_1^t(d, \theta_1^t) = \vartheta_0^t + \vartheta_1^t d + \vartheta_2^t d^2$, the minimally supported D-optimal design is a three-point design with equal masses at the points L, $(L + R)/2$ and R.*
  (ii) *If $\eta_1^t(d, \theta_1^t) = \vartheta_0^t + \vartheta_1^t d(\vartheta_2^t + d)^{-1}$, the minimally supported D-optimal design is a three-point design with equal masses at the points L, $\{(L + \vartheta_2^t)(R + \vartheta_2^t)\}^{1/2} - \vartheta_2^t$ and R.*

(b) *Assume that the model for efficacy is given by $\eta_1^e(d, \theta_1^e) = \vartheta_1^e d(\vartheta_2^e + d)^{-1}$. If $\eta_1^t(d, \theta_1^t) = \vartheta_1^t d(\vartheta_2^t + d)^{-1}$, the minimally supported D-optimal design is a two-point design with equal masses at the optimal points $L \vee [\{R\vartheta_2^e\vartheta_2^t(R + \vartheta_2^e + \vartheta_2^t) + (\vartheta_2^e\vartheta_2^t)^2\}^{1/2} - \vartheta_2^e\vartheta_2^t](R + \vartheta_2^e + \vartheta_2^t)^{-1}$ and R.*

(c) *Assume that the model for efficacy is given by $\eta_1^e(d, \theta_1^e) = \vartheta_0^e + \vartheta_1^e d(\vartheta_2^e + d)^{-1}$. If $\eta_1^t(d, \theta_1^t) = \vartheta_0^t + \vartheta_1^t d(\vartheta_2^t + d)^{-1}$, the minimally supported D-optimal design is a three-point design with equal masses at the points L, $[\{(L + \vartheta_2^e)(L + \vartheta_2^t)(R + \vartheta_2^e)(R + \vartheta_2^t)\}^{1/2} + LR - \vartheta_2^e\vartheta_2^t](L + R + \vartheta_2^e + \vartheta_2^t)^{-1}$ and R.*

4. OPTIMAL DESIGNS FOR ACTIVE CONTROLLED DOSE-FINDING TRIALS

We now extend the preceding results to active controlled dose-finding trials with a predetermined total number of patients N by determining the optimal number k of different dose levels for the new drug, their individual dose levels $d_1, \ldots, d_k$, and the optimal number n_1 of patients to be assigned to the new drug, along with the allocation scheme across the recommended doses. The remaining number $n_2 = N - n_1$ of the patients are assigned to the active control, usually a marketed drug administered at a fixed dose level C. Thus, we have designs of the form

$$\tilde{\xi} = \begin{pmatrix} (d_1, 0) & \ldots & (d_k, 0) & (C, 1) \\ \tilde{\omega}_1 & \ldots & \tilde{\omega}_k & \tilde{\omega}_{k+1} \end{pmatrix}, \tag{4}$$

where $\tilde{\omega}_i$ and $\tilde{\omega}_{k+1}$ denote the proportion of patients treated at the ith dose level of the new drug ($i = 1, \ldots, k$) and with the active control, respectively; thus, $n_2 \approx \tilde{\omega}_{k+1} N$. The second component of the design points in (4) specifies whether patients receive the new drug, 0, or the active control, 1. Note that the approximate design $\tilde{\xi}$ induces an approximate design of the form

$$\xi = \begin{pmatrix} d_1 & \ldots & d_k \\ \omega_1 & \ldots & \omega_k \end{pmatrix} \tag{5}$$

for the new drug by defining $\omega_i = \tilde{\omega}_i/(1 - \tilde{\omega}_{k+1})$. Extending the statistical model from Dette et al. (2014) to the efficacy-toxicity outcomes considered here, we have

$$Y_{ij} = (Y_{ij}^e, Y_{ij}^t)^{\mathrm{T}} \sim \mathcal{N}_2\{\eta_1(d_i, \theta_1), \Sigma_1\} \quad (j = 1, \ldots, n_{1i}), \tag{6}$$

$$Z_j = (Z_j^e, Z_j^t)^{\mathrm{T}} \sim \mathcal{N}_2\{\eta_2(\theta_2), \Sigma_2\} \quad (j = 1, \ldots, n_2), \tag{7}$$

where Y_{ij} denotes the outcome of the jth patient treated with the new drug at dose level d_i, and Z_j the outcome from the jth patient treated with the active control.

The two-dimensional vector $\eta_2(\theta_2)$ is the expected outcome, where the parameter θ_2 varies in a compact parameter space, say $\Theta_2 \subset \mathbb{R}^2$, and Σ_2 is a 2×2 covariance matrix. The function η_2 that maps Θ_2 to $\mathbb{R}^2$ is

assumed to be continuously differentiable. Assuming that all observations are independent, the information matrix of a design $\tilde{\xi}$ defined in (4) has a block structure of the form

$$M(\tilde{\xi},\theta) = \begin{pmatrix} (1-\tilde{\omega}_{k+1})M_1(\xi,\theta_1) & 0 \\ 0 & \tilde{\omega}_{k+1}\mathcal{I}_2(\theta_2) \end{pmatrix}, \tag{8}$$

where $\theta = (\theta_1^{\mathrm{T}}, \theta_2^{\mathrm{T}})^{\mathrm{T}}$ and

$$\mathcal{I}_2(\theta_2) = \left\{ \frac{\partial}{\partial\theta_2}\eta_2(\theta_2) \right\}^{\mathrm{T}} \Sigma_2^{-1} \left\{ \frac{\partial}{\partial\theta_2}\eta_2(\theta_2) \right\}$$

is the 2×2 Fisher information matrix corresponding to the active control. Following Dette et al. (2015), locally optimal designs for active controlled dose-finding trials can be obtained from locally optimal designs for dose-finding trials without an active control. We extend this result to the class of admissible designs introduced in the preceding sections.

THEOREM 6. *If ξ is an admissible design of the form (5) in model (1) and $\tilde{\omega}_{k+1} \in (0,1)$, the design $\tilde{\xi}$ defined in (4) is an admissible design for the model (6) with an active control (7).*

In a similar way, ϕ_p-optimal designs for active controlled trials with efficacy-toxicity outcomes can be obtained. For this purpose we state the following result, which can be proved in a similar way to Theorem 1 in Dette et al. (2015) using the block structure of the matrix $M(\tilde{\xi},\theta)$ in (8).

THEOREM 7. *Let ξ^* denote the locally ϕ_p-optimal design of the form (5) in the dose-response model (6) with masses $\omega_1^*,\ldots,\omega_k^*$ at the points $d_1^*,\ldots,d_k^*$, respectively. The design $\tilde{\xi}^*$ with masses $\tilde{\omega}_1^* = \rho_p(1+\rho_p)^{-1}\omega_1^*,\ldots,\tilde{\omega}_k^* = \rho_p(1+\rho_p)^{-1}\omega_k^*$ and $\tilde{\omega}_{k+1}^* = (1+\rho_p)^{-1}$ at the points $(d_1^*,0),\ldots,(d_k^*,0)$ and $(C,1)$, respectively, is locally ϕ_p-optimal in the dose-response model (6) with an active control (7), where*

$$\rho_p = \begin{cases} \dfrac{(\mathrm{tr}[\{\mathcal{I}_2^{-1}(\theta_2)\}^{-p}])^{1/(p-1)}}{(\mathrm{tr}[\{M_1^{-1}(\tilde{\xi}^*,\theta_1)\}^{-p}])^{1/(p-1)}}, & p \in (-\infty,1)\setminus\{0\}, \\ \dfrac{s_1}{2}, & p = 0, \\ \dfrac{\lambda_{\min}\{\mathcal{I}_2(\theta_2)\}}{\lambda_{\min}\{M_1(\tilde{\xi}^*,\theta_1)\}}, & p = -\infty. \end{cases}$$

Theorem 7 can be extended to construct minimally supported designs. In particular, any minimally supported ϕ_p-optimal design of the form (5) for the dose-response model (6) yields a minimally supported ϕ_p-optimal design for the dose-response model (6) with an active control (7) by the transformation described in Theorem 7.

Example 1. Assume that efficacy and toxicity are described over $\mathcal{D} = [0,7]$ by a quadratic model with parameter $\theta_1^e = (0{\cdot}5, 0{\cdot}01, 0{\cdot}1)^{\mathrm{T}}$ and an Emax model with parameter $\theta_1^t = (0{\cdot}1, 2{\cdot}4, 1{\cdot}2)^{\mathrm{T}}$, respectively; see Jin & Barker (2016) and Thomas & Roy (2017) for the choice of dose-response models in clinical trials. Assume further that $\sigma_e = 0{\cdot}1$ and $\sigma_t = 0{\cdot}4$. It follows from Theorem 2(b) and Theorem 6 that only designs with at most six support points have to be considered for the corresponding model with an active control. We first used a metaheuristic particle swarm optimization algorithm to generate the locally D-optimal design for model (1); see Wong et al. (2015) for an explanation of the algorithm. We then applied Theorem 7 to determine the locally optimal design for the model with an active control. The results for the locally D-optimal designs are listed in the left part of Table 1 for $\rho = 0{\cdot}1, 0{\cdot}5, 0{\cdot}9$. The D-optimal designs are not minimally supported and the support points and weights depend on the correlation. The minimally supported D-optimal designs can be found by an application of Theorem 5 and are presented in the right part of Table 1. Note that these designs do not depend on ρ since they are equally weighted

Table 1. *Locally and minimally supported D-optimal designs for the models in Example 1 and for various values of the correlation coefficient ρ*

ρ	Locally D-optimal design					Minimally supported D-optimal design			
0·1	$(0, 0)$	$(0{\cdot}86, 0)$	$(3{\cdot}58, 0)$	$(7, 0)$	$(C, 1)$	$(0, 0)$	$(1{\cdot}94, 0)$	$(7, 0)$	$(C, 1)$
	0·225	0·15	0·15	0·225	0·25	0·25	0·25	0·25	0·25
0·5	$(0, 0)$	$(0{\cdot}8, 0)$	$(3{\cdot}73, 0)$	$(7, 0)$	$(C, 1)$	$(0, 0)$	$(1{\cdot}94, 0)$	$(7, 0)$	$(C, 1)$
	0·2175	0·1575	0·1575	0·2175	0·25	0·25	0·25	0·25	0·25
0·9	$(0, 0)$	$(0{\cdot}7, 0)$	$(3{\cdot}99, 0)$	$(7, 0)$	$(C, 1)$	$(0, 0)$	$(1{\cdot}94, 0)$	$(7, 0)$	$(C, 1)$
	0·21	0·165	0·165	0·21	0·25	0·25	0·25	0·25	0·25

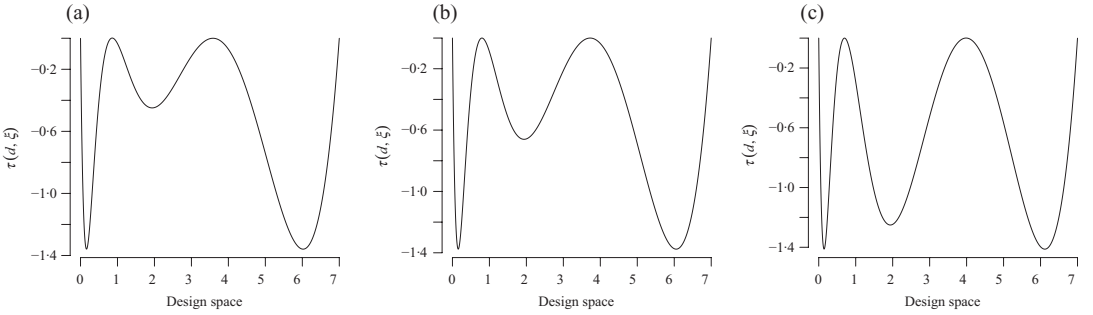

Fig. 1. Sensitivity functions of the locally D-optimal designs in Table 1 for (a) $\rho = 0{\cdot}1$, (b) $\rho = 0{\cdot}5$ and (c) $\rho = 0{\cdot}9$.

by Theorem 4. The optimality of the numerically calculated locally D-optimal designs was checked by Theorem 1 and the corresponding sensitivity functions are displayed in Fig. 1 for different values of the correlation. All designs calculated by the particle swarm optimization algorithm are in fact D-optimal. Moreover, the minimally supported designs are not optimal and their D-efficiencies are given by 0·97, 0·95 and 0·82 for the cases $\rho = 0{\cdot}1, 0{\cdot}5$ and 0·9, respectively. This indicates that minimally supported designs are only efficient if the bivariate outcomes are weakly correlated. For a strong correlation between efficacy and toxicity, minimally supported designs cannot be recommended. A direct calculation shows the uniform design with dose levels $0{\cdot}00, 0{\cdot}35, 1{\cdot}40, 2{\cdot}80, 4{\cdot}20, 5{\cdot}60, 7{\cdot}00$ has D-efficiencies $0{\cdot}89, 0{\cdot}89, 0{\cdot}88$ when $\rho = 0{\cdot}1, 0{\cdot}5$ and 0·9, respectively.

5. Conclusions

There are several lines of future research. First, the results of this paper on locally optimal designs require a priori information about the unknown model parameters and an interesting direction is to further develop the methodology to accommodate more sophisticated optimality criteria, which are robust against a misspecification of the unknown parameters. Second, while the mentioned applications typically consider bivariate outcomes, it is also of interest to extend these results to multivariate responses. The results of Mukhopadhyay & Khuri (2008) indicate that such an extension is very challenging. Finally, this paper considers models with additive normally distributed random errors. This assumption is mainly made for the sake of transparent notation, and using similar techniques to those described by Dette et al. (2015) the results can be extended to more general models involving exponential families. On the other hand, a further important and very challenging direction of future research is the development of a corresponding methodology which is applicable in random effect models.

Acknowledgement

This work was supported in part by the German Research Foundation. Dette, Kettelhake, Schorning and Wong were partially supported by the National Institute of General Medical Sciences of the U.S. National

Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We would also like to thank the reviewers for their constructive comments on an earlier version of this paper. Frank Bretz is also affiliated with the Shanghai University of Finance and Economics, China.

SUPPLEMENTARY MATERIAL

Supplementary material available at *Biometrika* online contains proofs, additional results for exponential models and additional examples.

REFERENCES

CHANG, S. (1997). An algorithm to generate near *D*-optimal designs for multiple response surface models. *IEEE Trans.* **29**, 1073–81.

DETTE, H. (1996). Lower bounds for efficiencies with applications. In *Research Developments in Probability and Statistics: Festschrift zum 65sten Geburtstag von M.L. Puri*, E. Brunner & M. Denker, eds. Utrecht: VSP, pp. 111–24.

DETTE, H., KETTELHAKE, K. & BRETZ, F. (2015). Designing dose-finding studies with an active control for exponential families. *Biometrika* **102**, 937–50.

DETTE, H., KISS, C., BENDA, N. & BRETZ, F. (2014). Optimal designs for dose finding studies with an active control. *J. R. Statist. Soc.* B **76**, 265–95.

DETTE, H. & MELAS, V. B. (2011). A note on the de la Garza phenomenon for locally optimal designs. *Ann. Statist.* **39**, 1266–81.

DETTE, H. & SCHORNING, K. (2013). Complete classes of designs for nonlinear regression models and principal representations of moment spaces. *Ann. Statist.* **41**, 1260–7.

DRAGALIN, V., FEDOROV, V. V. & WU, Y. (2008). Two-stage design for dose-finding that accounts for both efficacy and safety. *Statist. Med.* **27**, 5156–76.

FAN, S. K. & CHALONER, K. (2004). Optimal designs and limiting optimal designs for a trinomial response. *J. Statist. Plan. Infer.* **126**, 347–60.

HEISE, M. A. & MYERS, R. H. (1996). Optimal designs for bivariate logistic regression. *Biometrics* **52**, 613–24.

HELMS, H., BENDA, N., ZINSERLING, P., KNEIB, T. & FRIEDE, T. (2015). Spline-based procedures for dose-finding studies with active control. *Statist. Med.* **34**, 232–48.

JIN, B. & BARKER, K. (2016). Equivalence test for two Emax response curves. *Statist. Biopharm. Res.* **8**, 307–15.

KARLIN, S. & STUDDEN, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics*. New York: Wiley.

KIEFER, J. (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.* **2**, 849–79.

MAGNUSDOTTIR, B. T. (2013). *c*-optimal designs for the bivariate Emax model. In *mODa 10: Advances in Model-Oriented Design and Analysis*, D. Ucinski, A. Atkinson & M. Patan, eds. Cham: Springer.

MUKHOPADHYAY, S. & KHURI, A. I. (2008). Comparison of designs for multivariate generalized linear models. *J. Statist. Plan. Infer.* **138**, 169–83.

PUKELSHEIM, F. (2006). *Optimal Design of Experiments*. Philadelphia: SIAM.

SAGNOL, G. (2011). Computing optimal designs of multiresponse experiments reduces to second-order cone programming. *J. Statist. Plan. Infer.* **141**, 1684–708.

SPLAWINSKI, J. & KUZNIAR, J. (2004). Clinical trials: Active control vs placebo – what is ethical? *Sci. Eng. Ethics* **10**, 73–9.

TEMPLE, R. & ELLENBERG, S. S. (2000). Placebo-controlled trials and active-control trials in the evaluation of new treatments. *Ann. Internal Med.* **133**, 455–63.

THOMAS, N. & ROY, D. (2017). Analysis of clinical dose-response in small-molecule drug development: 2009–2014. *Statist. Biopharm. Res.* **9**, 137–46.

WONG, W. K., CHEN, R. B., HUANG, C. C. & WANG, W. C. (2015). A modified particle swarm optimization technique for finding optimal designs for mixture models. *PLoS ONE*, doi: 10.1371/journal.pone.0124720.

YANG, M. (2010). On the de la Garza phenomenon. *Ann. Statist.* **38**, 2499–524.

YANG, M. & STUFKEN, J. (2009). Support points of locally optimal designs for nonlinear models with two parameters. *Ann. Statist.* **37**, 518–41.

YANG, M. & STUFKEN, J. (2012). Identifying locally optimal designs for nonlinear models: A simple extension with profound consequences. *Ann. Statist.* **40**, 1665–85.

[*Received on* 13 *May* 2016. *Editorial decision on* 12 *August* 2017]