Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 137 (2007) 2316–2327
journal of statistical planning and inference
www.elsevier.com/locate/jspi

Cumulative cohort design for dose-finding

Anastasia Ivanova[a,*], Nancy Flournoy[b], Yeonseung Chung[a]

[a] *Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7420, USA*
[b] *Department of Statistics, University of Missouri-Columbia, 146 Middlebush Building, Columbia, MO 65211, USA*

Received 26 September 2005; received in revised form 19 July 2006; accepted 26 July 2006
Available online 10 October 2006

Abstract

We introduce a new design for dose-finding in the context of toxicity studies for which it is assumed that toxicity increases with dose. The goal is to identify the maximum tolerated dose, which is taken to be the dose associated with a prespecified "target" toxicity rate. The decision to decrease, increase or repeat a dose for the next subject depends on how far an estimated toxicity rate at the current dose is from the target. The size of the window within which the current dose will be repeated is obtained based on the theory of Markov chains as applied to group up-and-down designs. But whereas the treatment allocation rule in Markovian group up-and-down designs is only based on information from the current cohort of subjects, the treatment allocation rule for the proposed design is based on the cumulative information at the current dose. We then consider an extension of this new design for clinical trials in which the subject's outcome is not known immediately. The new design is compared to the continual reassessment method.

Keywords: Phase I clinical trials; Group up-and-down designs; Toxicity studies; Delayed responses; Continual reassessment method; Markov chains; Maximum tolerated dose; Non-parametric designs

1. Introduction

This work was motivated by the task of designing a Phase I lymphoma study in which patients receive a regimen of a new radioimmunotherapy and one cycle of a proteasome inhibitor. Three dose levels of the proteasome inhibitor were chosen for the study. The dose limiting toxicity was defined as grade 4 thrombocytopenia (reduced platelet count) from which the patient does not recover to at least grade 1 toxicity within 12 weeks after the treatment of radioimmunotherapy. Since radioimmunotherapy was administered on day 8, the follow-up time for toxicity was 13 weeks after the initiation of treatment. The goal was to find the dose with the toxicity rate closest to 0.1.

Phase I clinical trials in oncology are conducted to estimate the maximum tolerated dose (MTD) of a new agent. Typically, the MTD is defined explicitly as the dose corresponding to a prespecified "target" toxicity rate Γ. Many designs have been developed for dose-finding in oncology. For detailed comprehensive reviews of these methods we refer the reader to chapters in several excellent books that appeared recently: Chapters 1–3 in the book edited by Crowley (2001), the entire book edited by Ting (2006), and the book edited by Chevret (2006). A special issue of the Journal of Statistical Planning and Inference on Adaptive Designs in Clinical Trials (2006) contains a number of articles on dose-finding; included are articles authored by O'Quigley, Stallard, Dragalin and Fedorov.

* Corresponding author.
E-mail address: aivanova@bios.unc.edu (A. Ivanova).

0378-3758/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2006.07.009

Dose-finding designs in oncology are usually classified as non-parametric and parametric. Non-parametric designs for dose-finding studies are intuitive, easy to explain and easy to implement. No difficult computations are involved. These non-parametric designs include biased coin designs (Durham and Flournoy, 1994, 1995), group up-and-down designs (Wetherill, 1963; Gezmu and Flournoy, 2006), the traditional or 3+3 design (Korn et al., 1994), A + B designs (Lin and Shih, 2001), and others. The 3 + 3 design is frequently used in dose-finding studies in oncology. The 3+3 design selects as the MTD the dose with the toxicity rate of about 0.2 (Reiner et al., 1999). For this reason the 3 + 3 design cannot be used in the lymphoma trial where target toxicity rate is 0.1. The generalization of the 3 + 3 design, the A + B designs, can target a wide range of quantiles (see Ivanova, 2006). However, the A + B designs will not work well in lymphoma trial because long follow-up is required to observe toxicities which causes the duration of the experiment to be excessive. The biased coin designs and group up-and-down designs, or simply *group designs*, can target a wide range of toxicity levels. But the group designs also may result in excessively long experiments when there are delayed responses, and the randomization required by the biased coin designs has not been well received by clinicians. In addition, group designs require a fixed number of subjects in each cohort and this is not practical for the lymphoma trial because subjects may be lost to follow-up.

All of the non-parametric procedures cited above only use information from the most recent subject cohort to determine the new treatment assignment. Our challenge was to develop a non-parametric procedure that is flexible and utilizes more information than just from the most recent subject cohort to make a treatment assignment, and yet retains operational simplicity and good operating characteristics. A few authors have suggested using the isotonic estimate of dose–toxicity curve in the decision rule of a dose-finding design (Yuan and Chappell, 2004; Conaway et al., 2004). Yuan and Chappell (2004), for example, suggested repeating the dose if the estimated toxicity rate at the current dose is within $(\Gamma, 2\Gamma)$ for trials with $\Gamma = 0.2$. This decision rule is rather ad hoc and might not be appropriate for other values of Γ.

In this paper, we develop a treatment allocation rule for a non-parametric dose-finding procedure that is theoretically justified. Our new method will be compared with its parametric competitor, the continual reassessment method (CRM) proposed by O'Quigley et al. (1990). The CRM can target a wide rage of toxicity levels and hence could, alternatively, have been used in the motivating lymphoma trial. A time-to-event version of the CRM (Cheung and Chappell, 2000) reduces the trial duration for studies with long follow-up such as the lymphoma trial.

We start in Section 2 by describing classical group designs and giving some theoretical results regarding the operating characteristics of these designs. These theoretical developments are then used to construct the new cumulative cohort design in Section 3. In Section 4, we extend the cumulative cohort design to accommodate delayed responses. We discuss designs for the lymphoma trial in Section 5, and we make some concluding remarks in Section 6.

2. Group designs

2.1. Properties of group designs

Let $D = \{d_1, \ldots, d_K\}$ denote the set of ordered dose levels. Let $F(d)$ be the probability of toxicity at dose d and let $q_j = F(d_j)$, $j = 1, \ldots, K$. Assume that $F(d)$ is an increasing function of d with $0 < F(d_1) < \cdots < F(d_K) < 1$. The goal is to find a dose in D with toxicity rate closest to the target probability of toxicity Γ. The group design is defined as follows.

Group design. Subjects are treated in cohorts of size s. Let $X(d_j)$ be the number of toxic responses in the *most recent* cohort assigned at dose level d_j. Then given d_j, $X(d_j) \sim \text{Bin}(s, q_j)$, where $\text{Bin}(s, q)$ denotes a binomial random variable with parameters s and q. Using notation from Gezmu and Flournoy (2006), let c_L and c_U be two integers such that $0 \leqslant c_L < c_U \leqslant s$. The next cohort of s subjects is assigned to

(i) dose level d_{j+1}, if $X(d_j) \leqslant c_L$;
(ii) dose level d_{j-1}, if $X(d_j) \geqslant c_U$;
(iii) dose level d_j, if $c_L < X(d_j) < c_U$.

Applying this rule when the dose is changing from d_1 or d_K might cause the dose assignment to be outside of the set D of dose levels. Thus for $j = 1$ or K, when the rule would cause a treatment to be outside of D, the current dose is repeated instead. This design is denoted by $UD(s, c_L, c_U)$.

Because the next treatment assignment only depends on the treatment and outcomes of the current cohort of subjects, and assuming the probability of toxicity only depends on the dose, the group design induces a Markov chain on D. Let $N_j(m)$ be the number of subjects assigned to dose d_j by the time m subjects have been assigned. Then, since the Markov chain is aperiodic and positive recurrent, $\lim_{m\to\infty}(N_1(m)/m, \ldots, N_K(m)/m) = (\pi_1, \ldots, \pi_K) \equiv \boldsymbol{\pi}$ in probability. The vector $\boldsymbol{\pi}$ is the *stationary distribution*, which is also called the *asymptotic treatment distribution*. Gezmu and Flournoy (2006) point out that the stationary distribution for a $UD(s, c_L, c_U)$ design exists and prove it is unimodal. Consider the equation

$$\Pr\{\mathrm{Bin}(s, \Gamma^*) \leqslant c_L\} = \Pr\{\mathrm{Bin}(s, \Gamma^*) \geqslant c_U\}. \tag{1}$$

The following theorem (with proof in Appendix A) states that the mode of the stationary distribution is the dose level that has toxicity rate close to Γ^*. We say that the design $UD(s, c_L, c_U)$ *targets* Γ^*, if Γ^* is the solution of (1).

Theorem. *The solution Γ^* of Eq.* (1) *exists and is unique. Furthermore, if there exists dose d_k such that $q_k = \Gamma^*$, the mode of the stationary distribution of $UD(s, c_L, c_U)$ is at d_k. If $q_{k-1} < \Gamma^* < q_k$ for some $k = 2, \ldots, K$, the mode is at d_{k-1} if $\Pr\{X(d_{k-1}) \geqslant c_U\} > \Pr\{X(d_k) \leqslant c_L\}$, and the mode is at d_k if $\Pr\{X(d_{k-1}) \geqslant c_U\} < \Pr\{X(d_k) \leqslant c_L\}$. If the two probabilities are equal, the mode spans d_{k-1} and d_k. If $\Gamma^* < q_1$, the mode is at d_1, and if $\Gamma^* > q_K$, the mode is at d_K.*

A consequence of this theorem is that the target quantile for any group design can be computed using formula (1). To obtain Γ^* for $UD(6, 0, 1)$, for example, using formulae for binomial probabilities Eq. (1) is $(1-\Gamma^*)^6 = 1-(1-\Gamma^*)^6$ with the solution $\Gamma^* = 1-(0.5)^{1/6} \approx 0.109$. For any s in $UD(s, 0, 1)$, similarly, the solution of (1) is $\Gamma^* = 1-(0.5)^{1/s}$. Alternatively, with a group size of six, (1) can be used to show that $UD(6, 0, 3)$ and $UD(6, 1, 2)$ both target quantiles near $\Gamma = 0.25$:

$$\Pr\{\mathrm{Bin}(6, 0.253) \leqslant 0\} \approx \Pr\{\mathrm{Bin}(6, 0.253) \geqslant 3\},$$

$$\Pr\{\mathrm{Bin}(6, 0.264) \leqslant 1\} \approx \Pr\{\mathrm{Bin}(6, 0.264) \geqslant 2\}.$$

We use formula (1) to develop the decision rule for the cumulative cohort design. But first we reformulate the group design so that decisions to change treatment levels are based on the proportions of responses rather than on absolute numbers. The advantage of this formulation is that it does not contain the group size s and hence will allow us to construct a generalization of the group design.

2.2. An alternative definition of the group design

Consider $UD(s, c_L, c_U)$, and let Γ^* be the solution of Eq. (1). Assume that the most recent cohort of subjects was assigned to d_j. Let $\hat{q} = X(d_j)/s$ be the proportion of toxicities in the most recent cohort. Since $c_L/s < \Gamma^* < c_U/s$ (see Appendix B), and if the desired Γ is close enough to Γ^* that $c_L/s < \Gamma < c_U/s$, one can find non-unique cutoff values Δ_L and Δ_U such that the decision rule for a group design can be written in a simpler form:

(i) if $\hat{q} \leqslant \Gamma - \Delta_L$, the next group of subjects is assigned to dose d_{j+1},
(ii) if $\hat{q} \geqslant \Gamma + \Delta_U$, the next group of subjects is assigned to dose d_{j-1},
(iii) if $\Gamma - \Delta_L < \hat{q} < \Gamma + \Delta_U$, the next group of subjects is assigned to dose d_j.

For example, according to $UD(6, 0, 3)$ the dose is increased if the proportion of toxicities is less than or equal to $0/6 = 0$, and decreased if the proportion is greater than or equal to $3/6 = 0.50$. Alternatively, this rule can be described in terms of $\Gamma = 0.25$ and $\Delta_L = \Delta_U = 0.25$ as: increase the dose if the proportion of toxicities is less than or equal to $\Gamma - \Delta_L = 0.25 - 0.25 = 0$, decrease if greater than or equal to $\Gamma + \Delta_U = 0.25 + 0.25 = 0.50$. Parameters Δ_L and Δ_U were computed as $\Delta_L = \Gamma - c_L/s = 0.25 - 0/6 = 0.25$ and $\Delta_U = c_U/s - \Gamma = 3/6 - 0.25 = 0.25$. Note that $\Delta_L = \Delta_U$ because the cutoff values are symmetric around the target.

Because the only possible values for the proportion of toxicities in cohorts of size $s = 6$ trials are 0, 1/6, 2/6, 3/6, 4/6, 5/6, and 1, to target a given Γ, the values that can be used for Δ_L and Δ_U are not unique and can be chosen within a certain interval; the $UD(6, 0, 3)$ can be described using any Δ_L in the interval $(1/12, 1/4]$ and any Δ_U in the interval $(1/12, 1/4]$ to target $\Gamma = 0.25$, and similarly, $UD(6, 1, 2)$ can be defined using any Δ_L in $(0, 1/12]$ and any Δ_U in

$(0, 1/12]$ to target $\Gamma = 0.25$. For any $UD(s, c_L, c_U)$ design, such functionally equivalent intervals, $(\Delta_L^1, \Delta_L^2]$ for Δ_L and $(\Delta_U^1, \Delta_U^2]$ for Δ_U, can be computed as

$$\Delta_L^1 = \max\left(0, \Gamma - \frac{c_L + 1}{s}\right), \quad \Delta_L^2 = \Gamma - \frac{c_L}{s},$$
$$\Delta_U^1 = \max\left(0, \frac{c_U - 1}{s} - \Gamma\right), \quad \Delta_U^2 = \frac{c_U}{s} - \Gamma. \tag{2}$$

To further simplify the definition of group designs, let $(\Delta_1, \Delta_2]$ be the intersection of $(\Delta_L^1, \Delta_L^2]$ and $(\Delta_U^1, \Delta_U^2]$, where

$$\Delta_1 = \max(\Delta_L^1, \Delta_U^1) \quad \text{and} \quad \Delta_2 = \min(\Delta_L^2, \Delta_U^2). \tag{3}$$

Now any value of Δ from $(\Delta_1, \Delta_2]$ can be used to express the design in even simpler form: the dose is increased if $\hat{q} \leqslant \Gamma - \Delta$, decreased if $\hat{q} \geqslant \Gamma + \Delta$, and repeated if $\Gamma - \Delta < \hat{q} < \Gamma + \Delta$. For example, to target $\Gamma = 0.25$, $UD(6, 0, 3)$ and $UD(6, 1, 2)$ can be described using any Δ in the interval $(1/12, 1/4]$ and $(0, 1/12]$, respectively.

2.3. Comparing group designs

There can be several group designs with the same group size s to use for the same target toxicity rate. We have seen that both $UD(6, 1, 2)$ with $\Gamma^* \approx 0.264$, and $UD(6, 0, 3)$ with $\Gamma^* \approx 0.253$ can be used when the target toxicity rate is $\Gamma = 0.25$. Alternatively, $UD(6, 2, 4)$ and $UD(6, 1, 5)$ both target $\Gamma^* = 0.5$. Guidelines for selecting the best decision rule given a fixed cohort size s are needed, but this is not our goal. Instead of deciding to increase the dose based on the current cohort, we propose in Section 3 to use the cumulated data at the current dose. In the proposed design, we will use $s = s(n)$ to denote the increasing sample size at the current dose and use a group up-and-down rule that changes with the increasing "group" size at the current dose. Thus, for our purposes, we need comparisons of the group designs for much greater "cohort" sizes than would be used if we were actually going to use a group design. Also, if we were actually going to use a group design, we might well focus on starting with the lowest dose. However, in the proposed design, we are effectively continually reinitializing the group design as the clinical trial proceeds. Therefore, we need to compare the performance of the group designs across different starting doses.

Because a group design induces a Markov chain on D, we look to Markov chain theory for measures of comparison. The performance of designs based on Markov chains has been evaluated on the basis of the "peakedness" of the stationary distribution and the rate of convergence to the stationary distribution (Giovagnoli and Pintacuda, 1998). For two designs with stationary distributions $\tilde{\boldsymbol{\pi}}$ and $\boldsymbol{\pi}$ with the same mode, they say $\tilde{\boldsymbol{\pi}}$ is "more peaked", if it grows more quickly to the left of the mode and decreases more quickly to the right of the mode. That is, if the mode is at j, $\tilde{\pi}_i/\tilde{\pi}_{i-1} \geqslant \pi_i/\pi_{i-1} \geqslant 1$ for $i = 1, \ldots, j$, and $\tilde{\pi}_{i+1}/\tilde{\pi}_i \leqslant \pi_{i+1}/\pi_i \leqslant 1$ for $i = j, \ldots, K$. The rate of convergence is determined by the second largest eigenvalue in the absolute value of the transition matrix (*cf.* Mira, 2001).

Both the peakedness and the rate of convergence to a stationary distribution for a group design depend on the dose–toxicity relationship. Thus we suggest comparing two group designs that have target toxicity rates near Γ in a trial with K dose levels by assuming a *large* number (for example 10,000) of different toxicity scenarios with K dose levels, and summarizing the results. To simplify the comparison, scenarios are constructed in such a way that there is a dose level with toxicity rate exactly equal to Γ. For each dose–toxicity scenario and each pair of designs, we could compute the transition probability matrix; then compute the stationary distribution and the second largest eigenvalue of the transition probability matrix in the absolute value. Declare one design to be better than the other for a particular dose–toxicity scenario if its stationary distribution is more peaked and it converges faster to the stationary distribution (see Bortot and Giovagnoli (2005), for more details).

Unfortunately, for the majority of group designs that we have compared, the design that converges faster has a stationary distribution that is less peaked. Because of this and because our interest is in moderate sample size performance, we use a finite sample measure to compare designs instead. Such finite measure computed for every design can serve as a useful approximation for how the design will perform as a part of the cumulative cohort procedure proposed in Section 3. For each dose–toxicity scenario, we compare designs based on the expected proportion of subjects assigned to the MTD in a trial with n cohorts. The initial dose is considered equally likely to be any of the K possible doses. The measure is calculated (not simulated) by averaging the elements of the matrix $(\boldsymbol{I} + \boldsymbol{P} + \cdots + \boldsymbol{P}^n)/(n+1)$ in the column corresponding to the location of the MTD, where $\boldsymbol{P}$ is design's transition matrix for the scenario (see Appendix A) and

$\boldsymbol{I}$ is an identity matrix. This produces an average over all possible "starting" doses. Then we average this measure over all 10,000 dose–toxicity scenarios and call it the *average expected proportion allocated to the MTD*. This summary design characteristic is then used to select group procedures for the cumulative cohort design.

The choice of the total number of groups, n, is guided by the number of doses K. One wants $n \geqslant K-1$ so that the design is able to move from one dose level to the other levels even if these levels are far apart, for example, from d_1 to d_K. We took $n = K$. The scenarios were generated with the MTD position being equally likely at any of K doses. That is, about $1/K$ of all scenarios had the MTD at d_1, $1/K$ of all scenarios at d_2, etc. The toxicity rates at doses below the MTD were generated as ordered uniforms on $(0, \Gamma)$, and above the MTD as ordered uniforms on $(\Gamma, q_{\max})$ with $\Gamma < q_{\max} \leqslant 1$. The value $q_{\max}$ reflects the investigators' belief about the range of possible toxicity rates in a trial. For example, if the MTD location for a scenario was d_2, the toxicity rate at d_1 was generated as the uniform random variable on $(0, \Gamma)$, and toxicity rates at $(d_3, \ldots, d_K)$ were generated as $K-2$ ordered uniform random variables on $(\Gamma, q_{\max})$. We compared group designs when $K = 6$. The choice of $K = 6$ reflects the number of doses in many dose-finding experiments (Buoen et al., 2005).

A large number of dose-finding trials in oncology target quantiles in the neighborhood of $\Gamma = 0.25$ (Lin and Shih, 2001; Thall and Lee, 2003). For each s, we constructed all possible group designs that target quantiles in the neighborhood of $\Gamma = 0.25$. Then for each design, we computed the average expected proportion allocated to the MTD, and selected the design with the largest value of this measure. For example, $UD(6, 0, 3)$ and $UD(6, 1, 2)$ are the only two group designs with $s = 6$ and targets close to 0.25. The average expected proportions allocated to the MTD using $UD(6, 0, 3)$ and $UD(6, 1, 2)$ were 0.28 and 0.26, respectively. Hence $UD(6, 0, 3)$ is said to be the *best group design* for $\Gamma = 0.25$ and $s = 6$. The best group designs to use with target $\Gamma = 0.25$ and each $s = 3, \ldots, 25$ are $UD(3, 0, 2)$, $UD(4, 0, 2)$, $UD(5, 0, 2)$, $UD(6, 0, 3)$, $UD(7, 1, 3)$, $UD(8, 1, 3)$, $UD(9, 1, 4)$, $UD(10, 1, 4)$, $UD(11, 1, 4)$, $UD(12, 1, 5)$, $UD(13, 2, 5)$, $UD(14, 2, 5)$, $UD(15, 2, 6)$, $UD(16, 2, 6)$, $UD(17, 2, 6)$, $UD(18, 2, 7)$, $UD(19, 3, 7)$, $UD(20, 3, 7)$, $UD(21, 3, 8)$, $UD(22, 3, 8)$, $UD(23, 3, 8)$, $UD(24, 4, 8)$, and $UD(25, 4, 9)$.

3. Cumulative cohort design

3.1. Cumulative cohort design

The cumulative cohort design has a treatment allocation rule similar to that of the group design except it is based on the information from all subjects that have been assigned to the dose, not just the information from the most recent group of subjects. Let $s(n)$ be the number of subjects assigned to the current dose after a total of n subjects have been treated, and let $\Delta_{s(n)}$ be a value from the interval $(\Delta_1, \Delta_2]$ given by (2) and (3) for the best group design with group size $s(n)$. The cumulative cohort design uses the decision rule given in Section 2.2, but takes the current "cohort" size to be the cumulative number of subjects treated at the current dose and uses $\Delta = \Delta_{s(n)}$. Therefore, the treatment allocation decision of the cumulative cohort design is based on the cumulative toxicity data from all $s(n)$ subjects.

Fig. 1 presents the intervals $(\Delta_1, \Delta_2]$ of possible values Δ found by applying Eqs. (2) and (3) to the best group design that targets quantiles around $\Gamma = 0.25$ for each $s = 3, \ldots, 25$. For example, design $UD(6, 0, 3)$ is represented by the interval (1/12,1/4] in Fig. 1, because it can be defined by $\Gamma = 0.25$ and any Δ in this interval as described in Section 2.2. For the best designs used in Fig. 1 with $3 \leqslant s \leqslant 23$, the intervals $(\Delta_1, \Delta_2]$ have a non-empty intersection which is the interval $(1/12, 7/76] \approx (0.083, 0.092]$. Therefore, all these designs can be conveniently described by using a single Δ which can be any Δ from $(1/12, 7/76]$, for example, $\Delta = 0.09$. The value $\Delta = 0.09$ does not describe the best group design for $s = 24$: $\Delta = 0.09$ corresponds to the group design $UD(24, 3, 9)$, whereas the best group design is $UD(24, 4, 8)$. However, the average expected proportion allocated to the MTD is 0.357 for $UD(24, 3, 9)$ which is very close to that of 0.358 for $UD(24, 4, 8)$.

Fig. 2 displays the average expected proportion allocated to the MTD for the best group designs (which is equivalent to using $\Delta = 0.09$ for sample sizes from 3 to 23) and for an ad hoc group designs with $\Delta = 0.01$. The parameter $\Delta = 0.01$ defines the designs where, for moderate sample sizes, the dose is repeated if its estimated toxicity rate is equal to Γ and changed otherwise. Fig. 1 shows that for $\Gamma = 0.25$ only two of the group designs defined by $\Delta = 0.01$ are the best group designs (see group sizes four and eight). Except for these two group designs, the average expected proportion allocated to the MTD is substantially higher for the best group designs compared to designs with $\Delta = 0.01$ (Fig. 2). This should translate into a gain in the number of subjects allocated to the MTD when the best designs ($\Delta = 0.09$) are used in the proposed cumulative cohort procedure rather than using some arbitrary designs such as defined by parameter $\Delta = 0.01$.

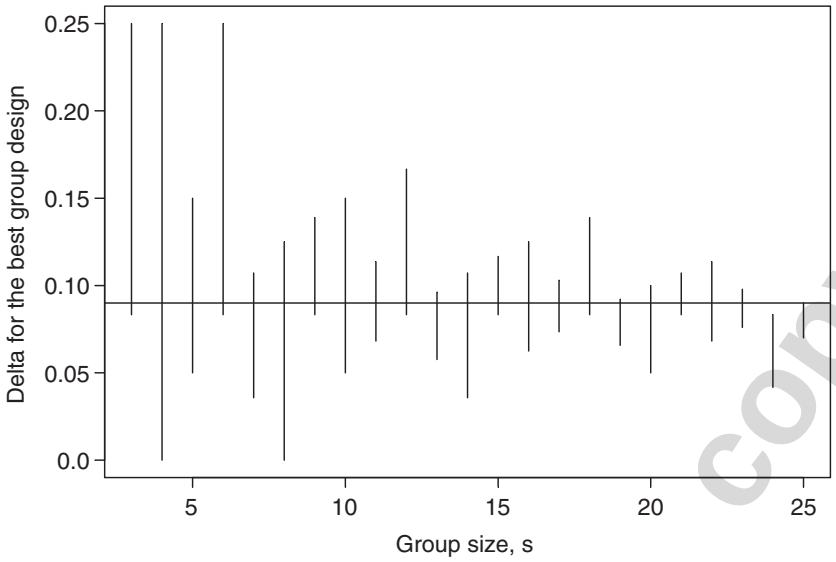

Fig. 1. The best group designs with $s = 3, \ldots, 25$ to target $\Gamma = 0.25$. The designs are $UD(3, 0, 2)$, $UD(4, 0, 2)$, $UD(5, 0, 2)$, $UD(6, 0, 3)$, $UD(7, 1, 3)$, $UD(8, 1, 3)$, $UD(9, 1, 4)$, $UD(10, 1, 4)$, $UD(11, 1, 4)$, $UD(12, 1, 5)$, $UD(13, 2, 5)$, $UD(14, 2, 5)$, $UD(15, 2, 6)$, $UD(16, 2, 6)$, $UD(17, 2, 6)$, $UD(18, 2, 7)$, $UD(19, 3, 7)$, $UD(20, 3, 7)$, $UD(21, 3, 8)$, $UD(22, 3, 8)$, $UD(23, 3, 8)$, $UD(24, 4, 8)$, and $UD(25, 4, 9)$. Each design is represented by a range of Δ values to use in the simplified definition of the design. The ranges of Δ were obtained for each design using formulae (2) and (3). A possible common value of $\Delta = 0.09$ that can be used to define the best design for $3 \leqslant s \leqslant 23$ is shown by the horizontal line.

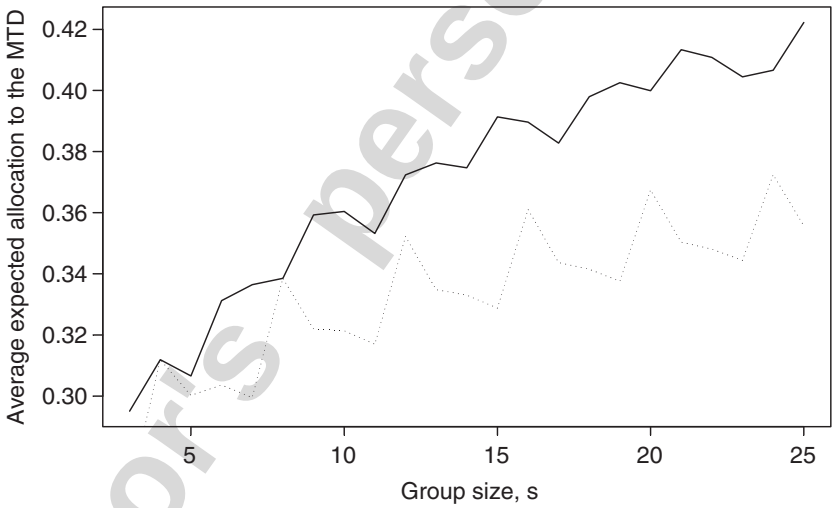

Fig. 2. Average (over 10,000 scenarios) expected allocation to the MTD for the best group designs (solid line) and for the group designs with $\Delta = 0.01$ (dotted line) for each $s = 3, \ldots, 25$. The designs were constructed to target $\Gamma = 0.25$.

Note that the average expected proportion allocated to the MTD plotted in Fig. 2 is affected not only by a group size but also by the value Γ^*, the design's target quantile. This explains why the average expected proportion can be sometimes lower for designs with higher group size, for example, $UD(4, 0, 2)$ and $UD(5, 0, 2)$ may be used in practice to target $\Gamma = 0.25$. But design $UD(4, 0, 2)$ actually targets $\Gamma^* \approx 0.266$ whereas $UD(5, 0, 2)$ actually targets $\Gamma^* \approx 0.216$ which is further from the desired $\Gamma = 0.25$, and this results in lower average expected proportion for $UD(5, 0, 2)$.

To summarize, the decision rule for the cumulative cohort design for $\Gamma = 0.25$ and moderate sample sizes (i.e., $s(n) \leqslant 23$ at a dose) can be written in a very simple form with $\Delta = 0.09$: that is, increase the dose if the proportion of toxic responses at the current dose is lower than or equal to 0.16, decrease the dose if higher than or equal to 0.34, and repeat the dose otherwise. Similarly, we calculated sequences of the best designs for trials with six doses for a range of other values Γ. Values of Δ that correspond to most of the best group designs for a given Γ for moderate sample sizes

Table 1
Recommended $\Delta = \Delta_{s(n)}$ to use in the cumulative cohort design to target Γ for moderate sample sizes with $s(n) \leqslant 20$ and six dose levels

Γ	Δ
0.10	0.09
0.15	0.09
0.20	0.09
0.25	0.09
0.30	0.10
0.35	0.10
0.40	0.12
0.45	0.13
0.50	0.13

are presented in Table 1. For example, for $\Gamma = 0.1$, $\Delta = 0.09$ defines all the best group designs for $s(n)$ up to 20. For $\Gamma = 0.5$, $\Delta = 0.13$ yields the best group designs for $s(n)$ up to 21, except when $s = 11$.

In toxicology studies where $\Gamma = 0.5$, $UD(1, 0, 1)$ is often used. However, the cumulative cohort design with $\Delta = 0.13$ has better targeting capabilities than $UD(1, 0, 1)$. It is also easy to understand and implement, just increase the dose if the current proportion of toxicities is greater than or equal to 0.63, decrease the dose if it is less than or equal to 0.37 and repeat the current dose otherwise. We recommend this design for toxicology studies.

We compared group designs when $K = 6$, $q_{\max} = 0.6$ for $\Gamma < 0.40$, and $q_{\max} = 1.0$ for $\Gamma \geqslant 0.40$. If the number of dose levels in a trial is very different from $K = 6$ or $q_{\max}$ differs from the values we considered, it is wise to repeat the computations using the desired K and $q_{\max}$. Our computations for the same test model and $n = K + 1$ and $K + 2$, however, yielded the same Table 1.

The cumulative cohort approach allows more flexibility because subjects can be assigned in groups of any desired size, possibly varying and including one at a time. For example, the smallest group size that yields Γ^* close to 0.1 as in the lymphoma trial is $s = 6$ with $\Gamma^* \approx 0.109$. Yet, the cumulative cohort design can be used with any group size.

When a trial is completed, we recommend using isotonic regression to estimate the MTD. It has been shown recently that isotonic regression estimator performs better than other estimators (Stylianou and Flournoy, 2002; Ivanova et al., 2003). Stylianou and Flournoy (2002) described in detail how to compute the isotonic estimator after a dose-finding trial. In short, at the end of the trial consider dose levels d_j such that $N_j(n) > 0$. Estimate q_j by $\hat{q}_j = X_j(n)/N_j(n)$ for such d_j. In the case $\hat{q}_j$ are not non-decreasing with d, adjust $\hat{q}$'s by using the *pool adjacent violators algorithm* (*cf.* Barlow et al., 1972). The dose with isotonic toxicity rate estimate closest to Γ is the estimated MTD. If there are two or more such doses, the lowest of the doses is chosen except for the case where all such doses have estimated toxicity rate lower than Γ, in which case the highest of the doses is chosen.

3.2. Simulation study

Whereas, for a given dose–response scenario, group designs can be compared theoretically among themselves (asymptotically or using the finite sample approach described in Section 2.3), in order to compare the cumulative cohort design to the CRM we must resort to simulations. Thall and Lee (2003) recently compared the 3 + 3 design with the CRM in the trial with $\Gamma = 0.25$ where subjects were assigned in groups of 3. We simulated the cumulative cohort design in all six scenarios considered in Thall and Lee (2003) and show the results of these simulations in Table 2. Scenarios 1–3 are derived from a logistic relation between dose and toxicity. Scenarios 4–6 follow the model $Q(d_i) = b_i^{\theta}$ with $\{b_1, \ldots, b_6\} = \{0.15, 0.25, 0.40, 0.60, 0.75, 0.85\}$. This later model was the underlying working model used in the CRM (Thall and Lee, 2003) in which the parameter θ followed normal prior with mean 0 and $\mathrm{Var}(\theta) = 1.8$. The starting dose was d_2 and the total sample size was 36. The results for the CRM are reproduced from Thall and Lee (2003). Results for the cumulative cohort design were obtained by simulations with 4000 repetitions.

The methods were compared with respect to the proportion of subjects assigned to each dose level, the average toxicity rates observed, and the percentage of times the MTD was correctly selected. Considering the allocation of subjects in this simulation study, the CRM was more conservative and assigned fewer subjects on average to the doses higher than the

Table 2
Proportion of times each dose is recommended as the MTD and the average number of subjects allocated to each dose for the cumulative cohort design (CCD) with $\Delta = 0.09$ and the CRM

	Proportion recommended						Allocation					
	d_1	d_2	d_3	d_4	d_5	d_6	d_1	d_2	d_3	d_4	d_5	d_6
Scenario 1												
$F(d_j)$	0.25	0.53	0.69	0.79	0.84	0.88						
CCD	**0.96**	0.04	0.00	0.00	0.00	0.00	25.0	10.5	0.5	0.0	0.0	0.0
CRM	**0.95**	0.05	0.00	0.00	0.00	0.00	29.1	6.5	0.4	0.0	0.0	0.0
Scenario 2												
$F(d_j)$	0.01	0.09	0.26	0.47	0.64	0.76						
CCD	0.00	0.16	**0.76**	0.08	0.00	0.00	0.2	9.5	20.2	5.7	0.4	0.0
CRM	0.00	0.16	**0.79**	0.05	0.00	0.00	0.9	10.0	21.7	3.2	0.2	0.0
Scenario 3												
$F(d_j)$	0.00	0.01	0.05	0.13	0.24	0.36						
CCD	0.00	0.00	0.01	0.22	**0.54**	0.23	0.0	3.2	4.7	9.7	12.1	6.2
CRM	0.00	0.00	0.06	0.41	**0.47**	0.06	0.1	3.3	6.7	12.8	11.6	1.4
Scenario 4												
$F(d_j)$	0.15	0.25	0.40	0.60	0.75	0.85						
CCD	0.16	**0.68**	0.16	0.00	0.00	0.00	5.4	21.6	8.4	0.8	0.0	0.0
CRM	0.22	**0.64**	0.14	0.00	0.00	0.00	9.7	18.8	7.0	0.5	0.0	0.0
Scenario 5												
$F(d_j)$	0.00	0.02	0.08	0.24	0.45	0.63						
CCD	0.00	0.00	0.10	**0.75**	0.14	0.00	0.0	3.5	7.8	17.9	6.2	0.6
CRM	0.00	0.00	0.20	**0.74**	0.06	0.00	0.2	3.6	11.0	17.8	3.4	0.0
Scenario 6												
$F(d_j)$	0.00	0.00	0.02	0.12	0.30	0.50						
CCD	0.00	0.00	0.00	0.29	**0.66**	0.05	0.0	3.0	3.7	11.3	14.7	3.3
CRM	0.00	0.00	0.00	0.46	**0.53**	0.01	0.0	3.0	4.3	15.3	12.8	0.6

Subjects were assigned three at a time. The total sample size is 36. The results for the CRM are reproduced from Thall and Lee (2003). Results at the MTD are in bold.

MTD. The average proportions of toxicities for each of the six scenarios were 0.34, 0.25, 0.19, 0.28, 0.23 and 0.21 for the cumulative cohort design and 0.31, 0.23, 0.15, 0.26, 0.19 and 0.17 for the CRM. These toxicity proportions are slightly lower for the CRM; they are below the target 0.25 for both designs in four scenarios and above the target for two scenarios. In terms of correct selection probabilities, the performance of the cumulative cohort design and the CRM is similar (Table 2): the cumulative cohort design performs better to some extent in five scenarios, and slightly worse in one scenario. It is remarkable to see comparable or better estimation precision and only slightly worse targeting of the cumulative cohort design compared to the CRM in scenarios 4–6 since these scenarios were derived from the underlying model assumed in the CRM simulations. Furthermore, reducing the total sample size from 36 to 24, the cumulative cohort design still performs pretty well, correctly selecting the MTD in 0.91, 0.65, 0.49, 0.62, 0.62, 0.58 of the simulations for the six scenarios, respectively.

To verify that the better performance of a group design transfers to better performance of the cumulative cohort design, we simulated the cumulative cohort design with the suboptimal parameter $\Delta = 0.01$ for scenarios 1–6. The correct MTD was selected in 0.96, 0.74, 0.55, 0.58, 0.72, 0.66 of the simulations for the six scenarios, respectively, which is only very slightly worse than was found for the cumulative cohort design that uses the best decision rules. The design with $\Delta = 0.01$ allocated 3–5 less subjects on average to the MTD compared to the cumulative cohort design with $\Delta = 0.09$. For example, in Scenario 4 the cumulative cohort design with $\Delta = 0.01$ allocated 15.8 subjects to the MTD versus 21.5 for the cumulative cohort design with $\Delta = 0.09$. The cumulative cohort design with suboptimal parameter $\Delta = 0.15$ was significantly inferior in targeting and estimating the MTD (data are available from the authors). The group design $UD(3, 0, 1)$ with $\Gamma^* \approx 0.206$ did not perform as well as the cumulative cohort design either.

4. Time to event modifications of the cumulative cohort design

If the follow-up time is long, the cumulative cohort design and the CRM may result in dose-finding trials of excessive duration. Cheung and Chappell (2000) proposed a time-to-event modification of the CRM, called TITE-CRM, for dose-finding trials requiring long follow-up times. This method has been used in radiation therapy trials (e.g., Muler et al., 2004). Let x_i be the dose level received by subject i, $x_i \in D$, and $y_i = 1$ if the ith subject had toxicity and 0 otherwise. In the original CRM (O'Quigley et al., 1990), the calculation of posterior mean of θ at the time when $(n+1)$th subject enters the trial is based on the likelihood

$$L_n(\theta) = \prod_{i=1}^{n} F(x_i, \theta)^{y_i} \{1 - F(x_i, \theta)\}^{1-y_i},$$

where $F(x_i, \theta) = b_{x_i}^{\theta}$ and $(b_1, \ldots, b_K)$ is a set of positive constants. In clinical trials that require long follow-up times, the toxicity rate at dose x_i is defined as the probability of observing toxicity at x_i during a time period of length T after initiation of therapy. Data for the ith subject, $i = 1, \ldots, n$, when $(n+1)$th subject is assigned to a treatment, consist of dose x_i, toxicity indicator $y_{i,n}$ and the time $u_{i,n}$ that has elapsed from the time of the ith subject's treatment assignment to the time of the $(n+1)$th subject's treatment assignment.

For TITE-CRM, Cheung and Chappell (2000) suggested the weighted likelihood

$$\tilde{L}_n(\theta) = \prod_{i=1}^{n} \{w_{i,n} F(x_i, \theta)\}^{y_i} \{1 - w_{i,n} F(x_i, \theta)\}^{1-y_i},$$

where $w_{i,n}$ is the weight assigned to the ith observation prior to the entry of the $(n+1)$th subject. For example, setting $w_{i,n} = \min(u_{i,n}/T, 1)$ reflects an assumption that the density of time to toxicity is flat in $(0, T)$. Other choices for the weight function can be considered (see Cheung and Chappell, 2000). Let $(\hat{q}_{1,n}, \ldots, \hat{q}_{K,n})$ be the vector that maximizes the likelihood

$$\tilde{L}_n(\theta) = \prod_{i=1}^{n} \{w_{i,n} q_{x_i,n}\}^{y_{i,n}} \{1 - w_{i,n} q_{x_i,n}\}^{1-y_{i,n}}$$

over all vectors such that $\hat{q}_{j,n} \in [0, 1]$, $j = 1, \ldots, K$, and $(\hat{q}_{1,n}, \ldots, \hat{q}_{K,n})$ are non-decreasing.

Motivated by the TITE-CRM, we now define the TITE version of the cumulative cohort design (CCD) or TITE-CCD.

TITE-CCD design. Assume that subject n was assigned to dose $x_n = d_j$. When the $(n+1)$th subject arrives, compute estimates of the toxicity rates, $\hat{q}_{1,n}, \ldots, \hat{q}_{K,n}$, as described below. Set $\hat{q} = \hat{q}_{j,n}$ and find Δ in Table 1.

(i) If $\hat{q} \leqslant \Gamma - \Delta$, the next subject is assigned to dose d_{j+1}.
(ii) If $\Gamma + \Delta \leqslant \hat{q}$, the next subject is assigned to dose d_i where $i = \max\{k : \hat{q}_{k,n} \leqslant \Gamma + \Delta\}$.
(iii) If $\Gamma - \Delta < \hat{q} < \Gamma + \Delta$, the next subject is assigned to dose d_j.

Appropriate adjustments are made at the lowest and highest doses.

To account for long follow-up time, the cumulative cohort design is extended in several ways. First, the toxicity proportions are computed using data from all subjects including subjects still under follow-up. Then the toxicity proportions are isotonized to provide estimates $\hat{q}_{1,n} \leqslant \cdots \leqslant \hat{q}_{K,n}$. Note that the dose can be decreased by more than one level because at the time of new assignment the estimated toxicity rates may have changed at doses other than the current dose. In the cumulative cohort design, by construction, all doses below the current dose have estimated toxicity rate no higher than $\Gamma + \Delta$. This is no longer the case in TITE-CCD. Therefore, (ii) and (iii) in the definition of TITE-CCD ensure that the estimated toxicity rate at the dose assigned to the next subject does not exceed $\Gamma + \Delta$.

When the TITE-CRM or TITE-CCD is used, a start-up procedure that uses complete follow-up times is needed to prevent the possibility of escalating to doses with very high rates of toxicity. As pointed out by Cheung and Chappell (2000),

the TITE-CRM might not perform adequately if the accrual is too rapid compared to the follow-up time. This is true for TITE-CCD as well and should be remembered when TITE-type designs are used in practice.

5. Design for lymphoma trial

The target quantile in the lymphoma trial was $\Gamma = 0.1$. A start-up procedure with a group size of 4 is used because it will bring the treatment allocations to doses with toxicity rate below $1 - 0.5^{1/4} \approx 0.159$ on average (Ivanova, 2006). During the start-up patients are assigned in groups of four at escalating doses until the first toxicity is seen. We compared the cumulative cohort and CRM designs for three dose–toxicity scenarios with three dose levels (Table 3). The three scenarios represent possible dose–toxicity curves in the lymphoma trial. The underlying working model for the CRM was $\{b_1, b_2, b_3\} = \{0.15, 0.20, 0.35\}$ with the same prior for parameter θ as in Thall and Lee (2003). Both designs were used with the start-up rule with groups of four, and patients assigned one at a time afterwards.

From Table 3, one can see that the cumulative cohort design performs better in scenario 2 and the CRM performs better in scenario 3. The CRM does not perform well in scenario 2 because the small number of toxicities at d_1 and d_2 makes the estimated toxicity at d_3 much lower than the observed proportion. This leads to d_3 being chosen as the MTD instead of d_2 in 26% of the runs for the CRM compared to only 10% for the cumulative cohort design. For example, if the trial results in 4, 15, and 5 assignments to each of the three doses with 0, 0, 2 toxicities, the MTD estimate according to the CRM is dose d_3.

Since the follow-up time for toxicity is 13 weeks and expected accrual rate in the lymphoma trial was about 1 patient per month, a time-to-event design will make the trial shorter compared to waiting until each follow-up is completed. The results of our comparison of the TITE-CRM and TITE-CCD are presented in Table 3 for trials with sample size of 24, the sample size that was guided by the amount of the resources available for the trial. According to the simulations (Table 3), sample size of 24 patients yielded satisfactory precision of the MTD estimates in case when the MTD is in the range of the three doses studied.

Table 3
Proportion of times each dose is recommended as the MTD and the average number of subjects allocated to each dose in a trial with $\Gamma = 0.10$ and the sample size of 24 for the cumulative cohort design (CCD), CRM, TITE-CCD, TITE-CRM

	Proportion recommended			Allocation		
	d_1	d_2	d_3	d_1	d_2	d_3
Scenario 1						
$F(d_j)$	0.10	0.35	0.60			
CCD	**0.95**	0.05	0.00	17.0	6.3	0.7
CRM	**0.94**	0.06	0.00	19.3	4.2	0.5
TITE-CCD	**0.93**	0.07	0.00	16.8	6.5	0.7
TITE-CRM	**0.91**	0.09	0.00	19.4	4.1	0.5
Scenario 2						
$F(d_j)$	0.01	0.04	0.30			
CCD	0.06	**0.84**	0.10	5.4	10.2	8.4
CRM	0.06	**0.67**	0.26	5.9	11.7	6.5
TITE-CCD	0.06	**0.83**	0.16	7.0	8.7	8.3
TITE-CRM	0.07	**0.60**	0.33	6.1	11.8	6.1
Scenario 3						
$F(d_j)$	0.01	0.04	0.08			
CCD	0.06	0.27	**0.67**	5.4	7.0	11.6
CRM	0.03	0.20	**0.77**	5.1	6.9	11.9
TITE-CCD	0.05	0.26	**0.69**	5.4	6.8	11.8
TITE-CRM	0.02	0.19	**0.78**	5.2	7.2	11.6

Results at the MTD are in bold.

6. Discussion

We described a design where the dose is repeated if the estimated toxicity rate at the dose is close to target, and changed otherwise. A window Δ to use in the treatment allocation rule for several target quantiles was specified and justified. We report the results when the decision rule uses the observed sample proportion to estimate the toxicity rate. Using the isotonic estimate of toxicity in the decision rule (as opposed to the final analysis) did not lead to improvements in the design performance in our simulation study (data are available from the authors). On the other hand, a cumulative cohort decision rule can offer benefit in trials where toxicity rates are estimated using isotonic methods as in Yuan and Chappell (2004) and Conaway et al. (2004). We recommend using the new design with the fixed sample size. Current experience is that designs that assign more subjects to the target dose, such as the CRM, do not perform well with a stopping rule (*cf.* Rosenberger and Haines, 2002).

Acknowledgments

The authors thank Dr. Thomas Shea from the Lineberger Comprehensive Cancer Center for the permission to use the process of developing the lymphoma trial as the motivating example. The authors thank the referees for their constructive comments.

Appendix A. Proof of Theorem

Let α_j, β_j, and γ_j denote the probabilities to decrease the dose from d_j to d_{j-1}, to repeat d_j, or increase the dose from d_j to d_{j+1} in $UD(s, c_L, c_U)$. Here $\alpha_j + \beta_j + \gamma_j = 1$ for $j \in \{1, \ldots, K\}$ are the elements of jth row of transition matrix P, with β_j being a diagonal element, and α_j, and γ_j being to the left and to the right of β_j. These probabilities can be computed as follows:

$$\alpha_1 = 0, \ \beta_1 = 1 - \gamma_1, \quad \gamma_1 = \Pr\{\mathrm{Bin}(s, q_1) \leqslant c_L\},$$

$$\alpha_j = \Pr\{\mathrm{Bin}(s, q_j) \geqslant c_U\}, \quad \beta_j = \Pr\{c_L < \mathrm{Bin}(s, q_j) < c_U\}, \quad \gamma_j = \Pr\{\mathrm{Bin}(s, q_j) \leqslant c_L\},$$

$$\alpha_K = \Pr\{\mathrm{Bin}(s, q_K) \geqslant c_U\}, \quad \beta_K = 1 - \alpha_K, \ \gamma_K = 0,$$

where $j \in \{2, \ldots, K-1\}$. The stationary distribution $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K)$ can be obtained by solving the balance equations, $\pi_j = \pi_{j-1}\gamma_{j-1} + \pi_j\beta_j + \pi_{j+1}\alpha_{j+1}$, $j \in \{1, \ldots, K\}$ (here for convenience $\gamma_0 = \alpha_{K+1} = 0$). The solution is

$$\pi_j = \prod_{i=1}^{j} \lambda_i, \quad \lambda_1 = \left(1 + \sum_{j=2}^{K} \prod_{i=2}^{j} \lambda_i\right)^{-1}, \quad \lambda_i = \frac{\gamma_{i-1}}{\alpha_i},$$

where $j \in \{2, \ldots, K\}$. Let $\lambda_2 \geqslant 1$. Gezmu and Flournoy (2006) showed that γ_j decreases with j while α_j increases with j, so similar to Durham and Flournoy (1994), the stationary distribution is log-concave with the mode at d_k if $\lambda_k > 1$, or the mode that spans d_{k-1} and d_k if $\lambda_k = 1$. If $\lambda_2 < 1$, the mode is at dose d_1. If $\lambda_K > 1$, the mode is in d_K. Denote $\alpha(q) = \Pr\{\mathrm{Bin}(s, q) \geqslant c_U\}$, and $\gamma(q) = \Pr\{\mathrm{Bin}(s, q) \leqslant c_L\}$. Since $\alpha(q)$ increases from 0 to 1 as q increases from 0 to 1 and since $\gamma(q)$ decreases from 1 to 0 as q increases from 0 to 1, the solution of Eq. (1) exists and is unique. Let Γ^* be the solution of Eq. (1), that is, to $\alpha(q) = \gamma(q)$. First assume that Γ^* is not equal to any of d_k and $q_{k-1} < \Gamma^* < q_k$, then

$$\lambda_{k-1} = \frac{\gamma(d_{k-2})}{\alpha(d_{k-1})} > \frac{\gamma(\Gamma^*)}{\alpha(\Gamma^*)} = 1 \quad \text{and} \quad \lambda_{k+1} = \frac{\gamma(d_k)}{\alpha(d_{k+1})} < \frac{\gamma(\Gamma^*)}{\alpha(\Gamma^*)} = 1.$$

Hence the mode is either in d_{k-1} or d_k. The mode is in d_k if $\lambda_k = \gamma(d_{k-1})/\alpha(d_k) > 1$. If $\lambda_k = 1$, the mode spans d_{k-1} and d_k. In the case $q_k = \Gamma^*$, similarly, we can show that $\lambda_k > 1$ and $\lambda_{k+1} < 1$, hence the mode is at d_k.

Appendix B. Proof of the fact that $c_L < \Gamma^* s < c_U$, where Γ^* is the solution of Eq. (1)

If $c_U - c_L = 1$, Γ^* and c_L are such that $\Pr\{\mathrm{Bin}(s, \Gamma^*) \leqslant c_L\} = 0.5$ and the result immediately follows from Uhlmann (1966). If $c_U - c_L > 1$, from (1) $\Pr\{\mathrm{Bin}(s, \Gamma^*) \leqslant c_L\} < 0.5$. Let k be the minimum integer such that $P\{\mathrm{Bin}(s, \Gamma^*) \leqslant k\}$

> 0.5, and hence $c_L \leqslant k-1$. From Hamza (1995) $|s\Gamma^* - k| < \log 2$. Hence $k < s\Gamma^* + \log 2$ and $k-1 < s\Gamma^*$, and consequently $c_L \leqslant k-1 < s\Gamma^*$. The statement regarding c_U can be obtained similarly using the fact that $\Pr\{\mathrm{Bin}(s, \Gamma^*) \geqslant c_U\} = \Pr\{\mathrm{Bin}(s, 1-\Gamma^*) \leqslant s - c_U\} < 0.5$.

References

Barlow, R.E., Bartholomew, D.J., Bremner, J.M., Brunk, H.D., 1972. Statistical Inference under Order Restrictions. Wiley, London, New York.

Bortot, P., Giovagnoli, A., 2005. Up-and-down experiments of first and second order. J. Statist. Plann. Inference 134, 236–253.

Buoen, C., Bjerrum, O.J., Thomsen, M.S., 2005. How first-time-in-human studies are being performed: a survey of phase I dose-escalation trials in healthy volunteers published between 1995 and 2004. J. Clinical Pharmacology 45, 1123–1136.

Cheung, Y.K., Chappell, R., 2000. Sequential designs for phase I clinical trials with late-onset toxicities. Biometrics 56, 1177–1182.

Chevret, S. (Ed.), 2006. Statistical Methods for Dose Finding. Wiley, New York.

Conaway, M., Dunbar, S., Peddada, S.D., 2004. Designs for single or multiple agent phase I trials. Biometrics 60, 661–669.

Crowley, J. (Ed.), 2001. Handbook of Statistics in Clinical Oncology. Marcel Dekker, New York, Basel.

Dragalin, V., Fedorov, V., 2006. Adaptive designs for dose-finding based on efficacy—toxicity response. J. Statist. Plann. Inference 136, 1800–1823.

Durham, S.D., Flournoy, N., 1994. Random walks for quantile estimation. In: Berger, J., Gupta, S. (Eds.), Statistical Decision Theory and Related Topics, vol. 1. Springer, New York, pp. 467–476.

Durham, S.D., Flournoy, N., 1995. Up-and-down designs I. Stationary treatment distributions. In: Flournoy, N., Rosenberger, W.F. (Eds.), Adaptive Designs. Institute of Mathematical Statistics, Hayward, California, pp. 139–157.

Gezmu, M., Flournoy, N., 2006. Group up-and-down designs for dose finding. J. Statist. Plann. Inference 136, 1749–1764.

Giovagnoli, A., Pintacuda, N., 1998. Properties of frequency distributions induced by general 'up-and-down' methods for estimating quantiles. J. Statist. Plann. Inference 74, 51–63.

Hamza, K., 1995. The smallest uniform upper bound on the distance between the mean and the median of the binomial and Poisson distributions. Statist. Probab. Lett. 23, 21–25.

Ivanova, A., 2006. Escalation, up-and-down and A + B designs for dose-finding trials. Statist. Medicine 25, 3668–3678.

Ivanova, A., Montazer-Haghighi, A., Mohanty, S.G., Durham, S.D., 2003. Improved up-and-down designs for phase 1 trials. Statist. Medicine 22, 69–82.

Korn, E.L., Midthune, D., Chen, T.T., Rubinstein, L.V., Christian, M.C., Simon, R.M., 1994. A comparison of two phase I trial designs. Statist. Medicine 13, 1799–1806.

Lin, Y., Shih, W.J., 2001. Statistical properties of the traditional algorithm-based designs for phase I cancer clinical trials. Biostatistics 2, 203–215.

Mira, A., 2001. Ordering and improving the performance of the Monte Carlo Markov Chains. Statist. Sci. 16, 340–350.

Muler, J.H., McGinn, C.J., Normolle, D., Lawrence, T., Brown, D., Hejna, G., Zalupski, M.M., 2004. Phase I trial using a time-to-event continual reassessment strategy for dose escalation of cisplatin combined with gemcitabine and radiation therapy in pancreatic cancer. J. Clinical Oncology 22, 238–243.

O'Quigley, J., 2006. Theoretical study of the continual reassessment method. J. Statist. Plann. Inference 136, 1765–1780.

O'Quigley, J., Pepe, M., Fisher, L., 1990. Continual reassessment method: a practical design for phase I clinical trials in cancer. Biometrics 46, 33–48.

Reiner, E., Paoletti, X., O'Quigley, J., 1999. Operating characteristics of the standard phase I clinical trial design. Comput. Statist. Data Anal. 30, 303–315.

Rosenberger, W.F., Haines, L.M., 2002. Competing designs for phase I clinical trials: a review. Statist. Medicine 21, 2757–2770.

Stallard, N., 2006. Optimal adaptive designs for acute oral toxicity assessment. J. Statist. Plann. Inference 136, 1781–1799.

Stylianou, M., Flournoy, N., 2002. Dose-finding using the biased coin up-and-down design and isotonic regression. Biometrics 58, 171–177.

Thall, P.F., Lee, S.J., 2003. Practical model-based dose-finding in phase 1 clinical trials: methods based on toxicity. Internat. J. Gynecological Cancer 13, 251–261.

Ting, N. (Ed.), 2006. Dose Finding in Drug Development. Springer, New York.

Uhlmann, W., 1966. Verleich der hypergeometrischen mit der Binomial-Verteilung. Metrika 10, 145–158.

Wetherill, G.B., 1963. Sequential estimation of quantal response curves. J. Roy. Statist. Soc. B 25, 1–48.

Yuan, Z., Chappell, R., 2004. Isotonic designs for phase I clinical trials with multiple risk groups. Clinical Trials 1, 499–508.