Contemporary Clinical Trials 36 (2013) 218–227

Contents lists available at SciVerse ScienceDirect

Contemporary Clinical Trials

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

Up-and-down designs for phase I clinical trials

Suyu Liu [a,*], Chunyan Cai [b], Jing Ning [a]

[a] Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, United States
[b] Biostatistics/Epidemiology/Research Design Core, Center for Clinical and Translational Sciences, The University of Texas Health Science Center at Houston, United States

ARTICLE INFO

Article history:
Received 8 May 2013
Received in revised form 2 July 2013
Accepted 3 July 2013
Available online 13 July 2013

Keywords:
Optimal design
Upper bound
Dose finding
Maximum tolerated dose

ABSTRACT

Various up-and-down designs have been proposed to improve the operating characteristics of the traditional "3 + 3" design, but they have been of limited use in practice. A major impediment to the adoption of the improved up-and-down designs is a lack of general guidance and a comprehensive assessment of the operating characteristics of these designs under practical clinical settings. To fill this gap, we review six up-and-down designs: the "3 + 3" design, accelerated titration design, biased coin design, *k*-in-a-row design, group up-and-down design and cumulative group up-and-down design. We conduct comprehensive simulation studies to evaluate their operating characteristics under various practical settings, and compare their performance to a theoretical optimal bound of nonparametric designs. The results show that the cumulative group up-and-down design has the best overall performance in terms of selecting the maximum tolerated dose (MTD), assigning patients to the MTD and patient safety. Its performance is generally close to the upper bound of nonparametric designs, but improvement seems possible in some cases.

Published by Elsevier Inc.

1. Introduction

The primary objective of a phase I clinical trial is to identify the maximum tolerated dose (MTD) of a new drug, which is defined as the dose with a dose-limiting toxicity (DLT) probability that is closest to the target toxicity rate. A phase I clinical trial is important because it determines the MTD that will be further investigated in the ensuing phase II or III trials. Misidentification of the MTD could result in misleading results and serious consequences in the subsequent larger-scale trials.

A class of phase I trial designs widely used in practice are the so-called algorithm-based designs, which are also known as up-and-down designs. The major advantage of this class of designs is their simplicity of implementation. Algorithm-based designs do not require any parametric assumptions on the dose–toxicity curve and strictly conduct dose escalation and deescalation according to prespecified algorithms. Because clinicians know under what precise circumstances dose escalation and deescalation will occur a priori, before the onset of the trial, it is often easy for them to understand and evaluate up-and-down designs based on their clinical experience.

The most well-known up-and-down design is the "3 + 3" design [1]. Although dominant in practice, the "3 + 3" design has been widely criticized for its poor operating characteristics [1–3]. Examples of this include a tendency for the resulting estimators of the MTD to be biased or inconsistent and for a large percentage of patients to be treated at doses below the MTD, and severe restrictions as to the choice of the targeted toxicity probability. Other up-and-down designs have been proposed to achieve better operating characteristics. Simon et al. [4] proposed accelerated titration designs to reduce the number of patients treated at subtherapeutic dose levels in the "3 + 3" design. Durham and Flournoy [5] proposed the biased coin design (BCD), in which the decision of dose escalation and deescalation is based on the toxicity outcome from the most recently treated patient. Wetherill [6] and Gezmu [7] investigated a *k*-in-a-row design to use the toxicity outcomes from the *k* most recently treated patients to determine dose escalation and deescalation. Lin and Shih [8] studied statistical properties of general "A + B" designs.

* Corresponding author. Tel.: +1 7135634280.
E-mail address: syliu@mdanderson.org (S. Liu).

1551-7144/$ – see front matter. Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.cct.2013.07.002

Leung and Wang [9] proposed an up-and-down design based on isotonic regression without making any parametric assumptions on the dose–toxicity curve. Stylianou and Flournoy [10] studied dose finding using the BCD and isotonic regression. Ivanova et al. [11] discussed several up-and-down designs that used more information from the most recently treated patients to make the dose assignment. Stylianou and Follmann [12] extended the BCD to handle delayed toxicities. Gezmu and Flournoy [13] developed the group up-and-down design in which the patients are treated in cohorts. Ivanova, Flournoy and Chung [14] proposed a cumulative cohort design to utilize the cumulative information at the current treating dose to make dose assignment. Ivanova and Kim [15] proposed a more general up-and-down design for dose finding with continuous and ordinal outcomes based on a t-statistic. Comprehensive reviews of dose-finding methods for phase I clinical trials have been provided by Chevret [16] and Ting [17].

Although most of the aforementioned up-and-down designs provide better operating characteristics than the "3 + 3" design, they have had limited use in practice. Rogatko et al. [18] reviewed 1235 phase I cancer trials published between 1991 and 2006, and found that an overwhelming 98.4% of the clinical trials used the "3 + 3" design. A major impediment to the adoption of the improved up-and-down designs in practice is a lack of general guidance and a comprehensive assessment of the operating characteristics of these designs under practical clinical settings. Ivanova [19] provided an excellent review of up-and-down designs, but mainly from the methodological perspective.

To fill this gap, herein we review six up-and-down designs and conduct comprehensive simulation studies to evaluate their operating characteristics under various practical settings. In particular, we compare the performance of these designs to a theoretical optimal bound of nonparametric designs, which elucidates not only the relative but also the absolute performance of the designs. In addition, this comparison sheds light on the potential for improving the existing up-and-down designs.

The remainder of this article is organized as follows. In Section 2, we review six up-and-down designs and the nonparametric optimal design. In Section 3, we present comprehensive simulation studies to assess the operating characteristics of these designs under a wide range of practical settings. We conclude with a brief discussion in Section 4.

2. Methods

Let $(d_1,\ldots,d_J)$ denote a set of J prespecified doses for the drug under investigation with corresponding toxicity probabilities $(p_1,\cdots,p_J)$. Let ϕ denote the target toxicity rate specified by physicians, and Y_i denote the binary toxicity outcome for the ith subject with $Y_i = 1$ indicating that the patient experiences a DLT. We assume that a total of $n = m \times s$ patients will be treated in the trial with a cohort size of s, $s \in \{1,2,\ldots,n\}$, and the first cohort of patients receives the lowest dose d_1. When $s = 1$, patients are assigned one at a time to a dose.

2.1. The "3 + 3" design

In the "3 + 3" design, patients are treated in cohorts of size $s = 3$. The dose escalation rule for the "3 + 3" design can be described as follows: among three patients treated at the current dose, if none experiences DLT, then the next cohort of three patients is treated at the next higher dose; if two or more patients experience DLT, then the next cohort of patients is treated at the next lower dose unless six patients have already been treated at that dose; and if one out of the three patients experiences DLT, then three more patients are treated at that same dose level. In general, if fewer than one of six patients treated at a dose level experiences DLT, then the next cohort is treated at the next higher dose; and if two or more of the six patients treated at a dose level experience DLT, then the MTD is considered to have been exceeded. The MTD is defined as the highest dose at which fewer than two out of six patients experience DLT. Although not explicitly defined in the design, the target toxicity rate of the "3 + 3" design is approximately $\phi = 0.25$.

2.2. The accelerated titration design

The accelerated titration design (ATD) [4] is an extension of the "3 + 3" design that aims to speed up the trial and reduce the number of patients assigned to low doses by adding an accelerated dose-assignment phase. The ATD is conducted in two phases: the trial starts with the titration phase, in which one patient is treated per dose level until one patient exhibits DLT, and then the trial switches to the second phase, which is the traditional "3 + 3" design phase described previously. Similar to the "3 + 3" design, one limitation of the ATD is that it targets only a toxicity rate of $\phi = 0.25$.

2.3. The biased coin design

Unlike the "3 + 3" design and the ATD, the biased coin design (BCD) is more flexible and can target any prespecified toxicity rate ϕ [5]. The BCD is based on the theory of random walk and assigns patients to a dose level one at a time, i.e., $s = 1$. Suppose that the ith patient is treated at dose level j. To determine a dose for the next patient,

- if $Y_i = 1$, we deescalate the dose level to $j - 1$;
- if $Y_i = 0$, we escalate the dose level to $j + 1$ with a probability of $\phi/(1 - \phi)$, otherwise we retain the current dose level j.

One drawback of the BCD is its low efficiency as it uses only the outcome of the most recently treated patient to determine the dose assignment, discarding the treatment information for all the other previously treated patients.

2.4. The k-in-a-row design

The k-in-a-row (KIR) design was proposed to address the low efficiency of the BCD by utilizing the treatment information from the last k ($k > 1$) patients treated, rather than only the last one patient treated, to make the decision of dose assignment [6,7]. In the KIR design, patients are also assigned to a dose level one at a time with $s = 1$. Suppose that the ith patient is treated at dose level j. To determine a dose for the next patient,

- if $Y_i = 1$, we deescalate the dose level to $j - 1$;
- if $Y_i = Y_{i-1} = \ldots = Y_{i-k+1} = 0$, that is, the k most recently treated patients were all treated at dose level j and none of them experienced toxicity, we escalate the dose level to $j + 1$; otherwise we retain the current dose level j.