



Published in final edited form as:

*Stat Methods Med Res.* 2022 June ; 31(6): 1104–1119. doi:10.1177/09622802221080753.

## A Bayesian Phase I/II Biomarker-based Design for Identifying Subgroup-Specific Optimal Dose for Immunotherapy

Beibei Guo<sup>1</sup>, Yong Zang<sup>2,3</sup>

<sup>1</sup>Department of Experimental Statistics, Louisiana State University, Baton Rouge, LA 70803, U.S.A.

<sup>2</sup>Department of Biostatistics and Health Data Science, Indiana University

<sup>3</sup>Center for Computational Biology and Bioinformatics, Indiana University, Indianapolis, IN 46202, U.S.A.

### Abstract

Immunotherapy is an innovative treatment that enlists the patient's immune system to battle tumors. The optimal dose for treating patients with an immunotherapeutic agent may differ according to their biomarker status. In this article, we propose a biomarker-based phase I/II dose-finding design for identifying subgroup-specific optimal dose for immunotherapy (BSOI) that jointly models the immune response, toxicity, and efficacy outcomes. We propose parsimonious yet flexible models to borrow information across different types of outcomes and subgroups. We quantify the desirability of the dose using a utility function and adopt a two-stage dose-finding algorithm to find the optimal dose for each subgroup. Simulation studies show that the BSOI design has desirable operating characteristics in selecting the subgroup-specific optimal doses and allocating patients to those optimal doses, and outperforms conventional designs.

### Keywords

Immunotherapy; subgroups; biomarker; phase I/II trial; dose finding; immune response; risk-benefit tradeoff; Bayesian adaptive design

## 1. Introduction

Cancer immunotherapy is an innovative treatment that stimulates a patient's immune system to attack tumor cells. It represents the most promising new cancer treatment approach since the first chemotherapies were developed in the late 1940s.<sup>1–3</sup> The recent success of immunotherapeutic agents, such as ipilimumab, nivolumab, and pembrolizumab, has made immunotherapy a topical area in cancer research.<sup>4–6</sup>

Numerous phase I/II trial designs have been proposed in the literature that consider both toxicity and efficacy, including Thall and Russell,<sup>7</sup> Thall and Cook,<sup>8</sup> Yin et al.,<sup>9</sup> Jin et al.,<sup>10</sup> Zang and Lee,<sup>11</sup> and Guo and Yuan.<sup>12</sup> Since immunotherapy achieves its therapeutic effect

by activating the immune system, the immune response is closely related to the treatment effect of immunotherapy and so should be incorporated into the trial design. To account for the immune response, a few dose-finding designs have been developed recently,<sup>13–16</sup> which jointly model the immune response, toxicity, and efficacy to identify the optimal dose for immunotherapy. The trivariate continual reassessment method described in Zhong, Koopmeiners, and Carlin<sup>17</sup> serves a similar purpose by considering surrogate markers.

One limitation of the preceding designs is that they all assume patient homogeneity and use the “one-dose-fits-all” rule for patient allocation and optimal dose identification. There is emerging evidence that a variety of factors may affect immunotherapy effectiveness. For example, the programmed cell death-ligand 1 (PD-L1) expression has been demonstrated as a predictive biomarker for checkpoint inhibitor-based immunotherapy in various cancer types, with PD-L1 positive patients showing higher response rates, better progression-free survival and overall survival, compared with PD-L1 negative patients.<sup>18–22</sup> As summarized in Bai et al.,<sup>23</sup> several studies have found significant gender difference in the response to immune checkpoint inhibitors in patients with various cancer types. Kugel et al.<sup>24</sup> found that melanoma patients over 60 years old had a significantly higher tumor response rate to pembrolizumab than patients under 60, and the likelihood of response increased with age. Nishijima et al.<sup>25</sup> also reported different efficacy of immune checkpoint inhibitors between younger and older patients. The CheckMate 171 trial showed that patients with performance status value 2 had inferior efficacy to the overall population.<sup>26</sup> Murphy and Longo<sup>27</sup> noticed a strong association between obesity and cancer immunotherapy efficacy. Wang et al.<sup>28</sup> reported a similar finding between obesity and PD-1 checkpoint blockade. In this paper, we will focus on the biomarker-defined subgroups. The optimal dose for treating patients with an immunotherapeutic agent may differ according to the patient’s biomarker status. For example, in 2016, the U.S. Food and Drug Administration (FDA) approved pembrolizumab for the treatment of metastatic non-small cell lung cancer (NSCLC) for PD-L1 positive patients with recommended dose 200 mg every three weeks. Multiple studies<sup>29–33</sup> have demonstrated equivalent efficacy with weight-based personalized doses. Goldstein et al.<sup>34</sup> demonstrated that weight-based dosing would lead to a significant reduction of drug cost and called for personalized immunotherapy and emphasized the importance of it. Ingles Garces et al.<sup>35</sup> commented that for immunotherapy, “More biomarker data may be useful in selection of dose and regimen in early clinical development, ...” in the conclusion part of the paper. Therefore, optimizing the treatment benefit for immunotherapy requires the dose-finding method to identify the optimal dose for each subgroup, stratified by patient’s biomarker status.

Our research is motivated by a phase I/II trial conducted at Indiana University Simon Comprehensive Cancer Center, which uses an anti-PD-L1 immune checkpoint inhibitor to treat patients with advanced refractory solid tumors. Five doses (0.1, 0.3, 0.5, 0.7, 0.9 mg/kg) of the inhibitor will be investigated, and a maximum of 60 patients will be accrued to the trial. The immune response is measured by the count of CD8+ T-cells on day 14 into the treatment. The dose-limiting toxicity is defined using National Cancer Institute Common Terminology Criteria for Adverse Events, Version 4.0. Patient efficacy response is characterized as complete response (CR), partial response (PR), stable disease (SD), or progressive disease (PD) based on the Response Evaluation Criteria in Solid Tumors. Our

motivating trial's goal is to identify the optimal dose for each of the PD-L1 positive and negative subgroups.

We propose a biomarker-based Bayesian phase I/II trial design (BSOI) to optimize subgroup-specific dose for immunotherapy. The BSOI jointly models the continuous immune response, binary toxicity outcome, and ordinal efficacy outcome as functions of dose and biomarker status. To capture the distinct features and relationships among the three endpoints, we model the marginal distribution of the immune response using a plateau dose-response model; and conditional on the immune response, we model the distributions of the binary toxicity outcome and the ordinal efficacy outcome. Following our motivating trial, in this article, we focus on the situation where the two biomarker subgroups are ordered in terms of efficacy while discussing how to handle unordered subgroups. To accommodate the small sample size in typical early-phase trials, we develop a parsimonious yet reasonably flexible model to borrow information across different types of outcomes and subgroups. We elicit the numerical utility to quantify the desirability of the dose based on the risk-benefit tradeoff. A two-stage dose-finding algorithm is proposed to guide dose assignment and selection.

The remainder of this article is organized as follows. In Section 2, we describe the probability models and present the dose-finding algorithm. In Section 3, we investigate the operating characteristics of the BSOI design through simulation studies. We provide concluding remarks in Section 4.

## 2. Method

### 2.1. Probability Models

Consider a phase I/II trial with  $D$  prespecified doses,  $d_1 < \dots < d_D$ , of an immunotherapeutic agent. Let  $Y_I$ ,  $Y_T$ , and  $Y_E$  denote the immune response, toxicity outcome, and efficacy outcome, respectively. We consider a binary biomarker with a subgroup indicator  $Z = 0$  or  $1$  corresponding to the marker-negative and marker-positive subgroup. The objective of the trial is to identify the optimal dose for each marker subgroup. To reflect the fact that in immunotherapy, clinical outcomes rely on the activation of the immune system, we factorize the joint distribution of  $(Y_I, Y_T, Y_E)$  into the product of the marginal distribution of  $Y_I$  and the distributions of  $Y_T$  and  $Y_E$  conditional on  $Y_I$ .

The immune response  $Y_I$  is taken to be the increase in a log-transformed immune activity (e.g., T-cell resistance) from baseline to post-treatment, which is generally a continuous outcome. We assume the observed immune response  $Y_{I,i}$  for patient  $i$  with subgroup indicator  $Z_i$  treated at dose  $d_i$  follows a normal distribution with mean  $\mu_I(d_i, Z_i)$  and variance  $\sigma^2$ , that is,

$$Y_{I,i} \mid d_i, Z_i \sim N(\mu_I(d_i, Z_i), \sigma^2) \quad (1)$$

where the expected value  $\mu_I(d_i, Z_i)$  is modeled as

$$\log \left\{ \frac{\mu_I(d_i, Z_i)}{\alpha \exp(\delta Z_i) - \mu_I(d_i, Z_i)} \right\} = \eta_0 + \eta_1 d_i, \quad (2)$$

$\alpha$  and  $\alpha \exp(\delta)$  are the maximum mean immune responses that the immunotherapeutic agent possibly achieves for marker-negative and positive subgroups, respectively. The specific value  $\delta = 0$  corresponds to the absence of patient heterogeneity. We restrict  $\alpha > 0$  to reflect that the mean immune response is positive as immunotherapy typically increases immune activity. Since in most clinical practices, marker-positive patients have a higher maximum mean immune response than marker-negative patients, we assign a prior distribution to  $\delta$  that puts most probability mass to positive values while allowing a small, say 5%, probability on negative values to accommodate situations where this ordering assumption is violated. If the subgroups' ordering is unknown, no restriction would be placed on the sign of  $\delta$ . We constrain  $\eta_1 > 0$  such that the immune response increases or first increases and then plateaus when the dose increases. Since  $\mu_I(d_i, Z_i)$  represents the mean pre-post difference in the immune activity, it is expected to be 0 when there is no drug, i.e.,  $d_i = 0$ . So we set  $\eta_0 = -3$  so that  $\mu_I(d_i, Z_i)$  is virtually 0 when  $d_i = 0$ . Compared with the Emax model used in Liu, Guo, and Yuan<sup>13</sup>, the mean immune response model (2) is more parsimonious yet very flexible. Figure 1 shows the immune response curves generated from this model with different values of  $\alpha$ ,  $\delta$ , and  $\eta_1$ . In the left panel, the mean immune response increases for both subgroups; in the middle panel, the mean immune response first increases and then plateaus for both subgroups; in the right panel, the mean immune response increases slowly and then plateau for marker-negative patients, and increases for marker-positive patients.

Taking the setting of the motivating trial, we model toxicity  $Y_T$  as a binary outcome, with  $Y_T = 1$  indicating toxicity (or severe adverse events), and  $Y_T = 0$  otherwise. For patient  $i$  treated at dose  $d_i$  with subgroup indicator  $Z_i$  and immune response  $Y_{I,i}$ , we model  $Y_{T,i}$  using the logistic model

$$\text{logit}\{\Pr(Y_{T,i} = 1 \mid d_i, Z_i, Y_{I,i})\} = \beta_{0,Z_i} + \beta_1 d_i + \beta_2 Y_{I,i} \quad (3)$$

In equation (3),  $\beta_{0,0}$  and  $\beta_{0,1}$  are the intercepts for marker-negative and marker-positive patients, respectively,  $\beta_1$  is the dose effect, and  $\beta_2$  is the effect of the immune response. To reflect that an immunotherapeutic agent's toxicity is closely related to the immune response but not necessarily entirely caused by the immune response, we include both dose  $d_i$  and the immune response  $Y_{I,i}$  in the model. That is, the immune response cannot fully explain the toxicity effect at different doses. If the dose-effect on the drug's toxicity is presumably entirely mediated by the immune response in certain trials, we may drop  $d_i$  from the model. In some trials, the toxicity outcome may be coded as a  $K$ -level ordinal variable, for example,  $Y_T = 1, 2, \text{ and } 3$  (so  $K = 3$ ) denoting toxicity grades 1–2, 3, and 4, respectively. In those cases, we can use the proportional odds model as for the ordinal efficacy outcome described in the next paragraph.

The efficacy outcome  $Y_E$  is modeled as an ordinal outcome with  $Y_E = 1, 2, \dots, J$  corresponding to increasingly desirable efficacy. In immunotherapy, the tumor response is often classified as CR, PR, SD, or PD, so we define  $Y_E$  as an ordinal outcome with  $Y_E =$

1, 2, and 3 indicating PD, SD, and CR/PR, respectively. This is based on the consideration that although CR and PR are generally more desirable, SD is often considered a positive immunotherapy response. Some immunotherapies prolong survival by achieving SD without significant tumor shrinkage. For patient  $i$  treated at dose  $d_i$  with subgroup indicator  $Z_i$  and immune response  $Y_{I,i}$ , we model  $Y_{E,i}$  using the proportional odds model as follows,

$$\text{logit}\{\Pr(Y_{E,i} \leq j \mid d_i, Z_i, Y_{I,i})\} = \gamma_{0,j} + \gamma_1 Z_i + \gamma_2 Y_{I,i} + \gamma_3 Y_{I,i}^2 \quad j = 1, \dots, J \quad (4)$$

where  $\gamma_{0,1} < \dots < \gamma_{0,J-1}$  are intercepts, and the quadratic term  $Y_{I,i}^2$  is used to accommodate the possibility that efficacy may not monotonically increase with the immune response. In equation (4), we assume that in each subgroup, conditional on the immune response  $Y_I$ ,  $Y_E$  is independent of dose  $d$  to reflect the consideration that the working mechanism of immunotherapy is to eliminate cancer cells by activating patient's immune system and therefore the treatment effect of immunotherapy is mostly mediated by the immune response. For cases where such an assumption is violated, we can add the dose  $d$  as a covariate in the model. Since the correlation between toxicity and efficacy is often weak for immunotherapy, we assume that conditional on the immune response  $Y_I$ ,  $Y_T$  and  $Y_E$  are independent. Previous studies<sup>14,36</sup> show that ignoring the correlation between toxicity and efficacy has little impact on the performance of phase I/II designs. If needed, we can add a patient-specific random effect in both equations (3) and (4) to model the correlation between  $Y_T$  and  $Y_E$ .

For the  $i$ th patient, denote the observed outcome by  $\mathbf{Y}_i = (Y_{I,i}, Y_{T,i}, Y_{E,i})$ , the assigned dose by  $d_i$ , and the subgroup indicator by  $Z_i$ . Letting  $\Theta = (\alpha, \delta, \sigma^2, \eta_1, \beta_{0,0}, \beta_{0,1}, \beta_1, \beta_2, \gamma_{0,1}, \dots, \gamma_{0,J-1}, \gamma_1, \gamma_2, \gamma_3)$  represent all parameters, the likelihood for the  $i$ th patient is

$$L_i(\mathbf{Y}_i; \Theta) = f(Y_{I,i} \mid \Theta) f(Y_{T,i} \mid Y_{I,i}, \Theta) f(Y_{E,i} \mid Y_{I,i}, \Theta) \quad (5)$$

where

$$f(Y_{I,i} \mid \Theta) = \phi\left(Y_{I,i}; \frac{\alpha \exp(\delta Z_i) \exp(\eta_0 + \eta_1 d_i)}{1 + \exp(\eta_0 + \eta_1 d_i)}, \sigma^2\right), \quad (6)$$

$$f(Y_{T,i} \mid Y_{I,i}, \Theta) = \frac{\exp\{Y_{T,i}(\beta_0 + \beta_1 d_i + \beta_2 Y_{I,i})\}}{1 + \exp(\beta_0 + \beta_1 d_i + \beta_2 Y_{I,i})} \quad (7)$$

$$f(Y_{E,i} \mid Y_{I,i}, \Theta) = \frac{\exp(\gamma_{0,Y_{E,i}} + \gamma_1 Z_i + \gamma_2 Y_{I,i} + \gamma_3 Y_{I,i}^2)}{1 + \exp(\gamma_{0,Y_{E,i}} + \gamma_1 Z_i + \gamma_2 Y_{I,i} + \gamma_3 Y_{I,i}^2)} \quad (8)$$

$$- \frac{\exp(\gamma_{0,Y_{E,i}-1} + \gamma_1 Z_i + \gamma_2 Y_{I,i} + \gamma_3 Y_{I,i}^2)}{1 + \exp(\gamma_{0,Y_{E,i}-1} + \gamma_1 Z_i + \gamma_2 Y_{I,i} + \gamma_3 Y_{I,i}^2)},$$

and  $\phi(\cdot; \mu, \sigma^2)$  denotes the probability density function for a normal distribution with mean  $\mu$  and variance  $\sigma^2$ . Let  $n = 1, \dots, N$  denote an interim sample size when a dose

assignment decision is to be made during the trial, and  $\mathcal{D}_n = (Y_1, \dots, Y_n)$  denote the observed data from the first  $n$  patients. The likelihood for the first  $n$  patients in the trial is  $L(\mathcal{D}_n; \Theta) = \prod_{i=1}^n L_i(Y_i; \Theta)$ .

Let  $p(\Theta)$  denote the joint prior distribution of  $\Theta$ . The joint posterior distribution based on the data from the first  $n$  patients is  $p(\Theta | \mathcal{D}_n) \propto L(\mathcal{D}_n; \Theta)p(\Theta)$ . We sample from this posterior distribution using the Markov chain Monte Carlo algorithm with Gibbs sampler.<sup>37</sup>

## 2.2. Prior Distributions

To specify the prior distribution for  $\eta_1$  in equation (2), we follow the idea of Gelman et al.<sup>38</sup>. Note that equation (2) is equivalent to  $\log\left\{\frac{\mu_I(d_i, Z_i)/(a\exp(\delta Z_i))}{1 - \mu_I(d_i, Z_i)/(a\exp(\delta Z_i))}\right\} = \eta_0 + \eta_1 d_i$ , which has the standard logistic regression form given that the range of  $\mu_I(d_i, Z_i)$  is  $(0, a\exp(\delta Z_i))$ . A change of 5 moves  $\mu_I(d_i, Z_i)/(a\exp(\delta Z_i))$  from 0.01 to 0.5 or from 0.5 to 0.99, which is considered unlikely for a typical change in a covariate. Therefore, we scale  $d_i$  to have standard deviation 0.5, and assign  $\eta_1$  a truncated normal prior distribution  $\eta_1 \sim \mathcal{N}(0, 2.5^2)I(\eta_1 > 0)$  so that a change in  $d_i$  from one standard deviation below the mean to one standard deviation above the mean will most likely result in a difference of less than 5 for the regression in equation (2).

Since  $a$  has the interpretation of the maximum immune response for marker-negative patients, and  $\exp(\delta)$  has the interpretation of the ratio of the maximum immune response of the marker-positive patients relative to the marker-negative patients, we elicit prior estimates of  $a$  and  $\exp(\delta)$  from clinicians, denoted as  $\hat{a}$  and  $\hat{r}$ , respectively.  $a$  is assigned a Gamma prior distributions with mean  $\hat{a}$  and a relatively large standard deviation (e.g.,  $3\hat{a}$ ) to obtain a vague prior, and  $\delta$  is assigned a Normal prior distribution with mean  $\log(\hat{r})$  and a variance such that there is a small, say 5%, probability on negative values of  $\delta$  to accommodate situations that the observed data are in contradiction with the ordering assumption. We assign  $\sigma^2$  a vague inverse Gamma prior distribution, e.g.,  $\sigma^2 \sim IG(0.1, 0.1)$  so that data will dominate the posterior distribution.

In toxicity model (3), we assume  $\beta_{0,0}$  and  $\beta_{0,1}$  independently follow Normal prior distributions with mean  $-4$  and standard deviation 1, i.e.,  $\mathcal{N}(-4, 1^2)$  so that a priori the toxicity probability in each subgroup is centered at 0.018 with 95% credible interval (0.0025, 0.12) when there is no drug, i.e., when  $d_j = 0$  and  $Y_{L,j} = 0$ . For  $\beta_1$  and  $\beta_2$ , we follow the same argument as for the immune response model (2) and assign normal prior distributions  $\mathcal{N}(0, 2.5^2)$  to  $\beta_1$  and  $\beta_2$  after scaling  $d_j$  and  $Y_{L,j}$  so that each has standard deviation 0.5. Here we do not restrict  $\beta_1$  and  $\beta_2$  to be positive to accommodate the fact that toxicity with immunotherapy agents typically increases slowly with the dose, so the true values of  $\beta_1$  and  $\beta_2$  are close to 0. Constraining the values to be positive using, e.g., the Gamma or truncated normal priors, tends to inflate the estimates of  $\beta_1$  and  $\beta_2$  especially at the beginning of the trial when data are sparse, which hinders dose escalation and thus hurts the performance of the design.

Likewise, in the efficacy model (4), we scale  $Y_{Li}$  to have mean 0 and standard deviation 0.5 and shift the binary  $Z_i$  to have a mean of 0 and to differ by 1 in their lower and upper conditions,<sup>38</sup> and assign  $\gamma_2$  and  $\gamma_3$  normal prior distributions  $\mathcal{N}(0, 2.5^2)$ . For the intercepts  $\gamma_{0,j}$  we use the same normal priors under the constraint that  $\gamma_{0,1} < \dots < \gamma_{0,j}$ . To accommodate the fact that the marker-positive subgroup tends to have a higher probability of efficacy,  $\gamma_1$  is assigned a normal prior distribution  $\mathcal{N}(-2.5, 1.5^2)$  so that a priori,  $\gamma_1$  is most likely between  $-5$  and  $0$ . However, there is still a small probability (5%) for  $\gamma_1$  to take positive values to allow the possibility that data contradicts the ordering assumption.

### 2.3. Subgroup-specific Optimal Dose

We measure a dose's desirability using a utility function that accounts for the risk-benefit tradeoff underlying medical decisions in practice. Using the motivating trial as an example, we elicit the utility from physicians as follows: we fix the desirability of the most desirable outcome pair (i.e.,  $Y_T = 0, Y_E = 3$ ) as  $\zeta(0, 3) = 100$  and the desirability of the least desirable outcome pair (i.e.,  $Y_T = 1, Y_E = 1$ ) as  $\zeta(1, 1) = 0$ , and then ask physicians to use these two pairs as references to score the desirability of the other elementary outcome pairs  $\{\zeta(Y_T, Y_E), Y_T = 0, 1, Y_E = 1, 2, 3\}$  using the scale of  $(0, 100)$ . This approach has been used in previous trial designs.<sup>12,39</sup> The elicited utility for our motivating trial is shown in Table 1.

Then for patients in biomarker subgroup  $Z$ , the true utility for a given dose  $d$  is

$$U_{true}(d, Z) = \sum_{Y_T=0}^1 \sum_{Y_E=1}^3 \zeta(Y_T, Y_E) \Pr(Y_T, Y_E | d, Z, \Theta). \quad (9)$$

Since the true mean utility  $U_{true}(d, Z)$  depends on unknown parameters  $\Theta$ , we need to estimate it based on the observed data. During the trial, given the interim data  $\mathcal{D}_n$  collected from  $n$  patients, we evaluate the desirability of dose  $d$  for patients in biomarker subgroup  $Z$  based on the posterior mean utility,

$$U_n(d, Z) = \sum_{Y_T=0}^1 \sum_{Y_E=1}^3 \zeta(Y_T, Y_E) E_{\Theta}[\Pr\{Y_T, Y_E | d, Z, \Theta\} | \mathcal{D}_n].$$

Let  $\pi_T(Z, d) \equiv \Pr(Y_T = 1 | Z, d)$  denote the toxicity rate and  $\pi_E(Z, d) \equiv \Pr(Y_E > 1 | Z, d)$  denote the response rate of SD/PR/CR for subgroup  $Z$  at dose  $d$ . Let  $\phi_T$  denote the upper limit of the toxicity rate, and  $\phi_E$  denotes the lower limit of the response rate specified by physicians. We define the subgroup-specific optimal dose for subgroup  $Z$  as the dose with the highest utility  $U_{true}(d, Z)$  while satisfying  $\pi_T < \phi_T$  and  $\pi_E > \phi_E$ . The definition of acceptable efficacy accounts for the practical consideration that for an immunotherapeutic agent, a dose is considered promising if it can achieve a certain rate of CR/PR or a certain rate of SD (if the rate of CR/PR is low). SD is often regarded as a favorable outcome for immunotherapy as many immunotherapeutic agents prolong survival without much tumor shrinkage.



## 2.4. Dose-finding Algorithm

At the beginning of an early-phase clinical trial, decision-making is difficult because of the small available data. This is especially true for dose findings involving subgroups. To alleviate this issue, we propose to use a two-stage dose-finding algorithm. In stage I, we perform dose escalation based on only the toxicity outcome without considering the immune response or efficacy. This stage aims to quickly explore the dose space and accumulate preliminary data to facilitate stage II model-based dose-finding. Stage I dose-escalation continues until reaching the first unsafe dose for either subgroup. Specifically, we treat the first cohort of patients at the lowest dose level and then escalate until we reach a dose that violates the safety requirement  $\Pr(\pi_T(Z, d) < \phi_T | \mathcal{D}_n) > C_I$  for either subgroup  $Z$ , where  $C_I$  is a probability cutoff to be tuned through simulation studies. Due to the sparse data in stage I, we evaluate the safety requirement based on the simple Beta-Binomial model in stage I. Precisely, for subgroup  $Z$  at dose  $d$ , we assume that the toxicity probability  $\pi_T(Z, d)$  follows a Beta prior distribution  $\text{Beta}(\xi_1, \xi_2)$ , and the number of patients who experience toxicity  $m_d^Z$  among  $n_d^Z$  treated patients follows a binomial distribution  $\text{Binom}(n_d^Z, \pi_T(Z, d))$ . We set  $\xi_1 = 0.1$  and  $\xi_2 = 0.2$  so that the prior is vague and the observed data dominate the posterior distribution. Under the Beta-Binomial model, the posterior distribution of  $\pi_T(Z, d)$  is  $\text{Beta}(\xi_1 + m_d^Z, \xi_2 + n_d^Z - m_d^Z)$ . When we reach a dose that violates the safety requirement or reaches the highest dose  $d_D$ , stage I completes, and the trial moves on to stage II. In stage I dose-escalation, dose skipping is not allowed for either subgroup, so if none of the patients in a cohort is in subgroup  $Z$ , then the current dose is repeated for subgroup  $Z$  in the next cohort. Although the stage I dose-escalation strategy does not utilize patient efficacy data or immune response, these data are collected to facilitate the model fitting in stage II.

At stage II, we adaptively randomize incoming patients to admissible doses that satisfy the safety and efficacy requirements. Based on interim data  $\mathcal{D}_n$ , we define a dose  $d$  as admissible for biomarker subgroup  $Z$  if it satisfies both the safety requirement

$$\Pr(\pi_T < \phi_T | Z, d, \mathcal{D}_n) > C_T \quad (10)$$

and the efficacy requirement

$$\Pr(\pi_E > \phi_E | Z, d, \mathcal{D}_n) > C_E, \quad (11)$$

where  $C_T$  and  $C_E$  are prespecified toxicity and efficacy cutoffs which should be calibrated to obtain good design operating characteristics. We denote the set of admissible doses by  $\mathcal{A}_{Z,n}$  for subgroup  $Z$ .

Stage II of the design can be summarized as follows.

1. Based on the currently observed data  $\mathcal{D}_n$ , determine the admissible set  $\mathcal{A}_{Z,n}$  for each subgroup based on safety and efficacy requirements (10) and (11).
2. If  $\mathcal{A}_{Z,n}$  is empty for both subgroups, terminate the trial and conclude that no dose is acceptable for both subgroups. If  $\mathcal{A}_{Z,n}$  is empty for one of the subgroups,



then conclude no dose is acceptable for that subgroup, and future patients in that subgroup are treated off protocol.

3. If  $\mathcal{A}_{Z,n}$  is not empty for at least one subgroup, then the patients in the next cohort in subgroup  $Z$  with non-empty  $\mathcal{A}_{Z,n}$  are adaptively randomized to dose  $\{d_j, d_j \in \mathcal{A}_{Z,n}\}$  with randomization probability  $\pi_j$  proportional to its posterior mean utility, i.e.,

$$\pi_j = \frac{U_n(d_j, Z)}{\sum_{d_{j'} \in \mathcal{A}_{Z,n}} U_n(d_{j'}, Z)}.$$

4. Repeat the above steps until reaching the maximum sample size  $N$  or terminating the trial early. Select the admissible dose with the largest posterior mean utility  $U_N(d, Z)$  as the subgroup-specific optimal dose for subgroup  $Z$ .

For safety considerations, any untried dose cannot be skipped when escalating.

### 3. Simulation

#### 3.1. Main Results

We conducted comprehensive simulation studies to evaluate the performance of the BSOI design. Taking the setting of the motivating trial, we considered five doses (0.1, 0.3, 0.5, 0.7, 0.9), with a maximum sample size of 60 and a cohort size of 3. The upper limit for toxicity  $\phi_T = 0.3$ , the lower limit for efficacy was  $\phi_E = 0.3$ , and the marker-positive prevalence  $\phi = 0.5$ . We set  $\hat{\alpha} = 20$  as a prior estimate of the maximum immune response for marker-negative patients elicited from clinicians, and  $\tau = 3\hat{\alpha}$  to obtain a vague prior for  $\alpha$ ,  $\text{Gamma}(1/9, 1/180)$ , so that the prior mean was 20. The prior standard deviation was 60, which was 3 times the prior mean. The prior for  $\delta$  was taken to be  $\text{Normal}(\log(1.5), 0.25^2)$  to match the physician-elicited ratio of the maximum immune response of the marker-positive patients relative to the marker-negative patients of 1.5 and 5% probability on negative values. We took probability cutoffs  $C_I = 0.3$ ,  $C_T = 0.12$  and  $C_E = 0.05$ . The utility elicited from physicians is displayed in Table 1.

We considered 14 scenarios in our simulation studies. For each scenario, we first specified the mean immune response  $\mu(d_i, Z)$  at each dose level  $d_i$  for each subgroup. Given these mean immune responses, the immune response  $Y_{I,i}$  for patient  $i$  treated at dose  $d_i$  was generated from the normal model (1). Conditional on  $Y_{I,i}$  the toxicity and efficacy outcomes  $Y_{T,i}$  and  $Y_{E,i}$  were simulated from models (3) and (4), respectively. The 14 scenarios varied in the number of target doses, location of the target doses, and the patterns of toxicity, efficacy, and immune response (see Table 2 and Table S1 in the Supplementary Materials). Figure 2 shows the true dose-response curves for the immune response, toxicity, and efficacy for the two subgroups under these scenarios. Data sharing is not applicable to this paper as all data in this article is computer simulated.

An important feature of the proposed BSOI design is its ability to borrow information across subgroups. To demonstrate this advantage, we compared the BSOI design with a separate

design approach that implemented the design described in Liu, Guo, and Yuan<sup>13</sup> (referred to as LGY design hereafter) separately for each subgroup. To make the comparisons more meaningful, we used model (2) for the immune response in the separate design approach, instead of the Emax model in LGY design. Specifically, for each subgroup  $Z = 0$  or  $1$ , we modeled the mean immune response at dose  $d_i$ ,  $\mu_I^Z(d_i)$ , as

$$\log \left\{ \frac{\mu_I^Z(d_i)}{\alpha_Z - \mu_I^Z(d_i)} \right\} = \eta_{0Z} + \eta_{1Z} d_i, \tag{12}$$

The prior distribution was Gamma(1/9, 1/180) for  $\alpha_0$  and Gamma(1/9, 1/270) for  $\alpha_1$  so that the elicited prior information was matched to that in the BSOI design. That is, the prior estimate of the maximum immune response was 20 for marker-negative patients and 30 for marker-positive patients; and the prior standard deviation was three times of the prior mean for each subgroup.  $\eta$ 's were also modeled separately for the two subgroups. We also compared the BSOI design with two conventional phase I/II designs. The first conventional design (conventional I) considered only efficacy and toxicity, as in most existing phase I/II designs. To make a fair comparison, we used the same toxicity and efficacy models as the proposed design, but with the immune response term dropped such that

$$\text{logit}\{\Pr(Y_{T,i} = 1 \mid d_i, Z_i, Y_{I,i})\} = \beta_{0,Z_i} + \beta_1 d_i \tag{13}$$

$$\text{logit}\{\Pr(Y_{E,i} \leq j \mid d_i, Z_i, Y_{I,i})\} = \gamma_{0,j} + \gamma_1 Z_i + \gamma_2 d_i + \gamma_3 d_i^2 \quad j = 1, \dots, J - 1 \tag{14}$$

To demonstrate the importance of accounting for the subgroup effect in immunotherapy dose finding, the second conventional design (conventional II) did not include the subgroup effects. Precisely, the immune response, toxicity, and efficacy models were

$$\log \left\{ \frac{\mu_I(d_i, Z_i)}{\alpha - \mu_I(d_i, Z_i)} \right\} = \eta_0 + \eta_1 d_i, \tag{15}$$

$$\text{logit}\{\Pr(Y_{T,i} = 1 \mid d_i, Z_i, Y_{I,i})\} = \beta_0 + \beta_1 d_i + \beta_2 Y_{I,i} \tag{16}$$

$$\text{logit}\{\Pr(Y_{E,i} \leq j \mid d_i, Z_i, Y_{I,i})\} = \gamma_{0,j} + \gamma_2 Y_{I,i} + \gamma_3 Y_{I,i}^2 \quad j = 1, \dots, J - 1 \tag{17}$$

Table 3 summarizes the operating characteristics of the BSOI design, the separate design approach, and the two conventional designs for the first 8 scenarios. The operating characteristics for scenarios 9 to 14 are provided in Table S2 in the Supplementary materials. In scenarios 1–7, each subgroup has only one target dose. In the first three scenarios, the target dose is the same for the two subgroups. In scenarios 1 and 2, both toxicity and immune response increase with dose, while efficacy almost remains constant across doses under scenario 1 and first increases and then decreases under scenario 2. Scenario 3 is a particular case where the two subgroups have the same dose-response relationships. In scenario 3, toxicity increases with dose, while both immune response and efficacy first

increase and then plateau. For these three scenarios, the BSOI design outperformed the separate design approach and the conventional I design and yielded comparable performance with the conventional II design regarding the percentage of correct optimal dose selection (PCS). For example, for scenario 1, the optimal dose is dose level 1 for both subgroups. The PCS of the target dose was 72.6% and 81.6% for marker-negative and marker-positive subgroups, respectively, under the BSOI design; 66.4% and 81.7% under the separate design approach; 60.8% and 72.4% under the conventional I design that ignored the immune response; and 76.2% and 76.2% under the conventional II design that ignored the subgroups. The proposed BSOI design treated a comparable number of patients at the target doses with the conventional designs. For example, in scenario 3, the BSOI design assigned 8.2 and 8 patients at the target doses for the two subgroups; the separate design approach treated 6.6 and 6.5 patients; and the two conventional designs assigned 7.6 and 7.4, and 8.3 and 7.9 patients on average.

In scenarios 4–7, each subgroup has only one target dose, but the target doses are different for the two subgroups. For these four scenarios, the proposed design resulted in higher PCS of the target doses and at least a comparable number of patients treated at the target doses than the separate design approach and both conventional designs. For example, in scenario 6, the target dose is dose level 3 for marker-negative patients and dose level 2 for marker-positive patients. The proposed design yielded much higher PCS of the target doses than the other three designs (85.8% and 71.2% under the proposed design, 69% and 58.9% under the separate designs, 68.8% and 21%, and 60.4% and 36.4% under the two conventional designs), and assigned more patients to the target doses (7.9 and 8.6 under the proposed design, 6.3 and 8.5 under the separate designs, and 7.6 and 7.5, and 7.3 and 7.7 under the two conventional designs).

In scenario 8, both subgroups have two target doses; in scenarios 9–12, one subgroup has one target dose and the other subgroup has two target doses. In these scenarios, the BSOI design resulted in overall better performance than the other three designs, and the improvement can be substantial. For example, in scenario 8, the target doses are dose levels 4 and 5 for marker-negative patients and dose levels 2 and 3 for marker-positive patients. The PCS under BSOI was 63.8% for marker-negative patients and 72% for marker-positive patients. The counterpart PCS were 58.8% and 61.7% for the separate designs, 24.8% and 48.2% for the conventional I design and 41.8% and 48.4% for the conventional II design.

In scenario 13, none of the doses is admissible for marker negative patients because the efficacy is lower than the efficacy lower bound  $\phi_E = 0.3$  at all dose levels, while the optimal doses are dose levels 3 and 4 for marker-positive patients. Consequently, no dose should be recommended for marker-negative patients. The proposed BSOI design correctly concluded none of the doses was acceptable for marker-positive subgroup 96.2% of the time, while this number was 100%, 89.6% and 55% for the other three designs. For marker-negative subgroup, the PCS was 59.8%, 51.8%, 51.4%, and 25.8% for the four designs. In scenario 14, all dose levels are too toxic for both subgroups. The trial was terminated early 100% of the time under all four designs.

### 3.2. Sensitivity Analyses

We carried out sensitivity analyses to assess the robustness of our proposed design's performance by using 1) less informative prior distributions, 2) an alternative data generating process, 3) alternative elicited values, 4) alternative utility functions, 5) smaller sample sizes, and 6) alternative marker prevalence ratios.

We made all the prior distributions more non-informative. Specifically, we assigned a  $\text{Gamma}(0.04, 0.002)$  prior for  $\alpha$  so that the prior standard deviation was five times the prior mean, i.e.,  $5\hat{\alpha} = 100$ .  $\sigma^2$  was assigned  $\text{InverseGamma}(0.01, 0.01)$  prior. The priors were set to be  $N(0, 5^2)$  for  $\beta_1, \beta_2, \gamma_{0,1}, \gamma_{0,2}, \gamma_2, \gamma_3$ ; truncated  $N(0, 5^2)$  for  $\eta_1$ , so that for each of these parameters, the prior scale parameter was twice the previous value. The results (see Table 4 and Table S5 in the Supplementary Materials) are very similar to those reported in Table 3, suggesting that the proposed design is robust to the choice of the prior distributions.

We conducted additional simulation studies to investigate the performance of the proposed BSOI design when data were generated from alternative models. Specifically, we matched the mean and variance of the immune response at each dose to those shown in Table 2. For toxicity, we used the probit model

$$\Phi^{-1}\{\Pr(Y_{T,i} = 1 \mid d_i, Z_i, Y_{I,i})\} = \beta_{0,Z_i} + \beta_1 d_i + \beta_2 Y_{I,i} \quad (18)$$

For efficacy, we used the cumulative probit model

$$\frac{\Phi^{-1}\{\Pr(Y_{E,i} \leq j \mid d_i, Z_i, Y_{I,i})\}}{-1} = \gamma_{0,j} + \gamma_1 Z_i + \gamma_2 Y_{I,i} + \gamma_3 Y_{I,i}^2 \quad j = 1, \dots, J \quad (19)$$

The parameters in the toxicity and efficacy models were selected to match the toxicity and efficacy probabilities and the utilities shown in Table 2. Both the admissible dose set and the optimal dose are the same as in the original setting. The operating characteristics of the proposed model are shown in Table 5 and Table S6 in the Supplementary Materials. As can be seen, the results are very similar to the original results in Table 3.

The BSOI design requires elicitation of prior estimates of  $\alpha$  and  $\delta$  from clinicians. Due to the typically small sample sizes of early phase trials, it is common practice for these trials to incorporate prior knowledge elicited from clinicians (e.g., O'Quigley, Pepe, and Fisher<sup>40</sup>; Thall and Cook<sup>8</sup>, Jin et al.<sup>10</sup>). To evaluate the robustness of the BSOI design to different values of these elicited values, we performed sensitivity analysis with alternative prior estimates of  $\alpha$  and  $\delta$ . In our original simulation setting, we set  $\hat{\alpha} = 20$  and  $\hat{r} = 1.5$ . For sensitivity analyses, we evaluated two alternative sets of prior estimates. For the first set,  $\hat{\alpha} = 15$  and  $\hat{r} = 1.3$ , with corresponding prior distributions  $\alpha \sim \text{Gamma}(1/9, 1/135)$  and  $\delta \sim \text{Normal}(0.26, 0.16^2)$ . For the second set,  $\hat{\alpha} = 25$  and  $\hat{r} = 1.7$ , with corresponding prior distributions  $\alpha \sim \text{Gamma}(1/9, 1/225)$  and  $\delta \sim \text{Normal}(0.53, 0.32^2)$ . The operating characteristics of the BSOI design with these alternative elicited prior estimates are given in Tables S7 and S8 in the Supplementary Materials. As these tables show, the new results are very similar to the original results given in Table 3, suggesting that the BSOI design is not sensitive to the elicited values as long as the prior distributions are very vague.

We also evaluated the robustness of the BSOI design with respect to the specification of the utility function. We considered two alternative utility functions displayed in Table S9 in the Supplementary Materials. Utility function 1 gave lower desirability scores to DLT outcomes than the original utility function whereas utility function 2 took the opposite way. As shown in Tables S10 and S11 in the Supplementary Materials, the results under these alternative utility functions are generally similar to those in Table 3, showing that the BSOI design is not sensitive to the utility function.

When the maximum sample size dropped from 60 to 45, our design's performance was slightly worse, as summarized in Table 6 and Table S3 in the Supplementary Materials, but the selection percentage of the target dose was still the highest among all doses. When the maximum sample size was 36, i.e., 18 patients in each subgroup, the BSOI design also resulted in a reasonable performance (See Table S4 in the Supplementary Materials).

Finally, we examined the sensitivity of the BSOI design to the marker prevalence ratio. We evaluated two additional values of marker-positive prevalence, 0.3 and 0.7 (18 and 42 patients in the two subgroups). Results, given in Tables S12 and S13 in the Supplementary Materials, are very similar to the original ones in Table 3 when the prevalence is 0.5.

The BSOI design may not always yield the highest selection percentage of the target dose if the sample size in a subgroup is less than 18, e.g., when the maximum sample size is 60 with marker prevalence ratio 0.25, or when the maximum sample size is 30 with marker prevalence ratio 0.5 (results not shown). In addition, the BSOI design may not be able to distinguish between doses with very similar utility values. For example, if the utility of a dose is within 95% of the highest utility value (at the target dose), the BSOI design may select that dose with a similar probability as the target dose. However, this is a reasonable result as that dose can as well be considered the target dose because its utility is very close to that of the target dose. In summary, we recommend to use the BSOI design when the sample size in each subgroup is at least 18. Under this setting, the BSOI design performs well under a wide range of situations as long as the priors are relatively vague and the utility function is elicited from physicians and/or patients to correctly account for the risk-benefit tradeoff that underlies medical decisions in practice.

#### 4. Discussion

We have proposed the BSOI design, a Bayesian adaptive phase I/II clinical trial design, to determine the subgroup-specific optimal dose for immunotherapy. We develop a parsimonious yet flexible model to jointly model the continuous immune response, binary toxicity, and ordinal efficacy, and utilize a utility function to quantify the doses' desirability within each biomarker subgroup. We propose a two-stage dose-finding algorithm to assign patients to desirable doses based on their biomarker status and select the subgroup-specific optimal doses for future patients. Simulation studies show that the proposed design has desirable operating characteristics. It reliably identifies the subgroup-specific optimal dose with a large probability and assigns more patients to the optimal doses.

The BSOI design is an extension of the LGY design by incorporating the subgroup structure. Due to the typically small sample sizes of early phase trials, especially when subgroups are involved, the BSOI design utilizes the more parsimonious plateau model (2) for the immune response (3 parameters with subgroup structure) instead of the more parameterized Emax model in LGY design. The Emax model in the original LGY design has 4 parameters. Under the current setting, the baseline immune activity can be set to 0 as the mean immune response is expected to be 0 when dose is 0. Therefore, there are 3 parameters without subgroups, so at least 4 parameters with subgroups involved under the Emax model. The toxicity and efficacy models (3) and (4) are also parsimonious to accommodate the small sample sizes while allowing for strength borrowing across subgroups. For the same reason, unlike the LGY design that employs a single stage design, the BSOI design implements a two-stage dose-finding design to alleviate the issue of difficult decision making at the beginning of the trial. The preliminary data collected from Stage I facilitate Stage II model-based dose finding.

In this article, we assume there are two pre-determined subgroups. In some situations, the subgroups may be unknown. It is of great interest to incorporate multiple biomarkers into the trial design and identify subgroups with different dose-response relationships. This is a topic of our future research.

As in most current phase I/II trial designs, the proposed design assumes that the outcomes are quickly ascertainable. The decision rules can be applied to determine the following new patient's dose assignment once (s)he enters the trial. However, late-onset outcomes are often expected in dose-finding studies such that responses often occur long after the treatment is finished. In the presence of such late-onset outcomes, a direct application of the proposed design may result in biased results. One practical solution to this late-onset outcome issue is to apply the methods proposed by Liu, Yin and Yuan,<sup>41</sup> Jin et al.,<sup>10</sup> or Zhang and Zang,<sup>42</sup> which accommodate delayed response outcomes using Bayesian data augmentation, expectation-maximization algorithm, or the conditional weighted likelihood method.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors thank two reviewers for their helpful comments which substantially improved the presentation of this paper. Beibei Guo's research was partially supported by the R & D Research Competitiveness Subprogram of Louisiana Board of Regents, Contract number LEQSF(2020-21)-RD-A-04. Yong Zang's research was partially supported by NIH/NCI grants P30 CA082709; R21 CA264257 and the Ralph W. and Grace M. Showalter Research Trust award.

## References

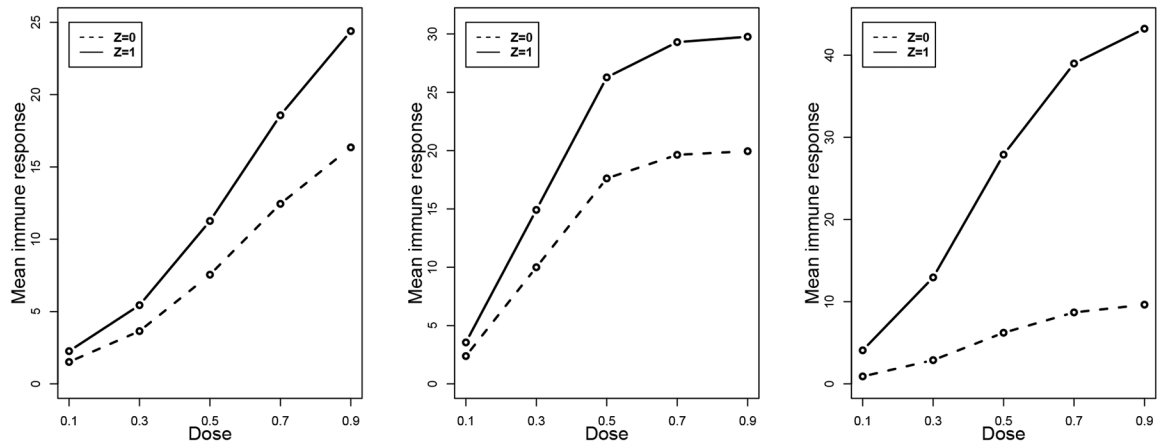
- [1]. Couzin-Frankel J (2013), Cancer immunotherapy. *Science*, 324, 1432–1433.
- [2]. Topalian SL, Weiner GJ and Pardoll DM (2011), Cancer immunotherapy comes of age. *Journal of Clinical Oncology*, 23, 4828–4836.



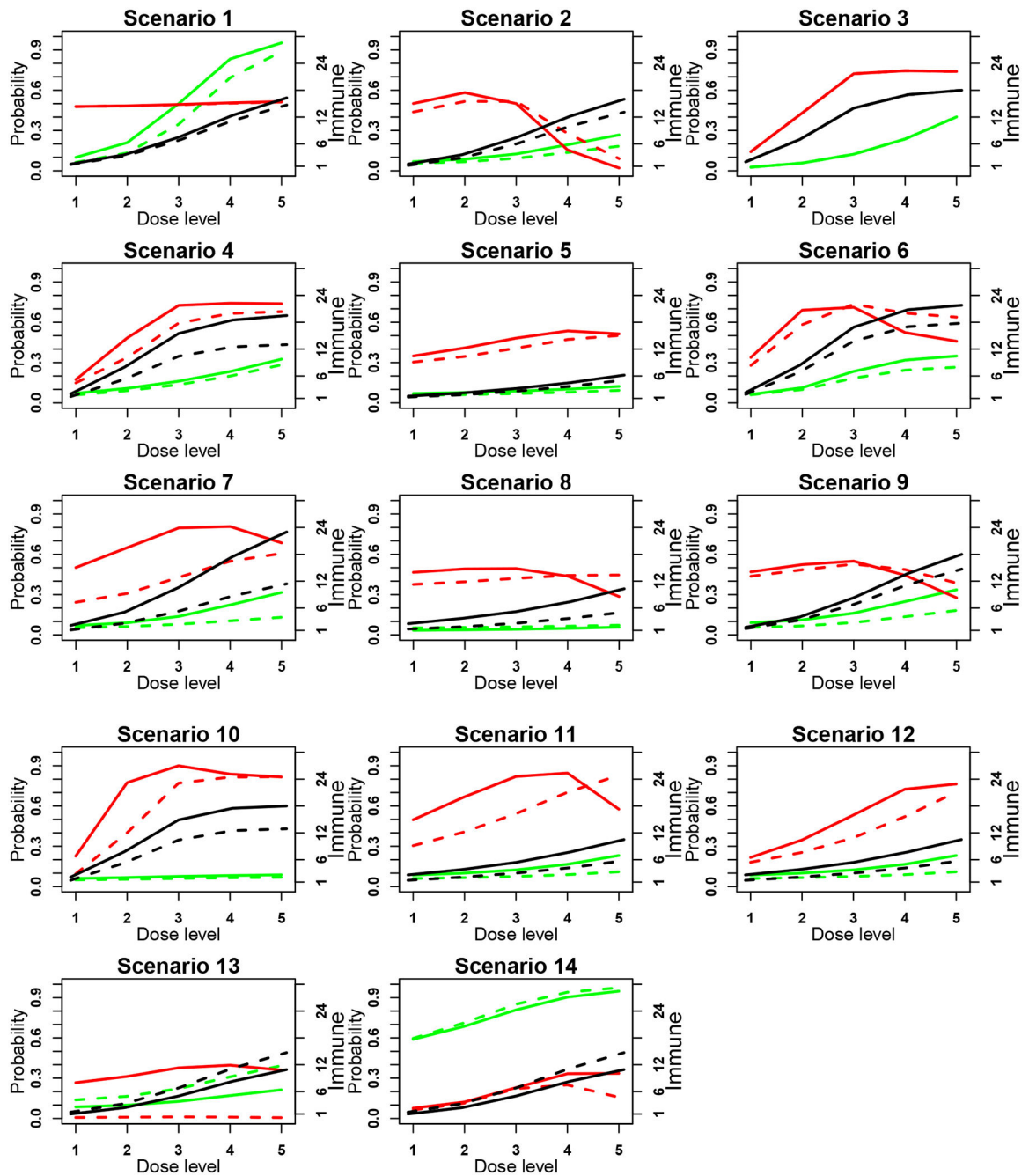
- [3]. Makkouk A, and Weiner GJ (2015), Cancer immunotherapy and breaking immune tolerance: new approaches to an old challenge. *Cancer Research*, 75, 5–10. [PubMed: 25524899]
- [4]. Hodi FS., O'Day SJ., McDermott DF., et al. (2010), Improved survival with ipilimumab in patients with metastatic melanoma. *The New England Journal of Medicine*, 363: 711–723. [PubMed: 20525992]
- [5]. Borghaei H, Paz-Ares L, Horn L, et al. (2015), Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *The New England Journal of Medicine*, 373: 1627–39. [PubMed: 26412456]
- [6]. Reck M, Rodriguez-Abreu D, Robinson AG. et al. , (2016) Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. *The New England Journal of Medicine* 375: 1823–1833. [PubMed: 27718847]
- [7]. Thall P, Russell K (1998) A strategy for dose-finding and safety monitoring based on efficacy and adverse outcomes in phase I/II clinical trials. *Statistics in Medicine*, 27, 4895–4913.
- [8]. Thall P, Cook J (2004) Dose-finding based on efficacy-toxicity trade-offs. *Biometrics*, 60, 684–693. [PubMed: 15339291]
- [9]. Yin G, Li Y, Ji Y (2006) Bayesian dose-finding in phase I/II clinical trials using toxicity and efficacy odds ratios. *Biometrics*, 62, 777–784. [PubMed: 16984320]
- [10]. Jin I, Liu S, Thall P, Yuan Y (2014) Using Data Augmentation to Facilitate Conduct of Phase I/II Clinical Trials with Delayed Outcomes. *Journal of the American Statistical Association*, 109: 525–536. [PubMed: 25382884]
- [11]. Zang Y, Lee JJ. (2017) A robust two-stage design identifying the optimal biological dose for phase I/II clinical trials. *Statistics in Medicine* 36: 27–42. [PubMed: 27538818]
- [12]. Guo B, Yuan Y (2017) Bayesian phase I/II biomarker-based dose finding for precision medicine with molecularly targeted agents. *Journal of the American Statistical Association*, 112: 508–520. [PubMed: 32863478]
- [13]. Liu S, Guo B, Yuan Y (2017) A Bayesian Phase I/II Design for Immunotherapy Trials. *Journal of the American Statistical Association*, 113: 1016–1027.
- [14]. Guo B, Li D, Yuan Y (2018) SPIRIT: A Seamless Phase I/II Randomized Design for Immunotherapy Trials. *Pharmaceutical Statistics*, 17: 527–540. [PubMed: 29882388]
- [15]. Guo B, Park Y, Liu S (2019) Utility-based Bayesian phase I/II design for immunotherapy trials with progression-free survival endpoint. *Journal of the Royal Statistical Society: Series C*, 68: 411–425.
- [16]. Guo B, Garrett E, Liu S (2021) A Bayesian phase I/II design for cancer clinical trials combining an immunotherapeutic agent with a chemotherapeutic agent. *Journal of the Royal Statistical Society: Series C*, In Press.
- [17]. Zhong W, Koopmeiners J, Carlin B (2012) A trivariate continual reassessment method for phase I/II trials of toxicity, efficacy, and surrogate efficacy. *Statistics in Medicine* 31: 3885–3895. [PubMed: 22807126]
- [18]. Topalian S, Hodi F, Brahmer J, et al. (2012), Safety, Activity, and Immune Correlates of Anti-PD-1 Antibody in Cancer. *The New England Journal of Medicine*, 366: 2443–2454. [PubMed: 22658127]
- [19]. Manhoney KM, Atkins MB (2014), Prognostic and predictive markers for the new immunotherapies. *Oncology*, 28: 39–48. [PubMed: 25384886]
- [20]. Garon EB, Rizvi NA, Hui R, et al. (2015), Pembrolizumab for the treatment of non-small-cell lung cancer. *The New England Journal of Medicine*, 372: 2018–28. [PubMed: 25891174]
- [21]. Larkin J, Chiarion-Sileni V, Gonzales R, et al. (2015), Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. *The New England Journal of Medicine*, 373: 23–34. [PubMed: 26027431]
- [22]. Aguiar P, Santoro S, Tadokoro H, et al. (2016), The role of PD-L1 expression as a predictive biomarker in advanced non-small-cell lung cancer: a network meta-analysis. *Immunotherapy*, 8, 479–488. [PubMed: 26973128]
- [23]. Bai R, Lv Z, Xu D, Cui J (2020), Predictive biomarkers for cancer immunotherapy with immune checkpoint inhibitors. *Biomarker Research*, 8–34. [PubMed: 32128213]



- [24]. Kugel CH., Douglass SM., Webster MR., et al. (2018), Age correlates with response to anti-PD1, reflecting age-related differences in intratumoral effector and regulatory T-cell populations. *Clinical Cancer Research*, 24: 5347–56. [PubMed: 29898988]
- [25]. Nishijima TF., Muss HB., Shachar SS., et al. (2016), Comparison of efficacy of immune checkpoint inhibitors (ICIs) between younger and older patients: a systematic review and meta-analysis. *Cancer Treatment Review*, 45: 30–7.
- [26]. Felip E, Ardizzoni A, Ciuleanu T et al. (2020), CheckMate 171: a phase 2 trial of nivolumab in patients with previously treated advanced squamous non-small cell lung cancer, including ECOG PS 2 and elderly population. *European Journal of Cancer*, 127, 160–72. [PubMed: 32028209]
- [27]. Murphy WJ., Longo DL. (2019), The surprisingly strong association between obesity and cancer immunotherapy efficacy. *JAMA*, 321: 1247–8. [PubMed: 30882850]
- [28]. Wang Z, Aguilar EG., Luna JL., et al. (2019), Paradoxical effects of obesity on T-cell function during tumor progression and PD-1 checkpoint blockade. *Nature Medicine*, 25: 141–51.
- [29]. Patnaik A, Kang SP, Rasco D, et al. (2015) Phase I study of pembrolizumab (MK-3475; anti-PD-1 monoclonal antibody) in patients with advanced solid tumors. *Clin Cancer Res*, 21(19):4286–4293. [PubMed: 25977344]
- [30]. Robert C, Ribas A, Wolchok JD, et al. (2014) Anti-programmed-death-receptor-1 treatment with pembrolizumab in ipilimumab-refractory advanced melanoma: A randomised dose-comparison cohort of a phase 1 trial. *Lancet*, 384(9948):1109–1117. [PubMed: 25034862]
- [31]. Ribas A, Puzanov I, Dummer R, et al. (2015) Pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory melanoma (KEYNOTE-002): A randomised, controlled, phase 2 trial. *Lancet Oncol*, 16(8):908–918. [PubMed: 26115796]
- [32]. Chatterjee M, Turner DC, Felip E, et al. (2016) Systematic evaluation of pembrolizumab dosing in patients with advanced non-small-cell lung cancer. *Ann Oncol*, 27(7):1291–1298. [PubMed: 27117531]
- [33]. Herbst RS, Baas P, Kim DW, et al. (2016) Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): A randomised controlled trial. *Lancet*, 387(10027): 1540–1550. [PubMed: 26712084]
- [34]. Goldstein D, Gordon N, Davidescu M et al. (2017) A Pharmacoeconomic Analysis of Personalized Dosing vs Fixed Dosing of Pembrolizumab in Firstline PD-L1-Positive Non-Small Cell Lung Cancer. *Journal of the National Cancer Institute*, 109(11): djx063.
- [35]. Ingles Garcés A, Au L, Mason R, Thomas J, Larkin J (2019), Building on the anti-PD1/PD-L1 backbone: combination immunotherapy for cancer. *Expert Opinion on Investigational Drugs*, 28: 695–708. [PubMed: 31359805]
- [36]. Cai C, Yuan Y, Ji Y (2014) A Bayesian dose finding design for oncology clinical trials of combinational biological agents. *Applied Statistics*, 63:159–173. [PubMed: 24511160]
- [37]. Robert C Casella G (2004). *Monte Carlo Statistical Methods*. New York: Springer-Verlag, 2nd edition.
- [38]. Gelman A, Jakulin A, Pittau MG, Su YS (2008) A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, 2: 1360–1383.
- [39]. Thall P, Nguyen (2012) Adaptive randomization to improve utility-based dose-finding with bivariate ordinal outcomes. *Journal of Biopharmaceutical Statistics*, 22: 785–801. [PubMed: 22651115]
- [40]. 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. [PubMed: 2350571]
- [41]. Liu S, Yin G, Yuan Y (2013) Bayesian Data Augmentation Dose Finding with Continual Reassessment Method and Delayed Toxicity. *Annals of Applied Statistics*, 4: 2138–2156.
- [42]. Zhang Y, Zang Y. (2021) CWL: A conditional weighted likelihood method to account for the delayed joint toxicity-efficacy outcomes for phase I/II clinical trials. *Statistical Methods in Medical Research* 30: 892–903. [PubMed: 33349166]



**Figure 1:** The mean immune response curves for the two subgroups generated from model (2). The parameters are  $a = 20, \delta = 0.4, \eta_0 = -3, \eta_1 = 5$ ;  $a = 20, \delta = 0.4, \eta_0 = -3, \eta_1 = 10$ ; and  $a = 10, \delta = 1.5, \eta_0 = -3, \eta_1 = 7$  for the three panels from left to right. Dashed and solid lines represent marker-negative ( $Z=0$ ) and marker-positive ( $Z=1$ ) subgroups, respectively.



**Figure 2:** Dose-response curves for the 14 scenarios in the simulation study. The dashed and solid lines are for subgroups  $Z=0$  and  $Z=1$ , respectively; and green, red, and black curves are the toxicity  $\pi_T$ , efficacy  $\pi_E$ , and mean immune response  $\mu_I$  curves, respectively. Toxicity and efficacy are plotted against the left y-axis, and the immune response is plotted against the right y-axis.

**Table 1:**

Utility for the motivating trial

	$Y_E=1$	$Y_E=2$	$Y_E=3$
$Y_T=0$	10	60	100
$Y_T=1$	0	20	30

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2:**

True mean immune response, toxicity, efficacy probabilities, and utility for the two subgroups at each dose for scenarios 1 to 8. The boldface numbers are the subgroup-specific optimal dose.

	Z=0					Z=1				
	Scenario 1									
$E(Y_d)$	<b>1.4</b>	3.3	6.8	11.2	14.7	<b>1.5</b>	3.6	7.5	12.4	16.3
$\pi_T$	<b>0.063</b>	0.133	0.346	0.693	0.887	<b>0.1</b>	0.21	0.5	0.832	0.953
$\pi_E$	<b>0.479</b>	0.483	0.492	0.503	0.512	<b>0.479</b>	0.484	0.494	0.506	0.516
$Pr(Y_E=3)$	<b>0.309</b>	0.311	0.314	0.319	0.322	<b>0.309</b>	0.311	0.315	0.32	0.324
$Pr(Y_E=2)$	<b>0.17</b>	0.173	0.178	0.184	0.19	<b>0.17</b>	0.173	0.179	0.186	0.192
Utility	<b>38.9</b>	37.1	31.3	21.3	15.6	<b>37.8</b>	34.9	26.8	17.1	13.7
	Scenario 2									
$E(Y_d)$	1.2	<b>2.9</b>	6	10	13.1	1.5	<b>3.6</b>	7.4	12.2	16
$\pi_T$	0.055	<b>0.067</b>	0.093	0.137	0.183	0.069	<b>0.086</b>	0.125	0.194	0.267
$\pi_E$	0.438	<b>0.517</b>	0.516	0.278	0.091	0.501	<b>0.582</b>	0.5	0.155	0.021
$Pr(Y_E=3)$	0.278	<b>0.318</b>	0.317	0.188	0.066	0.305	<b>0.339</b>	0.304	0.109	0.016
$Pr(Y_E=2)$	0.159	<b>0.199</b>	0.199	0.09	0.025	0.197	<b>0.243</b>	0.196	0.046	0.005
Utility	36.7	<b>41.7</b>	40.9	24.6	13.1	40.8	<b>45.8</b>	39	16.5	8.4
	Scenario 3									
$E(Y_d)$	2	7	<b>14</b>	17	18	2	7	<b>14</b>	17	18
$\pi_T$	0.027	0.058	<b>0.123</b>	0.237	0.401	0.027	0.058	<b>0.123</b>	0.237	0.401
$\pi_E$	0.143	0.431	<b>0.723</b>	0.745	0.74	0.143	0.431	<b>0.723</b>	0.745	0.74
$Pr(Y_E=3)$	0.133	0.389	<b>0.605</b>	0.617	0.614	0.133	0.389	<b>0.605</b>	0.617	0.614
$Pr(Y_E=2)$	0.009	0.042	<b>0.118</b>	0.128	0.126	0.009	0.042	<b>0.118</b>	0.128	0.126
Utility	17.1	31.8	<b>46.5</b>	43.8	37.6	17.1	31.8	<b>46.5</b>	43.8	37.6
	Scenario 4									
$E(Y_d)$	1.4	5.3	10.4	<b>12.5</b>	13	2	8	<b>15.5</b>	18.5	19.5
$\pi_T$	0.059	0.09	0.136	<b>0.199</b>	0.283	0.071	0.108	<b>0.161</b>	0.233	0.325
$\pi_E$	0.149	0.338	0.594	<b>0.666</b>	0.679	0.173	0.484	<b>0.726</b>	0.742	0.738
$Pr(Y_E=3)$	0.123	0.266	0.418	<b>0.445</b>	0.449	0.143	0.36	<b>0.459</b>	0.461	0.46
$Pr(Y_E=2)$	0.026	0.072	0.176	<b>0.221</b>	0.23	0.031	0.124	<b>0.267</b>	0.281	0.278
Utility	17.6	27.8	42.3	<b>44.8</b>	42.6	18.7	36.1	<b>50.5</b>	48.8	44.9
	Scenario 5									
$E(Y_d)$	1.3	1.8	2.6	3.7	<b>5.1</b>	1.5	2.2	3.2	<b>4.5</b>	6.2
$\pi_T$	0.057	0.062	0.069	0.079	<b>0.093</b>	0.069	0.077	0.087	<b>0.102</b>	0.122
$\pi_E$	0.304	0.345	0.407	0.472	<b>0.501</b>	0.349	0.409	0.482	<b>0.535</b>	0.513
$Pr(Y_E=3)$	0.219	0.248	0.289	0.332	<b>0.351</b>	0.244	0.282	0.328	<b>0.361</b>	0.348
$Pr(Y_E=2)$	0.084	0.097	0.118	0.14	<b>0.15</b>	0.105	0.126	0.153	<b>0.174</b>	0.165
Utility	27.3	29.7	33.3	36.9	<b>38.3</b>	30	33.5	37.7	<b>40.5</b>	38.6
	Scenario 6									

	Z=0					Z=1				
$E(Y_{\hat{d}})$	1.9	6.9	<b>13.8</b>	17	17.8	2.3	<b>8.4</b>	16.9	20.8	21.8
$\pi_T$	0.06	0.099	<b>0.184</b>	0.243	0.266	0.063	<b>0.113</b>	0.234	0.318	0.349
$\pi_E$	0.279	0.582	<b>0.734</b>	0.668	0.638	0.338	<b>0.69</b>	0.711	0.523	0.459
$Pr(Y_E=3)$	0.253	0.504	<b>0.609</b>	0.566	0.544	0.304	<b>0.575</b>	0.588	0.455	0.404
$Pr(Y_E=2)$	0.025	0.079	<b>0.126</b>	0.103	0.094	0.034	<b>0.116</b>	0.123	0.068	0.055
Utility	23.8	39.3	<b>45.1</b>	39.5	37.2	26.9	<b>45.3</b>	42.3	30.1	26.3
Scenario 7										
$E(Y_{\hat{d}})$	1.1	2.6	5.3	8.7	<b>11.4</b>	2.1	5.1	<b>10.6</b>	17.5	23
$\pi_T$	0.054	0.062	0.079	0.105	<b>0.131</b>	0.07	0.09	<b>0.138</b>	0.223	0.316
$\pi_E$	0.243	0.309	0.428	0.548	<b>0.607</b>	0.501	0.65	<b>0.797</b>	0.807	0.685
$Pr(Y_E=3)$	0.171	0.212	0.278	0.33	<b>0.348</b>	0.301	0.34	<b>0.326</b>	0.321	0.343
$Pr(Y_E=2)$	0.071	0.097	0.15	0.218	<b>0.258</b>	0.201	0.31	<b>0.471</b>	0.486	0.342
Utility	23.9	27.9	35.2	42.7	<b>46</b>	40.9	51.4	<b>62.1</b>	58.9	45
Scenario 8										
$E(Y_{\hat{d}})$	1.3	1.8	2.6	<b>3.7</b>	<b>5.1</b>	2.5	<b>3.7</b>	<b>5.2</b>	7.4	10.3
$\pi_T$	0.052	0.055	0.059	<b>0.065</b>	<b>0.072</b>	0.034	<b>0.037</b>	<b>0.042</b>	0.048	0.057
$\pi_E$	0.376	0.395	0.421	<b>0.443</b>	<b>0.446</b>	0.466	<b>0.492</b>	<b>0.494</b>	0.437	0.285
$Pr(Y_E=3)$	0.296	0.31	0.328	<b>0.344</b>	<b>0.346</b>	0.356	<b>0.373</b>	<b>0.375</b>	0.337	0.228
$Pr(Y_E=2)$	0.08	0.085	0.093	<b>0.099</b>	<b>0.1</b>	0.11	<b>0.118</b>	<b>0.119</b>	0.1	0.058
Utility	30.7	31.8	33.2	<b>34.4</b>	<b>34.4</b>	36.8	<b>38.3</b>	<b>38.3</b>	34.6	25.4

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3:**

Selection percentage and the average number of patients treated at each dose for the two subgroups under the BSOI design, the separate design approach, and the two conventional designs, which ignore the immune response  $Y_I$  or the subgroup  $Z$ , for scenarios 1 to 8. The numbers in boldface are target doses.

	<b>Z=0</b>					<b>Z=1</b>				
Scenario 1										
Sel % (BSOI)	<b>0.726</b>	0.206	0.06	0.004	0	<b>0.816</b>	0.174	0.01	0	0
# of patients	<b>10.4</b>	9.8	7.5	1.9	0.1	<b>12.7</b>	11.4	4.9	1.2	0
Sel % (separate)	<b>0.664</b>	0.216	0.1	0.018	0	<b>0.817</b>	0.154	0.027	0	0
# of patients	<b>12.7</b>	9.1	5.7	2.1	0.4	<b>14.2</b>	10	4.5	1.2	0.1
Sel % (conventional no $Y_I$ )	<b>0.608</b>	0.318	0.07	0	0	<b>0.724</b>	0.262	0.014	0	0
# of patients	<b>10.3</b>	9.5	7.4	2.1	0.2	<b>12</b>	11.2	5.8	1.4	0.1
Sel % (conventional no $Z$ )	<b>0.762</b>	0.182	0.03	0	0	<b>0.762</b>	0.182	0.03	0	0
# of patients	<b>11.5</b>	10.8	5.5	1.1	0.1	<b>11.5</b>	11.2	5.4	1.1	0.1
Scenario 2										
Sel % (BSOI)	0.108	<b>0.636</b>	0.252	0	0	0.19	<b>0.780</b>	0.03	0	0
# of patients	7.6	<b>7.8</b>	7.4	5.2	1.8	8.7	<b>8.5</b>	7.4	4.3	1.2
Sel % (separate)	0.152	<b>0.532</b>	0.306	0.01	0	0.251	<b>0.663</b>	0.086	0	0
# of patients	11	<b>7.8</b>	6	3.6	1.7	11.6	<b>8.4</b>	5.9	3	1.1
Sel % (conventional no $Y_I$ )	0.398	<b>0.548</b>	0.048	0.002	0	0.41	<b>0.542</b>	0.046	0.002	0
# of patients	8	<b>7.9</b>	7	4.8	2	8	<b>7.8</b>	6.8	5.2	2.2
Sel % (conventional no $Z$ )	0.112	<b>0.772</b>	0.116	0	0	0.112	<b>0.772</b>	0.116	0	0
# of patients	8.1	<b>8.2</b>	7.5	4.7	1.5	8.3	<b>7.9</b>	7.4	4.9	1.5
Scenario 3										
Sel % (BSOI)	0	0.002	<b>0.674</b>	0.298	0.022	0	0.002	<b>0.774</b>	0.208	0.016
# of patients	3.7	6.2	<b>8.2</b>	7.2	4.4	4.5	6.5	<b>8.0</b>	6.9	4.2
Sel % (separate)	0.004	0.014	<b>0.556</b>	0.346	0.08	0.002	0.015	<b>0.537</b>	0.377	0.069
# of patients	8.5	7	<b>6.6</b>	4.7	3.1	8.6	7.2	<b>6.5</b>	4.8	3
Sel % (conventional no $Y_I$ )	0	0	<b>0.286</b>	0.626	0.086	0	0	<b>0.344</b>	0.58	0.076
# of patients	4.3	6.2	<b>7.6</b>	7.4	4.4	4.8	6.6	<b>7.4</b>	7.1	4.2
Sel % (conventional no $Z$ )	0	0.002	<b>0.756</b>	0.232	0.01	0	0.002	<b>0.756</b>	0.232	0.01
# of patients	3.9	6.4	<b>8.3</b>	7.1	4.3	4.2	6.5	<b>7.9</b>	7.2	4.2
Scenario 4										
Sel % (BSOI)	0	0.004	0.17	<b>0.616</b>	0.202	0	0.046	<b>0.686</b>	0.22	0.048
# of patients	4.1	5.8	7.9	<b>7.0</b>	4.8	5.3	7	<b>7.8</b>	6.1	3.8
Sel % (separate)	0	0.02	0.236	<b>0.470</b>	0.272	0	0.04	<b>0.556</b>	0.303	0.101
# of patients	8.9	7	6.4	<b>4.6</b>	3.1	9	7.2	<b>6.5</b>	4.3	2.9
Sel % (conventional no $Y_I$ )	0	0	0.182	<b>0.548</b>	0.27	0	0.006	<b>0.276</b>	0.526	0.192
# of patients	4.1	6.3	7.5	<b>7.3</b>	5	5.1	6.6	<b>7.3</b>	6.7	4.2
Sel % (conventional no $Z$ )	0	0.002	0.384	<b>0.496</b>	0.11	0	0.002	<b>0.384</b>	0.496	0.11
# of patients	4.4	6.3	7.6	<b>6.8</b>	4.7	4.5	6.2	<b>7.5</b>	7.1	4.6
Scenario 5										



	Z=0					Z=1				
Sel % (BSOI)	0	0	0.028	0.272	<b>0.684</b>	0	0.006	0.148	<b>0.672</b>	0.174
# of patients	5.6	5.9	6.3	6.4	<b>5.1</b>	6	6.3	6.7	<b>6.2</b>	4.8
Sel % (separate)	0	0.01	0.112	0.314	<b>0.542</b>	0.004	0.023	0.187	<b>0.472</b>	0.305
# of patients	9.9	7	5.4	4.4	<b>3.0</b>	10	7.1	5.6	<b>4.4</b>	2.9
Sel % (conventional no Y)	0.082	0.026	0.22	0.208	<b>0.452</b>	0.096	0.042	0.238	<b>0.204</b>	0.42
# of patients	5.6	6.1	6.4	6.5	<b>5.2</b>	5.9	6.1	6.5	<b>6.4</b>	5
Sel % (conventional no Z)	0	0.002	0.042	0.51	<b>0.402</b>	0	0.002	0.042	<b>0.51</b>	0.402
# of patients	5.6	5.9	6.2	6.2	<b>4.9</b>	5.5	5.7	6.1	<b>6.2</b>	5
Scenario 6										
Sel % (BSOI)	0.002	0.114	<b>0.858</b>	0.016	0.004	0	<b>0.712</b>	0.288	0	0
# of patients	5.7	7.3	<b>7.9</b>	5.5	3.3	7	<b>8.6</b>	7.6	4.7	2.2
Sel % (separate)	0.002	0.164	<b>0.690</b>	0.108	0.036	0.004	<b>0.589</b>	0.398	0.004	0.002
# of patients	9.6	7.8	<b>6.3</b>	4	2.3	10.5	<b>8.5</b>	6.2	3.1	1.6
Sel % (conventional no Y)	0.008	0.16	<b>0.688</b>	0.13	0.014	0.012	<b>0.21</b>	0.708	0.066	0.004
# of patients	6.2	7.1	<b>7.6</b>	6	3.2	6.5	<b>7.5</b>	7.6	5.7	2.7
Sel % (conventional no Z)	0	0.364	<b>0.604</b>	0	0.002	0	<b>0.364</b>	0.604	0	0.002
# of patients	6.1	7.8	7.3	4.9	2.9	6.2	<b>7.7</b>	7.4	5	2.9
Scenario 7										
Sel % (BSOI)	0.006	0	0.018	0.202	<b>0.758</b>	0.004	0.038	<b>0.712</b>	0.218	0.028
# of patients	5.1	5.6	6.5	6.9	<b>5.5</b>	6.5	6.9	<b>7.5</b>	6.1	3.2
Sel % (separate)	0.022	0.034	0.122	0.228	<b>0.580</b>	0.021	0.08	<b>0.575</b>	0.267	0.057
# of patients	9.7	6.8	5.5	4.5	<b>3.3</b>	10.2	7.5	<b>6.1</b>	4.1	2.2
Sel % (conventional no Y)	0.012	0.018	0.244	0.328	<b>0.394</b>	0.022	0.046	<b>0.418</b>	0.346	0.168
# of patients	5.1	6.1	6.9	6.5	<b>5.2</b>	5.9	6.7	<b>7</b>	6.2	4.2
Sel % (conventional no Z)	0	0	0.062	0.564	<b>0.374</b>	0	0	<b>0.062</b>	0.564	0.374
# of patients	5.6	6	6.7	6.7	<b>5.1</b>	5.6	6	<b>6.6</b>	6.8	5
Scenario 8										
Sel % (BSOI)	0.11	0.064	0.17	<b>0.374</b>	<b>0.264</b>	0.208	<b>0.316</b>	<b>0.404</b>	0.05	0.022
# of patients	6.2	6.4	6.4	<b>6</b>	<b>4.4</b>	6.8	<b>6.7</b>	<b>6.7</b>	5.9	3.8
Sel % (separate)	0.108	0.106	0.186	<b>0.224</b>	<b>0.364</b>	0.24	<b>0.286</b>	<b>0.331</b>	0.093	0.051
# of patients	10.1	7.2	5.6	<b>4.1</b>	<b>2.7</b>	10.5	<b>7.3</b>	<b>5.7</b>	4.2	2.3
Sel % (conventional no Y)	0.274	0.192	0.282	<b>0.094</b>	<b>0.154</b>	0.264	<b>0.198</b>	<b>0.284</b>	0.088	0.166
# of patients	6.5	6.7	6.6	<b>5.9</b>	<b>4.3</b>	6.6	<b>6.4</b>	<b>6.5</b>	6.1	4.3
Sel % (conventional no Z)	0.08	0.1	0.384	<b>0.344</b>	<b>0.074</b>	0.08	<b>0.1</b>	<b>0.384</b>	0.344	0.074
# of patients	6.1	6.2	6.6	<b>6</b>	<b>4.6</b>	6.2	<b>6.1</b>	<b>6.5</b>	6	4.5

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 4:**

Selection percentage and the average number of patients treated at each dose for the two subgroups under the BSOI design under a less informative prior distribution for scenarios 1–8.

		<b>Z=0</b>				<b>Z=1</b>				
Scenario 1										
Sel %	<b>0.682</b>	0.226	0.08	0.008	0	<b>0.788</b>	0.192	0.02	0	0
# of patients	<b>7.7</b>	7.3	5.4	1.8	0.1	<b>8.7</b>	8.3	4.3	1.1	0.1
Scenario 2										
Sel %	0.172	<b>0.508</b>	0.306	0.004	0	0.258	<b>0.678</b>	0.062	0.002	0
# of patients	5.6	<b>5.9</b>	5.3	4	1.7	6.2	<b>6.2</b>	5.4	3.4	1.1
Scenario 3										
Sel %	0	0.008	<b>0.608</b>	0.334	0.046	0	0.014	<b>0.676</b>	0.284	0.026
# of patients	3.3	4.8	<b>5.8</b>	5.3	3.2	3.7	4.9	<b>5.7</b>	5.2	3
Scenario 4										
Sel %	0	0.004	0.216	<b>0.6</b>	0.164	0	0.072	<b>0.688</b>	0.21	0.03
# of patients	3.6	4.7	5.6	<b>5</b>	3.2	4.3	5.2	<b>5.9</b>	4.5	2.6
Scenario 5										
Sel %	0	0.004	0.034	0.344	<b>0.582</b>	0	0.008	0.244	<b>0.532</b>	0.216
# of patients	4.4	4.6	4.8	4.5	<b>3.4</b>	4.6	5	5.1	<b>4.9</b>	3
Scenario 6										
Sel %	0	0.158	<b>0.814</b>	0.022	0.006	0.004	<b>0.702</b>	0.294	0	0
# of patients	4.8	5.6	<b>5.8</b>	4	2.3	5.3	<b>6.5</b>	5.6	3.5	1.6
Scenario 7										
Sel %	0.022	0.006	0.04	0.288	<b>0.638</b>	0.018	0.076	<b>0.658</b>	0.216	0.032
# of patients	4.3	4.6	4.9	4.9	<b>3.6</b>	4.8	5.2	<b>5.5</b>	4.6	2.4
Scenario 8										
Sel %	0.138	0.068	0.192	<b>0.342</b>	<b>0.238</b>	0.23	<b>0.3</b>	<b>0.352</b>	0.094	0.022
# of patients	4.7	4.8	4.7	<b>4.7</b>	<b>3.2</b>	5	<b>5.3</b>	<b>5.1</b>	4.4	2.7

**Table 5:**

Selection percentage and the average number of patients treated at each dose for the two subgroups under the BSOI design with an alternative data-generating process for scenarios 1 to 8.

		<b>Z=0</b>				<b>Z=1</b>				
Scenario 1										
Sel %	<b>0.642</b>	0.274	0.074	0.002	0.002	<b>0.82</b>	0.174	0.006	0	0
# of patients	<b>9.8</b>	9.8	7.7	2.1	0.2	<b>12.5</b>	11.6	4.9	1.1	0.1
Scenario 2										
Sel %	0.024	<b>0.76</b>	0.208	0	0	0.072	<b>0.92</b>	0.008	0	0
# of patients	7.7	<b>8</b>	7.7	4.9	1.5	8.6	<b>8.8</b>	7.7	3.8	1
Scenario 3										
Sel %	0	0.002	<b>0.766</b>	0.182	0.05	0	0.004	<b>0.842</b>	0.122	0.032
# of patients	3.6	6.2	<b>8.1</b>	7.2	4.7	4.4	6.6	<b>7.9</b>	6.8	4.4
Scenario 4										
Sel %	0	0	0.15	<b>0.728</b>	0.118	0	0.008	<b>0.904</b>	0.08	0.008
# of patients	3.5	5.6	7.3	<b>7.7</b>	5.5	4.6	6.7	<b>7.8</b>	6.8	4.3
Scenario 5										
Sel %	0	0	0.024	0.266	<b>0.684</b>	0	0.002	0.164	<b>0.652</b>	0.182
# of patients	5.3	5.9	6.4	6.3	<b>5.1</b>	6	6.4	6.7	<b>6.4</b>	4.7
Scenario 6										
Sel %	0	0.258	<b>0.738</b>	0.002	0	0	<b>0.956</b>	0.044	0	0
# of patients	5.6	7.5	<b>8.2</b>	5.5	3	7.1	<b>9.5</b>	7.7	4	1.8
Scenario 7										
Sel %	0	0	0.002	0.032	<b>0.964</b>	0	0.002	<b>0.67</b>	0.326	0.002
# of patients	4	5.1	6.7	7.7	<b>6.6</b>	5.4	6.2	<b>7.2</b>	6.8	3.9
Scenario 8										
Sel %	0.022	0.03	0.2	<b>0.43</b>	<b>0.31</b>	0.098	<b>0.374</b>	<b>0.484</b>	0.04	0.004
# of patients	6	6.2	6.5	<b>6.5</b>	<b>4.8</b>	6.6	<b>6.5</b>	<b>6.7</b>	6	3.8

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 6:**

Selection percentage and the average number of patients treated at each dose for the two subgroups under the BSOI design when the maximum sample size is 45 for scenarios 1 to 8.

		<b>Z=0</b>				<b>Z=1</b>				
Scenario 1										
Sel %	<b>0.698</b>	0.236	0.052	0.002	0	<b>0.806</b>	0.182	0.012	0	0
# of patients	<b>10.9</b>	9.8	6.8	1.8	0.2	<b>12.9</b>	11.5	4.5	1.1	0
Scenario 2										
Sel %	0.068	<b>0.648</b>	0.28	0	0	0.112	<b>0.854</b>	0.034	0	0
# of patients	7.6	<b>7.9</b>	7.7	4.9	1.6	8.7	<b>8.8</b>	7.7	3.9	1
Scenario 3										
Sel %	0	0.008	<b>0.71</b>	0.244	0.028	0	0.004	<b>0.786</b>	0.182	0.026
# of patients	3.4	6.4	<b>8.6</b>	7.4	4	4.1	6.7	<b>8.2</b>	6.9	4
Scenario 4										
Sel %	0	0.002	0.226	<b>0.582</b>	0.154	0	0.034	<b>0.758</b>	0.162	0.038
# of patients	3.6	5.5	7.8	<b>7.3</b>	4.8	4.7	7.2	<b>7.9</b>	6.2	3.8
Scenario 5										
Sel %	0.002	0.008	0.028	0.352	<b>0.562</b>	0	0.01	0.194	<b>0.686</b>	0.106
# of patients	4.8	5.5	6.4	6.3	<b>4.9</b>	6	6.3	6.7	<b>6.3</b>	4.8
Scenario 6										
Sel %	0.002	0.106	<b>0.856</b>	0.014	0.014	0.002	<b>0.738</b>	0.258	0	0.002
# of patients	5.3	7.4	<b>7.8</b>	5.7	3.4	6.8	<b>9</b>	7.5	4.5	2.3
Scenario 7										
Sel %	0.012	0.004	0.022	0.206	<b>0.746</b>	0.01	0.028	<b>0.768</b>	0.176	0.018
# of patients	4.7	5.4	6.7	6.9	<b>6</b>	6.3	7	<b>7.4</b>	6.2	3.1
Scenario 8										
Sel %	0.102	0.076	0.182	<b>0.384</b>	<b>0.23</b>	0.176	<b>0.376</b>	<b>0.392</b>	0.038	0.018
# of patients	5.9	6.3	6.6	<b>6.1</b>	<b>4.4</b>	6.8	<b>6.8</b>	<b>6.6</b>	5.8	3.8