Contemporary Clinical Trials 71 (2018) 162–172

Contents lists available at ScienceDirect

Contemporary Clinical Trials

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

Efficacy/toxicity dose-finding using hierarchical modeling for multiple populations

Kristen M. Cunanan*, Joseph S. Koopmeiners

Memorial Sloan Kettering Cancer Center, Department of Epidemiology and Biostatistics, 485 Lexington Avenue 2nd Floor, New York, NY 10017, United States

ARTICLE INFO

Keywords:
Phase I-II
Dose-finding
Multiple populations
Continual reassessment method

ABSTRACT

Traditionally, Phase I oncology trials evaluate the safety profile of a novel agent and identify a maximum tolerable dose based on toxicity alone. With the development of biologically targeted agents, investigators believe the efficacy of a novel agent may plateau or diminish before reaching the maximum tolerable dose while toxicity continues to increase. This motivates dose-finding based on the simultaneous evaluation of toxicity and efficacy. Previously, we investigated hierarchical modeling in the context of Phase I dose-escalation studies for multiple populations and found borrowing strength across populations improved operating characteristics. In this article, we discuss three hierarchical extensions to commonly used probability models for efficacy and toxicity in Phase I-II trials and adapt our previously proposed dose-finding algorithm for multiple populations to this setting. First, we consider both parametric and non-parametric bivariate models for binary outcomes and, in addition, we consider an under-parameterized model that combines toxicity and efficacy into a single trinary outcome. Our simulation results indicate hierarchical modeling increases the probability of correctly identifying the optimal dose and increases the average number of patients treated at the optimal dose, with the under-parameterized hierarchical model displaying desirable and robust operating characteristics.

1. Introduction

Phase I oncology trials are primarily dose-escalation studies to evaluate the safety of a novel treatment and identify the maximum tolerable dose (MTD), defined as the highest dose with probability of dose limiting toxicity (DLT) less than some pre-specified threshold. Typically, efficacy is not examined until Phase II. Historically, clinicians believed the probabilities of toxicity and efficacy increase monotonically with dose and, subsequently, the highest dose with acceptable toxicity was thought to have the best chance to succeed in future trials. However, for contemporary biologically targeted agents, investigators often believe a drug's potential efficacy may level off or diminish before reaching the MTD, while potential toxicity increases with dosage. This motivates dose-finding based on the simultaneous evaluation of toxicity and efficacy. Furthermore, given the limited sample sizes in Phase I oncology trials, incorporating efficacy into dose-finding may improve identifying the optimal dose used in subsequent trials. Gooley et al. [8] were among the first to propose a dose-finding design based on simultaneous evaluation of toxicity and efficacy. Their results suggest that the additional dose-efficacy curve adds complexity (i.e., model parameters) to the dose-finding algorithm which is a cost that should be considered when designing a Phase I-II trial. Consequently, Thall and Russell [20] proposed a design combining toxicity and efficacy into one variable, reducing the parameter space. Alternatively, Braun [3] extends the continual reassessment method to account for two competing outcomes, while Thall and Cook [17] take a similar approach but also define a trade-off contour to guide dose-finding. A number of extensions to this basic approach have been discussed over the last decade [9, 12–15, 18, 19, 21–23]. Researchers are often interested in evaluating a novel treatment in a number of patient populations, which may have different background standards-of-care. For example, researchers at University of Minnesota College of Veterinary Medicine Animal Cancer Care and Research Program are interested in completing a Phase I-II trial of a novel targeted toxin. The trial will enroll dogs in two cohorts: a cohort focused on hemangiosarcomoa, for which the drug has previously shown promising results [2], and a cohort for other solid tumors. In this case, the hemangiosarcoma cohort will not utilize information found in the solid tumor cohort, resulting in a potential loss of efficiency, while the solid tumor cohort will collapse across multiple tumor types with potentially heterogeneous dose-response relationships. An alternate approach would be to use hierarchical modeling (HM) to allow each population to have separate dose-response relationships, while borrowing strength across populations to gain efficiency. Previously, we investigated HM in the context of Phase I dose-

* Corresponding author.
E-mail address: kristenmay206@gmail.com (K.M. Cunanan).

https://doi.org/10.1016/j.cct.2018.06.012
Received 11 December 2017; Received in revised form 5 May 2018; Accepted 13 June 2018
Available online 21 June 2018

escalation studies. We proposed extensions to commonly used dose-toxicity models and proposed dose-finding guidelines that protect patient safety, while allowing the design to fully realize the potential of HM [5]. Our simulation results indicate incorporating HM into Phase I dose-finding increases the probability of correctly identifying the MTD and the average number of patients treated at the MTD, with little impact on the rate of DLTs. In this article, we propose a Bayesian adaptive Phase I-II dose-escalation design that uses HM to estimate population-specific biologically optimal doses (BODs), while sharing both dose-toxicity and -efficacy information across populations.

2. Models

In this section, we present hierarchical extensions of three joint probability models for efficacy and toxicity that have been proposed for use in Phase I-II dose-finding trials. In each case, we define a two-level Bayesian hierarchical model where the first level specifies the population-level parameters and the second level facilitates borrowing across populations. Existing joint probability models for Phase I-II clinical trials can be broadly classified into two groups: bivariate outcome models, where separate dose-response models are specified for efficacy and toxicity and the correlation between efficacy and toxicity is incorporated into the model using a copula model or some other approach [3, 17, 21], and trinomial models, where efficacy and toxicity are combined into a trinomial outcome and a dose-response relationship is specified for the trinomial outcome [20, 23]. We begin by discussing hierarchical extensions of two bivariate binary outcome models and then discuss a hierarchical extension of the trinomial model proposed by Zhang et al. [23].

2.1. Bivariate binary outcomes

We use the following notation throughout Section 2.1. First, let T_{ikj} be a binary indicator for the presence or absence of DLT in subject i treated at dose j in population k, which takes the value 1 with probability $\pi_{T,\,kj}$, and let E_{ikj} be a binary indicator for the probability of tumor response in subject i treated at dose j in population k, which takes the value 1 with probability $\pi_{E,\,kj}$. We will consider two approaches for specifying a bivariate outcome model. First, we consider a parametric approach, where parametric dose-response models are specified for efficacy and toxicity. Next, we consider a non-parametric model that imposes a monotonicity constraint on the dose-toxicity model but avoids a formal parametric model.

2.1.1. Parametric model

For our parametric model, we extend a simple one-parameter power model for toxicity and a more flexible, quadratic logistic regression model for efficacy. Our hierarchical model for toxicity is specified as:

$$pr(T_{ikj} = 1 \mid population = k, dose = j) = \pi_{T,kj} = p_j^{exp(\alpha_k)} \tag{1}$$

$$\alpha_k \mid \mu_\alpha, \sigma_\alpha^2 \sim N(\mu_\alpha, \sigma_\alpha^2)$$

$$\mu_\alpha \sim Normal(0, 2^2) \quad \text{and} \quad \sigma_\alpha \sim Uniform(0.39, 3)$$

for dose level $j = 1, \ldots, D$ and population $k = 1, \ldots, K$. The vector $(p_1, \ldots, p_D)$ is referred to as the skeleton and its components are monotonically increasing and take values between 0 and 1. For our simulation results presented in Section 4, we set the power model skeleton equal to $(0.05, 0.15, 0.25, 0.35, 0.45)$. Our hierarchical model for efficacy is specified as:

$$pr(E_{ikj} = 1 \mid population = k, dose = j) = \pi_{E,kj} = \beta_{0k} + \beta_{1k}(dose - 1) + \beta_{2k}(dose - 1)^2 \tag{2}$$

$$\beta_{lk} \mid \mu_{\beta l}, \sigma_{\beta l}^2 \sim Normal(\mu_{\beta l}, \sigma_{\beta l}^2)$$

$$\mu_{\beta l} \sim Normal(m_l, s_l^2) \quad \text{and} \quad \sigma_{\beta l} \sim Uniform(0.39, 3),$$

for $l = 0, 1, 2$, dose level $j = 1, \ldots, D$ and population $k = 1, \ldots, K$. We originally fixed the intercept equal to -3 to reduce the number of unknown parameters, as suggested by Goodman et al. [7]. This reflects a 5% probability of tumor response at dose level 1, but we found that this model did not provide enough flexibility when the true optimal dose resides in the higher dose levels. The unknown m_0, m_1, and m_2 are the shared mean hyper-parameters for the intercept, linear and quadratic terms and are set equal to -2, 0.1, and 0, respectively, with shared variance hyper-parameters set to ${s_0}^2 = 4$, ${s_1}^2 = 9$, and ${s_2}^2 = 4$. This corresponds to a conservative, monotonic prior efficacy-skeleton of 0.12, 0.13, 0.14, 0.15, 0.17 for dose levels 1, 2, 3, 4, 5, respectively. The ${\sigma_{\beta l}}^2$ are our hierarchical variance parameters that control the amount of borrowing across populations, with smaller values indicating more borrowing. We specify a uniform prior distribution on the standard deviation, rather than the log standard deviation, as in [5], since this prior is well-received for other hierarchical applications and we are interested in exploring its use further in our dose-finding setting. In our previous investigation, a uniform prior on the standard deviation with a lower bound of 0 produced poor convergence and identifiability, given the small sample sizes early in a trial. The lower bound of our uniform prior was set to 0.39, based on our simulation results, which suggested that a lower bound < 0.39 results in over-borrowing and poor trial operating characteristics in settings where the true optimal dose varies by population. The toxicity and efficacy outcomes in Phase I-II clinical trials are thought to be correlated and a number of approaches have been proposed for jointly modeling efficacy and toxicity in Phase I-II clinical trials [3, 17, 21]. Recently, Iasonos et al. [11] provided an extensive evaluation of the effect of dimensionality on trial operating characteristics in early phase dose-finding studies. They found that more parsimonious models typically result in improved operating characteristic, even when some aspects of the data generating process are misspecified. A number of other authors have come to similar conclusions with respect to estimating the correlation between efficacy and toxicity in Phase I-II clinical trials [4, 10, 15, 21]. Therefore, we will proceed assuming independence between the toxicity and efficacy outcome for our parametric model.

2.1.2. Non-parametric model

The second model we consider is a hierarchical extension of the non-parametric model proposed by Yin et al. [21]. They specify a dose-response relationship for toxicity and efficacy through the following transformations. For population $k = 1, \ldots, K$, the dose-response model for toxicity is specified as,

$$\phi_{k1} = logit(\pi_{T,k1}), \qquad \phi_{kj} = log\left(\frac{\pi_{T,kj}}{1 - \pi_{T,kj}} - \frac{\pi_{T,k(j-1)}}{1 - \pi_{T,k(j-1)}}\right)$$

for $j = 2, \ldots, D$, and for efficacy, let

$$\psi_{k1} = logit(\pi_{E,k1}), \qquad \psi_{kj} = log\left(\frac{\pi_{E,kj}}{1 - \pi_{E,kj}}\right) - log\left(\frac{\pi_{E,k(j-1)}}{1 - \pi_{E,k(j-1)}}\right)$$

for $j = 2, \ldots, D$. The primary difference between the two parameterizations is that the model for toxicity enforces a monotonicity constraint on the dose-response relationship for toxicity, whereas the model for efficacy does not. Yin et al. [21] originally specified a bivariate normal prior for the efficacy and toxicity parameters to allow a priori correlation between the model parameters but found that setting the off-diagonal covariance elements to zero did not impact their results. We will specify independent normal priors for ϕ_{kj} and ψ_{kj} and facilitate borrowing strength across populations by specifying a