Cost-cautious designs for confirmatory bioassay

Alexander N. Donev\textsuperscript{a,}*, Randy Tobias\textsuperscript{b}, Farinaz Monadjemi\textsuperscript{c}

\textsuperscript{a}School of Mathematics, University of Manchester, Manchester, M13 9PL, UK
\textsuperscript{b}Linear Models R&D, SAS Institute Inc., Cary, NC 27513, USA
\textsuperscript{c}Department of Probability and Statistics, University of Sheffield, Sheffield, S3 7RH, UK

\textbf{ARTICLE INFO}

\textbf{Article history:}
Received 18 September 2006
Received in revised form 21 October 2007
Accepted 31 January 2008
Available online 28 March 2008

\textbf{Keywords:}
Bayesian experimental design
Bioassay
\(D\)- and \(G\)-optimality
Cost-cautious designs
General Equivalence Theorem
Nonlinear models

\textbf{ABSTRACT}

Confirmatory bioassay experiments take place in late stages of the drug discovery process when a small number of compounds have to be compared with respect to their properties. As the cost of the observations may differ considerably, the design problem is well specified by the cost of compound used rather than by the number of observations. We show that cost-efficient designs can be constructed using useful properties of the minimum support designs. These designs are particularly suited for studies where the parameters of the model to be estimated are known with high accuracy prior to the experiment, although they prove to be robust against typical inaccuracies of these values. When the parameters of the model can only be specified with ranges of values or by a probability distribution, we use a Bayesian criterion of optimality to construct the required designs. Typically, the number of their support points depends on the prior knowledge for the model parameters. In all cases we recommend identifying a set of designs with good statistical properties but different potential costs to choose from.

\(©\) 2008 Elsevier B.V. All rights reserved.

1. Introduction

In early stages of drug development many thousands, even millions of compounds are tested for biological activity by measuring a response of interest at one or two concentrations of each compound. Then, a cascade of bioassays is carried out to study various efficacy and safety measurements of the compounds. As the compounds go through the stages of drug development their properties and mechanism of action become well known. Many compounds fail to satisfy the requirements and the number of those that reach the latter stages is usually very small. At this point, it is important to retest these compounds in a confirmatory experiment, in order to make sure that their previously detected activity is genuine, and to compare them with existing drugs on the market as well as with each other. It is important that the design of a confirmatory experiment has good statistical properties, but also that it is practically feasible and cost efficient. The design of such studies is the subject of the current research.

Customarily the total number of observations of an experimental design is constrained by limitations on the available resources, such as time and cost. The design problem then consists of finding the support of the design (e.g. the unique combinations of the explanatory variables to be used) as well as the number of required replications. However, if the cost of the observations in the design region differs substantially this can be taken into account so that the budget for the study is not overspent. Elfving (1952) defines the problem, which is later addressed by many researchers including Cook and Fedorov (1995), Tack and Vandebroek (2004), and Fedorov and Leonov (2005). In all these cases a prespecified number of observations are selected using some design criterion of optimality, or a utility function, that incorporates the information about the cost of the observations. The problem

\* Corresponding author.
E-mail address: a.n.donev@manchester.ac.uk (A.N. Donev).
of choosing the total number of the observations taking their individual costs into account is much less studied but could be

**Example 1.** In an in vivo pharmacodynamic bioassay the efficiency of a small number of promising new compounds had to be

compared to a standard commercially available compound. Relatively large amounts of the compounds were used because the

animals receiving the treatment had to be periodically dosed over one month. The commercial cost of 1 g of the most expensive

compound was around £18,000, with the remaining compounds still markedly expensive. In order to put this figure in some real

perspective, we note that on the day of the experiment the cost of gold was traded on the stock market at around £11 per gram.

Clearly, the costs of observations using different dose levels were remarkably different. Therefore, a sensible way to define the

experimental resources was with the cost of the experiment, in particular the amount of compound that was going to be used,

rather than with the number of observations to be taken, which is the usual way to define a design problem.

An important feature of this and similar confirmatory experiments is that their aim is to refine the knowledge about the

new compounds of interest. There is no doubt about the form of the model that will summarize the data from the experiment.

Therefore, the experimental design has to be chosen in such a way that all available resources can be spent on estimating the

model parameters.

In Section 2 we provide more information about the background of the class of experiments of our interest and the appropriate

statistical models for describing data obtained in them. We also review some standard criteria of optimality that are relevant in

these cases. We discuss appropriate strategies for design of confirmatory experiments in Section 3, where we review some of the

useful properties of minimum support designs and describe the Bayesian approach to design construction. In Section 4 we return
to Example 1 and illustrate how the ideas presented in Section 3 can be extended to construct cost-efficient designs. The paper
ends with a discussion about the robustness of these designs and the opportunities to extend our approach.

### 2. Background

In a bioassay the effect of a compound is measured for different concentrations or doses. An experimental design is defined by

the doses \(d_1, d_2, \ldots, d_k\) for which the response have to be measured, and by the number of replicates \(w_1, w_2, \ldots, w_k\,

\(\sum_{i=1}^{k} w_i = n\), at each of these concentrations. The points \(d_1, d_2, \ldots, d_k\) form the support of the design. The model for the \(i\)th observation of the

response variable \(y_i\),

\[
E[y_i] = f(x_i, \theta) = \gamma + \frac{\delta - \gamma}{1 + 10^{(x_i - z)/\beta}},
\]

proposed by Hill (1913), is supported by scientific knowledge as well as proven ability to explain data collected in a variety of

biological studies. In (1) \(x_i\) is the base 10 logarithm of the \(i\)th dose \(d_i\), \(E[y_i]\) is the corresponding expected response, \(z\) is the base 10 logarithm of the ED50 or IC50, the dose required to achieve a response half way between the maximum and minimum possible responses \(\delta\) and \(\gamma\), respectively, \(\beta\) is known as the Hill slope, and \(\theta^T = (z, \beta, \gamma, \delta)\) is the vector of all model parameters. The smaller

the value of \(z\), the larger the potency of the compound. Homoscedastic additive independent errors in measuring the response

are often assumed, although there are important cases in practice when this assumption has to be relaxed as the observations

might be correlated and the variance of the experimental errors might increase with the value of the response. An example of

bioassay data and the prediction of the response with model (1) fitted to them are shown in Fig. 1. In this case the experimental

design consists of duplicates of eight doses equally distant on the logarithmic scale.

There are many situations when (1) can be simplified because \(\gamma\) and \(\delta\) are known. For example, if the response is the proportion

of cells affected by a full inhibitor compound, \(\gamma = 0\) and \(\delta = 1\). Then, the simpler inhibition model

\[
E[h_i] = f(x_i, \theta) = \frac{1}{1 + 10^{(x_i - z)/\beta}}
\]

can be used. In (2) the inhibition, \(h_i\), is the achieved proportion from the maximum possible effect and \(\theta^T = (z, \beta)\). Model (2) is

useful even when \(\gamma\) and \(\delta\) are not known precisely. One reason for that is that it is often useful to scale the data so that \(\gamma = 0\) and

\(\delta = 1\) approximately, as this allows for a simpler statistical analysis to be carried out as well as for a simpler interpretation of the

results. Another reason is that the estimation of \(\gamma\) and \(\delta\) does not require the use of compounds. Indeed, no compound is used at \(x \to -\infty\) (that is when the dose is zero) when \(h = \gamma\), and at \(x \to \infty\) when \(h = \delta\). In practice the latter measurements would

usually be obtained using a positive control, rather than by using a large dose of the compound. Taking observations at \(x \to -\infty\)

and \(x \to \infty\) is always useful but the decision about how many observations to collect at each of these dose levels is usually driven

by practicalities (e.g. number of available objects in an in vivo bioassay or number of available wells on the experimental plate in

an in vitro experiment). Therefore, without loss of generality, we focus our study on model (2).

Model (2) is nonlinear. Estimation of the parameters of such models usually employs iterative optimization algorithms which

use as a starting point of the search the initial values of the model parameters, say \(z = z_0\) and \(\beta = \beta_0\). For a review of some available

methods, see Bates and Watts (2007) and Seber and Wild (2005).
The statistical properties of the estimates of the model parameters depend on the information matrix. For example, the asymptotic information matrix is

$$M(\theta) = X^TV^{-1}X,$$  \hspace{1cm} (3)

where for model (2) $X$ is an $n \times 2$ matrix, so that the $i$th element of its two columns are

$$\frac{\partial f(x_i, \theta)}{\partial x} = \frac{(\gamma - \delta)10^{(x-x)_i} \beta}{(1 + 10^{(x-x)_i}\beta)^2},$$  \hspace{1cm} (4)

$$\frac{\partial f(x_i, \theta)}{\partial \beta} = \frac{(\gamma - \delta)10^{(x-x)_i}(x - x)_i \ln(10)}{(1 + 10^{(x-x)_i}\beta)^2},$$  \hspace{1cm} (5)

respectively. In (3) $V$ is the variance covariance matrix for the $n \times 1$ vector of observations. For simplicity, in the examples that follow we assume that $V = I_n\sigma^2$, where $I_n$ is the identity matrix. In Eqs. (4) and (5) $\ln(10)$ denotes the natural logarithm of 10. Clearly the elements of the information matrix depend on the true values of the model parameters, and therefore they are evaluated for the estimates of the parameters, $\bar{x} = \bar{x}_0$ and $\bar{\beta} = \beta_0$.

We are interested in the statistical properties of designs that are $D$-optimal for estimating $\theta$. For such designs the doses are chosen in such a way that the determinant of the information matrix $|M(\theta)|$ is maximised. The search for $x_1, x_2, \ldots, x_k$ and $w_1, w_2, \ldots, w_k$ is constrained within a range of doses of interest, i.e. the design region $\Omega$. The $D$-optimality criterion requires that the volume of the confidence ellipsoid for the model parameters is minimized. However, as the elements of the information matrix depend on the true values of the model parameters, $D$-optimal designs can be constructed only if information about their possible values is available. When such values are available and they are used in the design construction, resulting designs are called locally $D$-optimal. For more information about this approach see for instance Fedorov (1972), Silvey (1980, Chapter 6), Melas (2006, Chapter 2) and Atkinson et al. (2007, Chapter 17). When the values of the model parameters can only be specified by ranges of plausible values or, in general, by a distribution function, say $\psi(\theta)$, then a Bayesian $D$-optimality criterion can be used to construct the designs. We discuss this case further in Section 3.2.

Another useful property that an experimental design $\xi$ may possess is that of $G$-optimality. The approximate prediction variance of the response at a point $x$ is

$$d(x, \xi) = f^T(x)M^{-1}(\theta)f(x).$$  \hspace{1cm} (6)

A design which minimizes the maximum prediction variance of the response over the design region $\Omega$ is called locally $G$-optimal. In (6), $x \in \Omega$ and for model (2) $f(x)$ is an $2 \times 1$ vector with elements $\partial f(x_i, \theta)/\partial x$ and $\partial f(x_i, \theta)/\partial \beta$, as defined by Eqs. (4) and (5).

Kiefer and Wolfowitz (1960) prove that, in a certain general sense, if an experimental design is $D$-optimal, it is also $G$-optimal, a result known as the General Equivalence Theorem. They show that the variance of the prediction (6) for a design that satisfies this theorem has as many maxima as there are support points of the design and these maxima are located at the support points. This equivalence is only guaranteed if the number of observations is allowed to be non-integer. However, many studies have shown that even when this constraint is not satisfied, designs that are $D$-optimal perform well with respect to the $G$-optimality criterion.
As the main purpose of bioassay experiments is to evaluate the potency of the studied compounds, an important additional criterion to use in the comparisons of the designs is the precision of estimating the compound potency, i.e. the variance of the estimate of \( \hat{\alpha} \). However, we do not recommend using a criterion of optimality based only on this quantity (i.e. \( D_\alpha \)-optimality), as a design constructed using such a criterion may lack robustness and may not allow for assessing various important features of the experiment such as the variability and reproducibility of the results and the mechanism of action of the studies compounds.

3. Experimental strategies

Typically, the locally \( D \)- and \( G \)-optimal designs for non-linear models are minimum support designs, i.e. they have support at as many combinations of the explanatory variables as the number of the fixed parameters. For instance, Melas (2006, Chapter 6) proves this for a large class of exponential regression models. It is easy to verify that the support of the locally \( D \)-optimal design for model (2) consists of two different doses, hence a minimum support design consists of replications at those two doses.

The main disadvantage of the minimum support designs is that they are not robust in the case of model misspecification and a test for lack of fit for the model cannot be carried out. This is why these designs cannot be recommended in dose response studies in early stages of the drug discovery process. However, in a confirmatory experiment this is not an issue. Furthermore, they possess a number of useful characteristics that make them particularly attractive to use. We summarize these in two lemmas in Section 3.1. We do not provide proofs of these results as they are easy to find. The proof of Lemma 1 is given by Fedorov (1972, pp. 84–85). The result of Lemma 2, while less recognized, is most important for finding easy solutions to the design problem of our interest. It provides an important extension of the General Equivalence Theorem of Kiefer and Wolfowitz (1960) for the case where the optimal design cannot satisfy Lemma 1 because the weights are restricted to be unequal due to practical considerations. Its proof requires the use of similar ideas to that used in Fedorov (1972, p. 72). Its usefulness is illustrated in Section 4.

The statistical properties of the minimum support designs may not be good if the estimates of the model parameters used in the design construction differ considerably from their true values. Therefore, these designs may not be suitable in cases when there is no strong confidence in single values for the estimates of the model parameters. In Section 3.2 we describe a Bayesian approach to construction of \( D \)-optimal designs that is preferable in such situations.

3.1. Properties of minimum support designs

**Lemma 1.** If a minimum support design satisfies the General Equivalence Theorem, it allocates equal number of observations to all its support points.

**Lemma 2.** If the number of observations allocated at the support points are pre-specified as, say \( w_1^*, w_2^*, \ldots, w_p^* \), the \( D \)-optimal minimum support design \( \xi^* \), in the presence of these constraints, satisfies the following conditions:

1. \( \xi^* \) has the same support as \( \xi^* \), the design satisfying the General Equivalence Theorem in the absence of constraints on the number of observations;
2. design \( \xi^* \) minimizes \( d_{\text{max}} = \max_\Omega \lambda(x) d(x, \xi) \);
3. \( \lambda(x_i) d(x_i, \xi) = w_i^{* - 1} \) at the \( i \)th support design point, \( i = 1, 2, \ldots, p \); \( d_{\text{max}} = w_{\text{min}}^{* - 1} \), where \( w_{\text{min}} = \min(w_1^*, w_2^*, \ldots, w_p^*) \), and it is attained at the design point with a minimum number of observations.

The function \( \lambda(x) \) is called efficiency of the experiment (Fedorov, 1972, p. 39) and reflects the accuracy of the taken observations.

It can be shown that if observations from one support point of the \( D \)- and \( G \)-optimal design \( \xi^* \) are moved to one of the other support points, without causing singularity, the variances of the prediction at the remaining support points do not change. Also, the prediction variance has local maxima in the neighborhood of the support points, i.e. \( d(x_0, \xi^*) \sim p w_i^{* - 1} \). \( x_0 \in \Omega \), as \( x_0 \) tends to \( x_i \) for any \( i = 1, 2, \ldots, p \).

3.2. Bayesian \( D \)-optimal designs

As already noted, the Bayesian \( D \)-optimality criterion can be used to construct designs when the model parameters can be specified by ranges of plausible values or by a distribution function \( \psi(\theta) \). Various Bayesian \( D \)-optimality criteria are possible; the one we use requires the expectation of the determinant of the information matrix over \( \psi(\theta) \) to be maximized, i.e. an optimal design maximizes

\[
E_{\psi(\theta)} |M(\theta)|.
\]

For alternative Bayesian criteria see Atkinson et al. (2007, p. 293).

Using this approach requires numerical integration to evaluate the criterion of optimality (7). This is relatively simple for model (2) because there is a single explanatory variable, and the information matrix is of size \( 2 \times 2 \). Similarly to other cases
where the Bayesian approach has been used, the number of support points of the optimal designs depends on the strength of the knowledge of the values of the model parameters.

4. Cost-cautious designs

When the cost of measuring the response variable in different parts of the design region differs substantially, as in Example 1, and the size of the experimental design is defined by a constraint on its cost, then the use of expensive experimental settings has to be minimized. The saving made this way will allow cheaper experimental settings to be used more frequently in order to ensure efficient estimation of the model parameters. In this section we use the scenario of Example 1 to illustrate how the designs described in Section 3 can be adapted to keep the cost of the experiment low.

First, we use the properties of the minimum support designs presented in Section 3.1 to obtain cost-efficient designs with good statistical properties for several scenarios related to Example 1. In the absence of cost restrictions, according to Lemma 1 the experimental resources should be split equally between the two support points of the design. Suppose the values \( \alpha_0 = 0.1124 \) and \( \beta_0 = 0.7406 \) are obtained in a previous study of a compound of interest (namely, the data of Fig. 1). The locally \( D \)-optimal design for these values requires the doses 0.3166 and 5.3010 mg/kg to be used. If \( n = 12 \), six observations have to be taken at each dose level. In this case, the determinant of the information matrix is \( |M(\theta)| = 0.2456 \times 10^2 \), and the asymptotic variances for the estimates of \( \alpha \) and \( \beta \) are proportional to 4.095 and 5.997, respectively. The total amount of compound that is required is 33.708 mg/kg. The cost of the study would be proportional to this figure if apart from the cost of the compound the remaining expenditure is small, or virtually unaffected by the experimental design. Fig. 2 shows the plot of the variance of the prediction (6) with model (2) for this design. Clearly the design satisfies the General Equivalence Theorem.

An important interpretation of Lemma 2 in the context of this example is that the support of the experimental design does not depend on the numbers of the observations at different doses. Therefore, the task of choosing a suitable cost-efficient experimental design reduces to finding suitable numbers of the observations at the two doses. Taking unequal numbers of observations for each of the two doses would certainly result in a suboptimal design with respect to the \( D \)-optimality criterion, but this can be done so that the loss in precision of estimating \( \alpha \) is minimized. If the number of observations at the high dose is reduced to 4, while the number of the observations at the lower dose is increased to 12, the determinant of the information matrix for this design is reduced to \( |M(\theta)| = 0.1842 \times 10^2 \). However, interestingly the variances of the estimates of the model parameters \( \hat{\alpha} \) and \( \hat{\beta} \) are identical to those when the locally \( D \)-optimal design without constraints is used. Certainly, the correlation between \( \hat{\alpha} \) and \( \hat{\beta} \) is larger for the former design. Fig. 3 shows the plot of the variance of the prediction (6) with model (2) for this design. The location and the values of the optima are as stated by Lemma 2. The total amount of compound that is required now is 12.50 mg/kg, a reduction of over 24% which in most cases would vastly overweight the extra cost of increasing the number of observations from 12 to 16 and would therefore reduce substantially the cost of the experiment.

The Bayesian approach described in Section 3.2 is suitable if the experimenter cannot specify accurately single values for the model parameters to be used in the design construction. Table 1 shows \( D \)-optimal designs for several assumptions about the true values of the model parameters. Design D1 is that already discussed, while the remaining designs have been constructed assuming that \( \alpha \) and \( \beta \) are independent and equally likely to lie within specified ranges of values. For example, for design D5 the prior distribution for \( \alpha \) and \( \beta \) between 0.0924 and 0.1324 and 0.5406 and 0.9406, respectively. The designs are

![Fig. 2. Variance of the prediction with model (2) for the \( D \)-optimal design.](image-url)
Fig. 3. Variance of the prediction for a design allocating 12 observations at dose 0.3166 mg/kg and four observations at dose 5.3010 mg/kg.

Table 1
Bayesian $D$-optimal designs

<table>
<thead>
<tr>
<th>Design</th>
<th>$\psi(0)$</th>
<th>Doses (mg/kg)</th>
<th>Proportions of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0 = 0.1124$</td>
<td>$\beta_0 = 0.7406$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>$\pm 0$</td>
<td>$0.3166; 5.3010$</td>
<td>0.50; 0.50</td>
</tr>
<tr>
<td>D2</td>
<td>$\pm 0.2$</td>
<td>$0.3317; 5.2650$</td>
<td>0.50; 0.50</td>
</tr>
<tr>
<td>D3</td>
<td>$\pm 0.2$</td>
<td>$0.3596; 4.6480$</td>
<td>0.50; 0.50</td>
</tr>
<tr>
<td>D4</td>
<td>$\pm 0.5$</td>
<td>$0.3083; 5.4407$</td>
<td>0.50; 0.50</td>
</tr>
<tr>
<td>D5</td>
<td>$\pm 0.5$</td>
<td>$0.2636; 1.2936; 6.3377$</td>
<td>0.40; 0.20; 0.40</td>
</tr>
<tr>
<td>D6</td>
<td>$\pm 0.8$</td>
<td>$0.1692; 1.0309; 1.6198; 9.8644$</td>
<td>0.33; 0.17; 0.17; 0.33</td>
</tr>
<tr>
<td>D7</td>
<td>$\pm 1.0$</td>
<td>$0.1125; 0.6595; 2.5365; 14.8158$</td>
<td>0.26; 0.24; 0.24; 0.26</td>
</tr>
</tbody>
</table>

given by their support points and the proportions of total number of observations allocated to each of them. Designs D1, D2, D3, and D4 have two support points and splitting the observations equally between them is $D$-optimal. Design D5 has three support points, while designs D6 and D7 require four support points each.

In the case when $\psi(0)$ specifies the ranges as $0.1124 \pm 0.8$ for $\alpha$, and as $0.7406 \pm 0.5$ for $\beta$, the 4-support point design D6 is $D$-optimal. It requires twice as many replicates at doses 0.1692 and 9.8644 mg/kg as at doses 1.0309 and 1.6198 mg/kg. Hence, if $n = 12$, four replicates are required for doses 0.1692 and 9.8644 mg/kg, and two replicates are required for doses 1.0309 and 1.6198 mg/kg. The total amount of compound that is required now is 45.44 mg/kg, considerably higher than that for design D1 for any of the cases that were discussed earlier.

The increase in cost of this experiment can be seen as a result of the weaker knowledge about the true values of the model parameters. However, this cost can be reduced in a similar fashion to that used earlier for minimum support designs. The idea is simple: the $D$-optimal design is modified by reducing the number of replicates at the highest dose, and sometime also at the second highest dose; smaller doses can be used more times to partly compensate for the loss of precision. This considerably reduces the cost of the experiment.

A design constructed using the Bayesian design criterion (7) is robust as it ensures good statistical properties of the estimates of the model parameters for a range of their possible true values. For example, suppose the support of design D6 is used. Table 2 shows two designs obtained by altering design D6, each allocating different number of replications to the same support points and, hence, requiring different cost. Design D6a is obtained from design D6 by moving an observation from the highest dose to the second highest dose and making all replications the same, while design D6b requires only duplicates at the two highest doses and five observations at the two lowest doses. The required amount of compound is reduced in the two designs by over 16% and 36%, respectively, reducing substantially the cost of the experiment. The precise impact of using either of these designs on the precision of evaluating the compound potency depends on the true values of the model parameters. Table 2 shows the determinant of the information matrix and the variances of the estimates of $\alpha$ and $\beta$ if their true values were $\alpha_0 = 0.1124$ and $\beta_0 = 0.7406$. It turns out that in such a case the cost saving that designs D6a and D6b provide would not compromise the precision of estimating the true compound potency $\alpha$, as for design D6 its variance is 4.3076, while for designs D6a and D6b the variances are 3.6295 and 3.3869, respectively. Both D6a and D6b are worse than design D6 with respect to the local $D$-optimality criterion.
Table 2
Cost-efficient designs

<table>
<thead>
<tr>
<th>Design</th>
<th>Dose (replications)</th>
<th>Total amount (mg/kg)</th>
<th>(\mathbb{E}[\mathbf{M}(\mathbf{0})] \times 10^2)</th>
<th>Variance of (\hat{\alpha})</th>
<th>Variance of (\hat{\beta})</th>
</tr>
</thead>
<tbody>
<tr>
<td>D6</td>
<td>0.1692 (4)</td>
<td>45.44</td>
<td>0.3075</td>
<td>4.3076</td>
<td>7.1374</td>
</tr>
<tr>
<td></td>
<td>1.0309 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6198 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.8644 (4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D6a</td>
<td>0.1692 (3)</td>
<td>38.05</td>
<td>0.3805</td>
<td>3.6295</td>
<td>9.3858</td>
</tr>
<tr>
<td></td>
<td>1.0309 (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6198 (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.8644 (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D6b</td>
<td>0.1692 (5)</td>
<td>28.97</td>
<td>0.2897</td>
<td>3.3869</td>
<td>8.7250</td>
</tr>
<tr>
<td></td>
<td>1.0309 (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6198 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.8644 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Evaluation for \(\alpha_0 = 0.1124\) and \(\beta_0 = 0.7406\).

but this has negligible practical importance. Similar calculations can easily be carried out for a number of values of \(\alpha_0\) and \(\beta_0\) in order to take a final decision about which design to use in the study.

5. Concluding remarks

A common perception about the locally \(D\)-optimal designs is that they are not robust when the values of the model parameters used for their construction are different from the true parameter values. A comparison of the designs presented in Table 1 show that their robustness in typical scenarios for confirmatory experiments is satisfactory. Indeed, the minimum support designs D2, D3, and D4 are optimal for the Bayesian criterion of optimality (7) for typical inaccuracies in specifying the true values of the model parameters in confirmatory experiments. Yet, these designs are not much different from design D1. Design D5 requires three support points, while designs D6 and D7 require four support points. However, cases when these designs would be useful are rare. For instance D7 is intended to cover potencies between 0.13 and 12.95 mg/kg, which is unusually wide range for a confirmatory experiment.

It is perhaps surprising how small modifications in \(D\)-optimal designs can result in much cheaper experimental designs without compromising the statistical properties of the results. The practical value of any cost savings cannot be generalized, but if it contributes to taking a well-informed business decision, it is likely to be substantial.

It is certainly possible to complicate our approach further by introducing a utility function that incorporates the cost of the observations, the statistical design criterion and the size of the experimental designs, and then to use this function in the construction of the required experimental designs. Such an approach has been followed for instance by Cook and Fedorov (1995), Mentré et al. (1997), Tack and Vandebroek (2004) and Fedorov and Leonov (2005). However, this can be much more difficult to implement in the case that we consider than when the total number of observations is fixed. The reason for this is that changing the design size may require various practical considerations to be taken into account which might be difficult, or even impossible, to formalize. For instance, in Example 1, if the number of animals per group in an in vivo bioassay is increased beyond a particular figure (say 5), multiple cages for housing the animals receiving a particular dose have to be used and further potential sources of variability might be introduced. It would be difficult to quantify the increase in complexity of carrying out the experiment and of the following statistical analysis. Similarly, an experimental design for an in vitro bioassay where the required observations nicely fill in a plate with chosen size (e.g. 96-well plate) would appeal much more to an experimenter than one which complicates the study further and increases the probability of human error. In both cases, complex non-statistical considerations would have to be made in order to come up with a sensible utility function.

Our experience suggests that providing an opportunity to the experimenters to consider a set of experimental designs with similar statistical properties but requiring different experimental costs is a better practical alternative to the one based on a potentially esoteric utility function. The result of Lemma 2 considerably simplifies the choice of a minimum support design as it makes it easy to identify sets of designs with good statistical properties but with different costs. The experimenter can then choose the design that best suits the experimental situation. Developing software that automates and therefore eases this process would be useful.

The idea to reduce the cost by collecting larger number of observations at smaller doses than at higher doses can also be extended to studies when a Bayesian criterion of optimality is used. In in vivo studies similar to that of Example 1 this would also carry the additional ethical benefit of exposing fewer subjects to potentially toxic high doses.

Acknowledgments

The authors would like to thank the Editor, the Associate Editor, and the referees for their constructive criticism and suggestions which led to a considerable improvement of the paper.
References


