Contemporary Clinical Trials 31 (2010) 473–482

Contents lists available at ScienceDirect

Contemporary Clinical Trials

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

A novel toxicity scoring system treating toxicity response as a quasi-continuous variable in Phase I clinical trials

Zhengjia Chen [a,b,*], Mark D. Krailo [b,c], Stanley P. Azen [c], Mourad Tighiouart [a]

[a] Rollins School of Public Health, Emory University, Atlanta, GA 30322, United States
[b] Children's Oncology Group, Arcadia, CA 91066, United States
[c] Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, United States

ARTICLE INFO

Article history:
Received 20 January 2010
Accepted 26 May 2010

Keywords:
Toxicity scoring system
Isotonic regression
Maximum tolerance dose
Multiple toxicities
Quasi-continuous variable
Normalized equivalent toxicity score

ABSTRACT

In almost all current Phase I designs, toxicity response is treated coarsely as a binary indicator of dose limiting toxicity (DLT) and a lot of useful toxicity information is discarded. We are the first to establish a novel toxicity scoring system to treat toxicity response as a quasi-continuous variable and utilize all toxicities in Phase I trial. The generally accepted and objective parts, such as a logistic function, grade and type of toxicity, and whether the toxicity is DLT, are used so that the toxicity scoring system is relatively objective. Our toxicity scoring system has been successfully applied to an isotonic design (ID) [1] to develop an extended isotonic design (EID). Simulation study and application of EID to the data of a real Phase I trial demonstrate that EID can always estimate a more accurate maximum tolerated dose (MTD) according to the exact toxicity profile under any toxicity profiles without additional cost or length of the trial. These cannot be accomplished in designs using a binary indicator of DLT, such as Standard 3+3 design, ID, and continual reassessment method (CRM) [2]. Moreover, our EID is relatively objective, model free, and simple to use. Our toxicity scoring system can also be applied to other designs, such as CRM and escalation with overdose control (EWOC) [3], to improve their efficiency and accuracy in MTD estimation by utilizing all toxicities. Our novel toxicity scoring system and EID may help to begin a new era in which toxicity response is treated as a continuous variable.

1. Introduction

As one of the most important steps in drug development, a Phase I clinical trial is the first clinical trial in human subjects after the laboratory and animal study for a therapeutic agent showing a potential cure effect of disease. The main purpose of a cancer Phase I trial is to estimate its MTD under safe administration and acceptable level of adverse events using toxicity responses of a small number of patients treated at different dose levels [4,5].

In almost all of the current Phase I designs, toxicity response is reduced to be a binary indicator as 1 for DLT and 0 for no DLT. In the National Cancer Institute (NCI) common toxicity criteria [9], the DLT is defined as a group of grade 3 or 4 non-hematologic and grade 4 hematologic toxicities as well as death (grade 5) [4,5]. In practice, patients usually have multiple toxicities and there are some correlations between different toxicities, such as fever and fatigue. Some patients even have multiple DLTs and DLTs are not equally severe, for example a grade 4 non-reversible renal toxicity is much more severe than a grade 3 reversible neutropenia [6–8]. Similarly, among "non-toxicities", grades 0, 1, and 2 toxicities are not equally severe and a further differentiation of them will provide a more reliable basis for safety monitoring and dose allocation [6–8]. When toxicity response is treated as a binary variable, only the "worst" toxicity among all toxicities of each patient is considered and further dichotomized into a binary indicator of DLT so that a lot of useful toxicity information is ignored. A phase I trial is a small study with only a small amount of information so that all toxicity information of all

* Corresponding author. 1365-B Clifton Rd, Room B4110, Winship Cancer Institute, Emory University, Atlanta, GA 30322, United States. Tel.: +1 404 778 2017; fax: +1 404 778 5016.
E-mail address: zchen38@emory.edu (Z. Chen).

1551-7144/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.cct.2010.05.010

patients is very valuable and should be fully utilized in order to maximize its efficiency [6–8].

So far, there are only a few studies trying to propose a Phase I design in which toxicity response can be differentiated beyond binary [6–8]. In 2000, Wang et al. [6] first brought up the idea of differentiating toxicity beyond binary by proposing an extended CRM in which a weight is used to reflect the differentiation in the severity of grade 3 and 4 toxicities during dose allocation. Through simulation study, the extended CRM has been shown to reduce the chance of selecting the higher dose level as MTD by giving more impact to grade 4 toxicities. In 2004, Bekele and Thall [7] applied a total toxicity burden (TTB) to measure qualitatively the severity of multiple toxicities in real trials. The dose allocation procedure (increasing, staying at the same dose, or decreasing) is based on the comparison between the observed TTB and the average TTB value, TTBc, of the same outcome in a hypothetical collection of all possible outcomes. The estimated MTD is the dose with the average TTBc value of the "staying at the same dose". In 2006, Yuan et al. [8] proposed a quantitative method called Quasi-CRM in which a numeric equivalent toxicity score was employed to incorporate the impact of toxicity grade on the dose escalation decision of the standard CRM by using the quasi-Bernoulli likelihood. The Quasi-CRM was shown to be superior over the standard CRM and comparable to the Bekele and Thall method [7] in some simulation studies. Unfortunately, the fact that patients usually have multiple toxicities was not considered in their study [8].

A toxicity scoring system which calculates an equivalent score measuring the composite severity of multiple toxicities can be a good solution for the common cases of multiple toxicities per patient. To our knowledge, no such a comprehensive toxicity scoring system for Phase I trials has been proposed in the literature. In this study, we propose a novel toxicity scoring system to measure quantitatively and comprehensively the overall severity of multiple toxicities per patient. In order to reduce the arbitrariness and stay in the current track of Phase I clinical trial practice, the generally accepted and relatively objective components, such as a logistic function, grade and type of toxicity, and whether the toxicity is DLT, are used to establish our toxicity scoring system. At last we demonstrate that our system can be easily incorporated into the common designs to treat toxicity response as a quasi-continuous variable and increase the accuracy and efficiency of MTD estimation by simulation study and application to the data of a real Phase I clinical trial.

2. A novel toxicity scoring system

In the NCI common toxicity criteria (NCI 2003) [9], according to their severities and types, toxicities are classified into 5 grades as follows:

Grade 0: no toxicity;
Grade 1: mild toxicity;
Grade 2: moderate toxicity;
Grade 3: severe toxicity;
Grade 4: life-threatening toxicity; and
Grade 5: death.

The DLT is usually defined as a group of grade 3 or higher non-hematologic toxicities and grade 4 hematologic non-transient toxicities.

2.1. A mapping between adjusted grade and original toxicity

DLTs are usually pre-defined specifically in each trial. In order to take into account the classification of DLT and imitate what is done conventionally using the NCI grades, a mapping between adjusted grade and original toxicity has been proposed (Table 1). In the mapping, we further differentiate grade 3 toxicities into grade 3 non-DLT and grade 3 DLT, and grade 4 toxicities into grade 4 non-DLT and grade 4 DLT. It is assumed that low grade non-DLT is less severe than high grade non-DLT, non-DLT is less severe than DLT, and grade 3 DLT is less severe than grade 4 DLT. Therefore we assign an adjusted grade for toxicity: 0 for grade 0 toxicity, 1 for grade 1 toxicity, 2 for grade 2 toxicity, 3 for grade 3 non-DLT, 4 for grade 4 non-DLT, 5 for grade 3 DLT, and 6 for grade 4 DLT (Table 1). Drug-related death (grade 5) is not considered because when it happens the trial needs to be suspended and re-evaluated. But if necessary, a new highest adjusted grade, such as 7, can be assigned to death.

2.2. Equivalent toxicity score

Equivalent toxicity score (ETS) is defined as a quantitative measurement of the overall toxicity severity for each patient. The mapping between adjusted grade and original toxicity in Section 2.1 is flexible because the estimated ETS of each patient will be further normalized to a range from 0 to 1 in Section 2.3.

Suppose that there are $n_1, n_2,\ldots, n_K$ patients who received dose levels $d_1, d_2,\ldots, d_K$, $(d_1 \le d_2 \le \ldots \le d_K)$ respectively. Let $T_{j,k,i}$ be the ith $(1 \le i \le I)$ toxicity of the jth $(1 \le j \le n_k)$ patient among the n_k patients who received the dose level d_k $(1 \le k \le K)$. Its NCI toxicity grade is $G_{j,k,i}$ and the corresponding adjusted grade after the mapping is $G'_{j,k,i}$.

Let the maximum adjusted grade, $G'_{j,k,\max}$, among all I toxicities of the patient j at the dose level d_k be defined as:

$$G'_{j,k,\max} = \max_{1 \le i \le I}(G'_{j,k,i}).$$

Let the ETS for patient j at the dose level d_k be defined as $S_{j,k}$. In order to make the range of ETS start from 0, the ETS for a patient with no toxicity is defined as 0 as below:

$$S_{j,k} = 0.$$

Table 1
Mapping of adjusted grade and original toxicity. [a]

Original grade/whether DLT	Grade 0	Grade 1	Grade 2	Grade 3 non-DLT	Grade 4 non-DLT	Grade 3 DLT	Grade 4 DLT
Adjusted grade	0	1	2	3	4	5	6

[a] DLT: dose limiting toxicity.