Computational Statistics and Data Analysis 71 (2014) 288–297

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Optimal sequential designs in phase I studies

David Azriel *

Faculty of Industrial Engineering and Management, The Technion, Haifa 32000, Israel

ARTICLE INFO

Article history:
Received 28 June 2012
Received in revised form 26 February 2013
Accepted 13 May 2013
Available online 23 May 2013

Keywords:
Dynamic programming
Optimal designs
Phase I studies
Sequential methods

ABSTRACT

Phase I clinical trials are conducted in order to find the maximum tolerated dose of a given drug out of a set of doses, usually finite. In general, once a formal target function and a suitable probability structure are defined, optimization of sequential studies can theoretically be achieved using backward induction. This is a computationally heavy task and most of the proposed methods can be regarded as "myopic" strategies with respect to a certain loss function. Such designs are computationally feasible, but are not globally optimal. A Dynamic Programming algorithm that overcomes such computational difficulties is presented. It computes the global optimal designs with respect to different loss functions, which represent different purposes of a phase I study. Though the optimal designs provide an improvement over the standard designs, the improvement is not very significant. The expected loss of the global optimal design is about 3% (at most) less than in the "myopic" policies in the specific probability structure that have been considered. This is important as computationally feasible and simple algorithms provide designs that are very close to being optimal.

1. Introduction

It is generally believed that increasing the dosage of a certain drug increases both the probability of toxic reaction to the drug and its efficacy. It is therefore important to determine the maximum tolerated dose (MTD) of a given drug, that is, the highest dose that does not cause an unacceptable proportion of toxic reactions. This is particularly important in severe diseases, typically cancer, where strong and even lethal side effects may be present. Finding the MTD is one central aim of phase I studies, which are usually done by a sequential design for reasons of efficiency and for ethical reasons.

Numerous methods have been proposed and used for finding the MTD, see, e.g., the review papers of Rosenberger and Haines (2002) and Potter (2006), and a few are defined in the context of decision theory. The continual reassessment method (CRM) proposed by O'Quigley et al. (1990) is an improvement over earlier methods; see, e.g., O'Quigley and Chevret (1991). This method treats the next patient with the estimated MTD according to a one parameter working model. The escalation with overdose control (EWOC) method (Babb et al., 1998) is another popular method that uses, at each stage, an asymmetric loss function in which exceeding the MTD is overweighted.

These methods are regarded as "myopic strategies" in Rosenberger and Haines (2002):

> "The CRM and EWOC are, in some sense, myopic strategies in that assignment j is based on the jth stage posterior distribution. Myopic strategies are not always globally optimal; to determine globally optimal strategies, back-calculation (dynamic programming) is required, which is a computationally intensive determination of all possible strategies that could be taken each time a patient is assigned to a dose level". (p. 2764).

This work aims at overcoming such computational difficulties and constructing globally optimal policies.

* Tel.: +972 48294455; fax: +972 48295688.
E-mail addresses: davazriel@gmail.com, david.azriel@mail.huji.ac.il.

0167-9473/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.csda.2013.05.010

Whitehead and Brunier (1995) propose a method that minimizes the asymptotic variance of the estimator of the MTD, in which for each patient entering the trial, the dose that minimizes this variance at the current step is chosen. Again, this is a myopic strategy in which, however, the loss function pertains to the variance of the estimator of the MTD, rather than treatment at the MTD.

Haines et al. (2003) suggest constructing optimal Bayesian designs constrained to ethical limitations on the doses assigned at each stage. This is done in two stages: the first group's dose is assigned according to the constrained D-optimal design with respect to a uniform prior on the parameters. The second group follows the constrained D-optimal design with respect to the posterior of the first group observations. Recently, Drovandi et al. (2013) studied myopic Bayesian designs under different loss functions. They introduce an algorithm that efficiently computes the posterior expectation of the loss function and finds at each stage the action that minimizes this expectation.

In this work we consider a sequential design of N stages, that is, N subjects, in which a dose is assigned to each subject on the basis of the outcomes of previous subjects. The challenge is to find the optimal sequential design that minimizes a certain loss function. Leung and Wang (2002) suggest a Dynamic Programming (DP) algorithm for this problem, but computational complexities force them to use myopic policies. An approximate DP algorithm is introduced by Bartroff and Lai (2010, 2011), but a full DP that assigns subjects to treatments in an optimal way has yet to be developed. We show how the computational problems could be overcome and introduce a DP algorithm for a certain Bayesian model. We study optimal designs with respect to several loss functions and priors and compare them to the standard myopic designs.

The rest of the paper is organized as follows: in Section 2 the decision problem is defined and the DP equations are introduced. Section 3 presents the specific Bayesian framework considered here, in which computing the optimal design is feasible. We show how to compute the optimal design in an explanatory example in Section 4 and we study optimal designs with respect to different loss functions and several priors in Section 5. These optimal designs are compared to myopic policies in Section 6. A simulation study is presented in Section 7, and concluding remarks are given in Section 8.

2. Formulation of the problem

Let x be a dose of a certain drug and let y be a binary outcome, where $y = 1$ ($y = 0$) represents a toxic (non-toxic) response. Let $m(x) := P(y = 1|x)$ be the probability of a toxic response at dose x, where $m : \mathbb{R}^+ \to (0, 1)$ is an unknown strictly increasing function. The dose range D consists of only a few doses, $d_1 < d_2 < \cdots < d_K$, and the experimenter aims at finding the dose d_{j^*} which is closest to the known target toxicity level m^*, i.e., $j^* = \arg\min_j |m(d_j) - m^*|$; the dose d_{j^*} is called the MTD.

We consider sequential designs and denote by x_n and y_n the dose assigned to the n'th subject and his response, respectively. At stage n of the experiment the first $n-1$ results $\{x_1, y_1, x_2, y_2, \ldots, x_{n-1}, y_{n-1}\}$ are known and on the basis of this knowledge the experimenter decides which dose to assign to the n'th subject. A vector of the form $(x_1, y_1, \ldots, x_n, y_n)$ is called a *state* of the experiment, $n = 1, \ldots, N$. A *policy* τ (or a *design* or a *strategy*) is a function from all possible states to the dose space $D = \{d_1, \ldots, d_K\}$; the policy determines how the experiment will progress.

To define the minimization problem we assume that the response curve m is modeled by $m(x, \boldsymbol{\theta})$ where $\boldsymbol{\theta} \sim \pi$ and π is a given prior determined by the experimenter. At the end of the experiment an estimator $\hat{g}$ of $g(\boldsymbol{\theta})$ is obtained, where $g(\boldsymbol{\theta})$ may be the MTD or some other related dose, not necessarily in D. We consider a loss function with two components (called treatment and experimentation in Bartroff and Lai, 2010) L_1 and L_2 that relate to the two different purposes of phase I studies: L_1 corresponds to the estimation of $g(\boldsymbol{\theta})$ at the end of the trial, and L_2 represents, for each subject, the penalty for adverse reaction to a dose that is too high, or for a treatment at a dose that is too low and therefore non-efficient. Given L_1, L_2 and $\hat{g}$ one can compute the expected loss of each policy τ, which is

$$E_\tau \left[L_1\{\hat{g}(x_1, y_1, \ldots, x_N, y_N), g(\boldsymbol{\theta})\} + \sum_{i=1}^{N} L_2\{x_i, g(\boldsymbol{\theta})\} \right], \tag{1}$$

where the expectation is taken over the joint distribution of $(\boldsymbol{\theta}, x_1, y_1, \ldots, x_N, y_N)$. A policy τ that minimizes (1) is an *optimal policy*. Since the number of different policies is finite, an optimal policy exists for this problem.

Rather than a brute-force search over all possible policies, a natural way to find an optimal policy is by a backward induction scheme (usually called Dynamic Programming). First, for every state at the final stage we compute the following expected loss:

$$J_N(x_1, y_1, \ldots, x_N, y_N) := E[L_1\{\hat{g}(x_1, y_1, \ldots, x_N, y_N), g(\boldsymbol{\theta})\}|x_1, y_1, \ldots, x_N, y_N]; \tag{2}$$

note that the expectation does not depend on the policy. Now, for stage $N-1$, we compute for every vector $(x_1, y_1, \ldots, x_{N-1}, y_{N-1})$ and $x \in D$

$$\begin{aligned} \ell_{N-1}(x_1, y_1, \ldots, x_{N-1}, y_{N-1}, x) := {} & E[L_2\{x, g(\boldsymbol{\theta})\}|x_1, y_1, \ldots, x_{N-1}, y_{N-1}] \\ & + P(y_N = 1|x_1, y_1, \ldots, x_{N-1}, y_{N-1}, x_N = x) J_N(x_1, y_1, \ldots, x_{N-1}, y_{N-1}, x, 1) \\ & + P(y_N = 0|x_1, y_1, \ldots, x_{N-1}, y_{N-1}, x_N = x) J_N(x_1, y_1, \ldots, x_{N-1}, y_{N-1}, x, 0); \end{aligned} \tag{3}$$