**TSNP: A Two-Stage Nonparametric Phase I/II Clinical Trial Design for Immunotherapy**

1. *Department of Biostatistics, Indiana University, Indianapolis, Indiana*
2. *Center for Computational Biology and Bioinformatics, Indiana University, Indianapolis, Indiana*
3. *Department of Medical and Molecular Genetics, Indiana University, Indianapolis, Indiana*

*Corresponding Author: Yong Zang, Department of Biostatistics, Indiana University, 410 W 10th Street, Indianapolis, IN, 46202. Email:* [*zangy@iu.edu*](mailto:zangy@iu.edu)

***Abstract***

We develop a transparent and efficient two-stage nonparametric (TSNP) phase I/II clinical trial design to identify the optimal biological dose (OBD) of immunotherapy. We propose a nonparametric approach to derive the closed-form estimates of the joint toxicity-efficacy response probabilities under the monotonic increasing constraint for the toxicity outcomes. These estimates are then used to measure the immunotherapy's toxicity-efficacy profiles at each dose and guide the dose-finding. The first stage of the design aims to explore the toxicity profile. The second stage aims to find the OBD, which can achieve the optimal therapeutic effect by considering both the toxicity and efficacy outcomes through a utility function. The closed-form estimates and concise dose-finding algorithm make the TSNP design appealing in practice. The simulation results show that the TSNP design yields superior operating characteristics than the existing Bayesian parametric designs. User-friendly computational software is freely available to facilitate the application of the proposed design to real trials. We provide comprehensive illustrations and examples about implementing the proposed design with associated software.

***1. Introduction***

Thanks to improved understanding of cancer immunology and rapid developments in biomedicine, we have entered the era of immunotherapy in clinical oncology (1). Immunotherapy boosts the body's natural defenses to fight cancer and represents the most promising new cancer treatment approach since the development of the first chemotherapies in the late 1940s (2-4). There are several types of immunotherapy, including monoclonal antibodies, non-specific immunotherapies, oncolytic virus therapy, T-cell therapy, and cancer vaccines. Immunotherapy has already led to major treatment breakthroughs for several cancers, including brain cancer, breast cancer, bladder cancer, and other types of cancer (5).

The emergence of immunotherapy has challenged the “more is better” paradigm generally used by cytotoxic agents. The cytotoxic agents, such as chemotherapy and radiotherapy, assume that both efficacy and toxicity response rates increase monotonically with the dose. Therefore, the goal of a conventional phase I cancer clinical trial for a cytotoxic agent is to identify the maximum tolerated dose (MTD). However, immunotherapy uses different functional mechanisms to treat cancer. Specifically, immunotherapy enhances the immune system's innate power to stop the growth of the cancer cell and prevent cancer from spreading to other parts of the body. Consequently, to achieve an optimal therapeutic effect, immunotherapy is not necessarily administered at its MTD, making the conventional phase I trial design targeting MTD inappropriate for immunotherapy trials. To find the optimal biological dose (OBD) for immunotherapy achieving the optimal therapeutic effect requires monitoring both the toxicity and efficacy outcomes in one trial, which is typically rereferred to as the phase I/II clinical trial (6).

The topic of phase I/II clinical trial design has been intensively studied in the literature. Braun (7) proposed the bivariate continual reassessment method to jointly model the toxicity and disease progression outcomes. Thall and Cook (8) developed a Bayesian design based on the toxicity-efficacy tradeoffs contour. Yin et al. (9) proposed a Bayesian design based on the odds ratio of toxicity and efficacy probabilities. Zhang et al. (10) described a trinomial continual reassessment method. Zang et al. (11) proposed a double-sided isotonic regression method, and the idea was further extended by Zang and Lee to develop a two-stage design (12). Liu and Johnson (13) proposed a flexible Bayesian dynamic model to account for toxicity and efficacy outcomes simultaneously. Liu et al. (14) proposed a Bayesian dose-finding design for immunotherapy by simultaneously considering toxicity, efficacy, and immune response. Recently, Zhou et al. (15) proposed a Bayesian model-assistant phase I/II trial design named U-BOIN to find the OBD based on utility values. See Yuan et al. (6) for a comprehensive discussion of phase I/II designs. Despite the rich body of literature on phase I/II trial designs, these designs have been limited in practice. One possible reason is that most of the existing designs use sophisticated Bayesian probability model, which is non-transparent in the clinical community. Moreover, the performance of these designs heavily relies on the underlying model assumption, such as the dose-response curve and the correlation between the toxicity and efficacy outcomes. Unfortunately, these model assumptions are difficult to justify in practice, especially considering the limited sample size for an early phase clinical trial. Therefore, it is hazardous to use these parametric model-based phase I/II designs in practice. Last but not least, the lack of user-friendly software also prevents the widespread use of these designs.

In this paper, we propose a simple yet well-performing **t**wo-**s**tage **n**on**p**arametric (TSNP) phase I/II clinical trial design to identify the OBD for immunotherapy. In contrast to most existing Bayesian designs based on the parametric model, the TSNP design is purely nonparametric. We do not make any parametric assumptions on the dose-response curve and efficacy-toxicity correlation. Hence, the proposed design is robust because it performs favorably regardless of the underlying dose-response curve. Moreover, the closed-form estimates of the joint efficacy-toxicity response rates and the concise dose-finding algorithm developed in this paper make the proposed design transparent and easy to use in the clinical community. We compared the proposed TSNP design with two parametric designs developed by Thall et al. (8) and Yin et al. (9) and the Bayesian model-assistant U-BOIN design (15) through comprehensive simulation studies. The results show that the TSNP design outperforms the parametric designs and has comparable performance with the U-BOIN design. Also, to facilitate the widespread use of the proposed design, we provide user-friendly software for simulation and trial implementation with the detailed step-by-step illustration.

***2. Method***

*Probability Model and Nonparametric Estimate*

Let denote a set of J pre-specified increasing doses for the immunotherapy under investigation. Let denote the dose-limiting toxicity (DLT) endpoint and denote the tumor response (TR) efficacy endpoints. Then, at any dose level , we define as the DLT rate and as the response rate for the joint response outcome . For toxicity, it is generally reasonable to assume a monotonically increasing relationship between and as:

. (1)

Let be the highest acceptable toxicity rate; the maximum tolerated dose (MTD) is typically defined as the dose with a DLT rate that is closest to . We propose to use a utility function to measure the overall desirability of each dose, which accounts for the risk-benefit tradeoff based on the joint toxicity-efficacy response rate . Specifically, let us define as the desirability score for the response outcome , the utility function for dose is defined as:

. (2)

The desirability scores are elicited from physicians as follows: we first fix the score for the most desirable outcome as and the score for the least desirable outcome as . Then, we ask the physicians to elicit and using the scale (0, 100). Once the desirability scores are elicited, the optimal biological dose (OBD) is defined as the dose maximizing the utility function (2) with acceptable DLT and TR rate. This definition has also been used in several phase I/II trial designs (16-20).

The utility function involves unknown parameter , which is the response rate for the joint response outcome at dose level , and can be estimated from the accumulated data. Let denote the number of patients treated at dose and denote the number of the patients yielding response outcome at dose (k,l=0,1). The likelihood function for the joint toxicity-efficacy outcomes can be written as:

. (3)

Let be the maximum likelihood estimate (MLE) of . They can then be estimated by maximizing the likelihood function (3) subject to the order constraint (1).

We propose a two-step algorithm to derive . We first use the pool-adjacent-violators algorithm (PAVA) to derive the isotonic estimates satisfying the order constraint (1). Let denote the proportional estimate of . If satisfy the order constraint (1), then . Otherwise, for any dose such that , the isotonic estimates are the weighted average . We repeat the above process until the monotonically increasing restriction is satisfied by all the for . As pointed out by Ivanova and Flournoy (21), the isotonic estimates are equivalent to the MLE of under the order constraint (1). We can then derive the MLE by maximizing the likelihood function (3) under the condition . After some algebra, the closed form of can be explicitly written as:

. (4)

Hence, at any interim analysis stage in a clinical trial, after plugging in the MLE , we get as the utility value at dose , which can be used to guide the dose-finding process.

*Two-stage Dose-finding design*

We develop a **t**wo-**s**tage **n**on**p**arametric (TSNP) design for phase I/II clinical trials. The first stage of the design aims to explore the toxicity profile, and the second stage of the design aims to identify the OBD. We propose to use the cumulative cohort design (CCD) in the first stage. The CCD uses the isotonic estimates to guide the dose-finding procedure. In each interim analysis, the decision to increase, decrease, or retain the current dose level depends on a comparison between and the highest acceptable DLT rate and a pre-specified cut-off value τ. Ivanova et al. (22) provided a comprehensive discussion on selecting the cut-off value τ and a list of recommended τ based on different φ. Besides, we also construct an admissible set to safeguard patients from overly toxic doses. Specifically, we test the following hypothesis

for each dose that has been tried. Let be the p-value testing this hypothesis based on Fisher’s exact test and the nominal level to claim overly toxic. Then, during any interim analysis stage, if for any dose level we observe , dose level and higher are excluded from the admissible set, and no more patients will be allocated to these doses. The value of can be determined by evaluating the operating characteristics of the design through simulation studies.

As depicted in Figure 1, the first stage of the dose-finding design can be summarized as follows.

* 1. Treat the first cohort of patients at the lowest dose or any physician-specified dose.
  2. Update the admissible set as . If the admissible set is empty, terminate the trial early, and recommend no dose. Otherwise, restrict the dose-finding procedure within the admissible set.
  3. At the current dose level , estimate based on the toxicity outcomes. To assign a dose to the next cohort of patients,

a. if , escalate the dose level to ;

b. if , de-escalate the dose level to ;

c. otherwise, retain the current dose level .

1.4. Repeat until patients have been treated at the current dose.



***Figure 1. Flowchart of the first stage of the TSNP design.***

The original CCD does not contain a stopping rule. However, we need to save patients for the second stage of OBD finding. For this purpose, in step 1.4 we add a stopping rule to end the first stage and use the remaining patients for the second stage. We are aware that there are various early stopping rules available, targeting for different quantities. Based on our simulation study, the proposed simple rule yields satisfactory operating characteristics.

The second stage of the TSNP design aims to find the OBD, which yields the highest utility value for all the doses within the admissible set. We will update both the admissible set and OBD during the second stage using all the accumulated data. Besides that, if the highest dose has been tired is safe during any interim analysis, the true OBD may be located above all the tried doses, so we expand the admissible set by escalating the dose level.



***Figure 2. Flowchart of the second stage of the TSNP design.***

As depicted in Fig. 2, the second stage of the dose-finding design can be summarized as follows.

2.1 Update the admissible set using accumulated data.

2.2 If the admissible set is empty, terminate the trial early, and recommend no dose.

2.2 Otherwise, let be the highest dose that has been tried. To assign a dose to the next cohort of patients,

a. if j\* is safe, escalate the dose level to ;

b. if is overly toxic, calculate for all the doses within the admissible set and assign the next cohort of patients to the identified OBD.

2.3 Repeat until the maximum sample size of the trial has been reached.

At the end of the trial, using all the data, we construct the final admissible set and calculate the utility values for all the doses within the admissible set. The dose that yields the highest utility value within the admissible set is recommended as the final OBD.

In addition to overly toxic doses, in stage II we may want to add a futility rule to exclude less efficacious doses. For this purpose, let be the TR rate, and λ be the lowest acceptable TR rate, we test the following hypothesis

.

Let be the p-value testing this hypothesis based on Fisher’s exact test and the nominal level to claim futility. Then, during any interim analysis stage, any dose level with should be excluded from the admissible set.

***3. Results***

*Simulation studies*

We conducted simulation studies to evaluate the operating characteristics of the proposed TSNP design. We considered a phase I/II trial with 5 doses and a maximum sample size of 60 in cohorts of size 3. We investigated a commonly used highest acceptable DLT rate of φ=0.3 and specified the cut-off following Ivanova et al. (22). We propose to end the first stage when m=12 patients have been treated at the current dose. Based on our experience, letting m be 20% of the maximum sample size yields the most desirable operating characteristics. We elicited the desirability scores as and , suitable for a solid tumor trial (14, 19). We construct the admissible set solely based on the toxicity outcome and specify . We compared the proposed TSNP design with two Bayesian designs proposed by Thall et al. (8) and Yin et al. (9), which can jointly model the toxicity-efficacy outcomes based on the parametric models without assuming the monotonic dose-efficacy relationship. We are aware that these two designs use different definitions of the OBD. Hence, to make a fair comparison, we modified these two designs slightly to have the same definition of the OBD as the one used in this paper. For the data generation, we use the cross-ratio model (23) to simulate the joint toxicity-efficacy outcomes. This model uses a ratio to induce the correlation between the toxicity and efficacy outcomes with a value of indicating a positive correlation. We used for dose levels 1 to 5. This model has also been used by Yin et al. (9). Under the cross-ratio model, once the marginal DLT rate and marginal TR rate are fixed, we can explicitly determine the joint toxicity-efficacy response rate with given . Hence, to generate correlated toxicity-efficacy outcomes, we first fix , , and calculate the joint probability . Then, we draw random samples from a multinomial distribution with probability . We specified nine toxicity and efficacy scenarios and simulated 10,000 trials under each scenario. We used “P1” to denote the parametric design proposed by Thall et al. (8) and “P2” to denote the parametric design proposed by Yin et al. (9).

Table 1 shows the simulation results, including the OBD selection probability, the average number of patients treated at each dose level, the average percentages of patients experienced TR, and DLT under the TSNP, P1, and P2 designs. Under each scenario, the first three rows represented the underlying true DLT rate, true TR rate, and true utility values at each dose level. The following two rows represented the OBD selection percentage and the average number of patients treated at each dose level. All the results were based on 10,000 simulated trials, and a boldface font emphasizes the result for the true OBD.

In scenarios 1 to 3, the dose-efficacy curves are unimodal. In scenario 1, dose 4 is the MTD, and dose 3 is the OBD with the highest utility value and acceptable DLT rate. The TSNP design yielded the highest OBD selection percentage of 64.6%, whereas the P1 and P2 designs yielded 37.4% and 60.0% percentages of OBD selection. Also, the TSNP design assigned more patients to the OBD than the other two designs. All three designs yielded similar average percentages of TR, and the P1 design yielded a higher average percentage of DLT than the other two designs. The results for scenarios 2 and 3 were similar.

Scenarios 4, 5, and 6 simulate the cases in which the TR rate initially increases with the dose level and then plateaus. The TSNP design outperformed P1 and P2 in the OBD selection percentage. For example, in scenario 6, the OBD selection percentage for the TSNP design was 16% higher than that for P1 and 7% higher than that for P2. In terms of patient assignment, the TSNP design also performed best.

Scenarios 7 to 9 examined the performance of the designs when the dose–efficacy curve was monotonic. We can see that the TSNP design outperformed the other two designs with the highest selection percentages of the OBD. In general, the TSNP design outperformed the two parametric designs across all the scenarios investigated in Table 1.

We conducted a series of sensitivity analyses to further evaluate the TSNP design. First of all, in Table 2, we change the highest acceptable DLT rate from 0.3 to 0.2 with . The TSNP design is still the best in terms of the selection percentage of OBD, and the difference can be as large as 41.3% for the P1 design (scenario 3) and 8.1% for the P2 design (scenario 1). Scenario 4 represents the null case where all the doses under investigation are overly toxic. The TSNP has a 96.4% chance to terminate the trial early, which is similar to the P2 design and is better than the P1 design (93.3%).

***Table 1: The operating characteristics of the TSNP, P1, and P2 designs. The highest acceptable DLT rate is . The trial can only early be terminated for toxicity with nominal level . The bold font indicates the OBD with the highest utility value in the admissible set[[1]](#footnote-1).***

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | *Dose Level* | | | | | % of % of | |
| Design |  | 1 | 2 | 3 | 4 | 5 | TR | DLT |
| *Scenario 1.* |  |  |  |  |  |  |  |  |
|  | DLT rate | 0.05 | 0.1 | 0.15 | 0.25 | 0.4 |  |  |
|  | TR rate | 0.15 | 0.2 | 0.4 | 0.3 | 0.2 |  |  |
|  | Utility value | 34.8 | 36.7 | **49.2** | 38.2 | 26.9 |  |  |
| TSNP | Selection (%) | 7.6 | 11.5 | **64.6** | 15.5 | 0.8 |  |  |
|  | Patients | 7.1 | 9.3 | **24.1** | 12.9 | 6.7 | 29.6 | 18.0 |
| P1 | Selection (%) | 0.9 | 1.9 | **37.4** | 39.6 | 20.1 |  |  |
|  | Patients | 4.0 | 5.5 | **17.2** | 19.4 | 13.8 | 28.5 | 22.9 |
| P2 | Selection (%) | 7.8 | 9.8 | **60.0** | 21.2 | 1.2 |  |  |
|  | Patients | 7.0 | 9.1 | **23.6** | 13.6 | 6.7 | 29.3 | 18.1 |
| *Scenario 2.* |  |  |  |  |  |  |  |  |
|  | DLT rate | 0.1 | 0.15 | 0.25 | 0.4 | 0.5 |  |  |
|  | TR rate | 0.2 | 0.4 | 0.3 | 0.2 | 0.1 |  |  |
|  | Utility value | 37.0 | **49.2** | 38.7 | 26.9 | 18.2 |  |  |
| TSNP | Selection (%) | 13.1 | **66.9** | 18.8 | 1.0 | 0.0 |  |  |
|  | Patients | 10.1 | **25.7** | 14.3 | 6.6 | 3.2 | 30.3 | 21.2 |
| P1 | Selection (%) | 2.1 | **12.6** | 52.8 | 26.4 | 5.7 |  |  |
|  | Patients | 5.5 | **10.9** | 22.2 | 14.4 | 6.8 | 26.1 | 28.2 |
| P2 | Selection (%) | 12.1 | **63.3** | 20.4 | 1.0 | 0.2 |  |  |
|  | Patients | 9.7 | **25.8** | 14.2 | 6.8 | 3.3 | 30.2 | 21.4 |
| *Scenario 3.* |  |  |  |  |  |  |  |  |
|  | DLT rate | 0.25 | 0.4 | 0.45 | 0.5 | 0.6 |  |  |
|  | TR rate | 0.2 | 0.4 | 0.3 | 0.2 | 0.1 |  |  |
|  | Utility value | **32.5** | 40.0 | 32.0 | 24.0 | 15.5 |  |  |
| TSNP | Selection (%) | **62.8** | 24.6 | 3.2 | 0.3 | 0.1 |  |  |
|  | Patients | **28.4** | 17.5 | 5.6 | 2.7 | 1.8 | 26.6 | 36.4 |
| P1 | Selection (%) | **41.2** | 29.3 | 15.6 | 5.4 | 0.6 |  |  |
|  | Patients | **21.8** | 15.3 | 9.8 | 6.2 | 3.1 | 26.2 | 39.5 |
| P2 | Selection (%) | **60.5** | 24.8 | 2.5 | 0.5 | 0.0 |  |  |
|  | Patients | **28.9** | 17.9 | 5.0 | 2.5 | 1.8 | 26.9 | 36.5 |

***Table 1: continued***

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | *Dose Level* | | | | | % of % of | |
| Design |  | 1 | 2 | 3 | 4 | 5 | TR | DLT |
| *Scenario 4.* |  |  |  |  |  |  |  |  |
|  | DLT rate | 0.01 | 0.05 | 0.1 | 0.2 | 0.3 |  |  |
|  | TR rate | 0.05 | 0.2 | 0.4 | 0.4 | 0.4 |  |  |
|  | Utility value | 28.5 | 38.3 | **51.1** | 46.9 | 43.1 |  |  |
| TSNP | Selection (%) | 0.6 | 6.7 | **49.3** | 29.7 | 13.7 |  |  |
|  | Patients | 3.6 | 7.1 | **19.8** | 16.4 | 13.1 | 35.5 | 16.0 |
| P1 | Selection (%) | 0.4 | 0.4 | **28.1** | 40.8 | 30.3 |  |  |
|  | Patients | 3.3 | 4.0 | **14.1** | 19.5 | 19.1 | 37.0 | 19.2 |
| P2 | Selection (%) | 0.3 | 6.0 | **44.2** | 34.2 | 15.3 |  |  |
|  | Patients | 3.4 | 6.4 | **18.1** | 17.7 | 14.4 | 36.0 | 16.4 |
| *Scenario 5.* |  |  |  |  |  |  |  |  |
|  | DLT rate | 0.2 | 0.25 | 0.4 | 0.5 | 0.6 |  |  |
|  | TR rate | 0.05 | 0.3 | 0.4 | 0.4 | 0.4 |  |  |
|  | Utility value | 23.5 | **38.7** | 40.0 | 35.9 | 32.5 |  |  |
| TSNP | Selection (%) | 14.3 | **58.6** | 20.7 | 2.8 | 0.4 |  |  |
|  | Patients | 11.9 | **24.2** | 14.0 | 5.4 | 2.8 | 28.0 | 32.8 |
| P1 | Selection (%) | 9.7 | **43.2** | 34.7 | 8.5 | 0.2 |  |  |
|  | Patients | 10.4 | **17.9** | 17.1 | 8.7 | 4.0 | 29.6 | 35.8 |
| P2 | Selection (%) | 16.9 | **51.2** | 24.4 | 2.8 | 0.3 |  |  |
|  | Patients | 12.3 | **22.8** | 14.8 | 5.7 | 2.9 | 28.4 | 33.0 |
| *Scenario 6.* |  |  |  |  |  |  |  |  |
|  | DLT rate | 0.03 | 0.05 | 0.15 | 0.25 | 0.4 |  |  |
|  | TR rate | 0.05 | 0.1 | 0.2 | 0.4 | 0.4 |  |  |
|  | Utility value | 28.0 | 31.0 | 35.1 | **45.0** | 39.4 |  |  |
| TSNP | Selection (%) | 3.5 | 5.8 | 18.4 | **60.3** | 11.9 |  |  |
|  | Patients | 4.7 | 6.5 | 12.6 | **25.1** | 11.0 | 29.7 | 21.7 |
| P1 | Selection (%) | 2.8 | 2.9 | 32.4 | **44.3** | 17.6 |  |  |
|  | Patients | 4.2 | 4.8 | 16.6 | **21.0** | 13.4 | 29.9 | 22.6 |
| P2 | Selection (%) | 3.8 | 5.4 | 19.4 | **53.3** | 18.1 |  |  |
|  | Patients | 4.8 | 6.3 | 12.6 | **24.3** | 12.0 | 30.2 | 22.1 |
| *Scenario 7.* |  |  |  |  |  |  |  |  |
|  | DLT rate | 0.01 | 0.05 | 0.1 | 0.3 | 0.4 |  |  |
|  | TR rate | 0.1 | 0.2 | 0.4 | 0.45 | 0.5 |  |  |
|  | Utility value | 32.2 | 38.3 | **51.1** | 46.5 | 45.9 |  |  |
| TSNP | Selection (%) | 2.1 | 7.2 | **53.7** | 27.2 | 9.8 |  |  |
|  | Patients | 4.3 | 6.9 | **21.8** | 17.2 | 9.8 | 38.6 | 19.4 |
| P1 | Selection (%) | 1.0 | 1.3 | **41.3** | 38.3 | 18.1 |  |  |
|  | Patients | 3.4 | 4.4 | **18.9** | 19.9 | 13.4 | 41.3 | 22.4 |
| P2 | Selection (%) | 2.1 | 7.1 | **49.0** | 30.6 | 11.2 |  |  |
|  | Patients | 4.3 | 6.5 | **20.8** | 18.2 | 10.2 | 38.8 | 19.9 |

***Table 1: continued***

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | *Dose Level* | | | | | % of % of | |
| Design |  | 1 | 2 | 3 | 4 | 5 | TR | DLT |
| *Scenario 8.* |  |  |  |  |  |  |  |  |
|  | DLT rate | 0.1 | 0.2 | 0.4 | 0.45 | 0.5 |  |  |
|  | TR rate | 0.1 | 0.4 | 0.6 | 0.65 | 0.7 |  |  |
|  | Utility value | 29.8 | **47.3** | 53.0 | 53.7 | 54.9 |  |  |
| TSNP | Selection (%) | 7.3 | **57.0** | 24.5 | 8.0 | 2.9 |  |  |
|  | Patients | 7.8 | **24.1** | 15.9 | 7.5 | 4.5 | 46.8 | 29.6 |
| P1 | Selection (%) | 7.5 | **40.8** | 29.9 | 17.2 | 4.3 |  |  |
|  | Patients | 8.0 | **19.0** | 16.3 | 10.1 | 6.5 | 48.9 | 31.5 |
| P2 | Selection (%) | 11.2 | **50.8** | 28.1 | 7.3 | 2.3 |  |  |
|  | Patients | 7.6 | **24.0** | 16.1 | 7.7 | 4.5 | 47.1 | 29.7 |
| *Scenario 9.* |  |  |  |  |  |  |  |  |
|  | DLT rate | 0.01 | 0.05 | 0.1 | 0.15 | 0.25 |  |  |
|  | TR rate | 0.05 | 0.1 | 0.3 | 0.4 | 0.7 |  |  |
|  | Utility value | 28.5 | 31.0 | 43.9 | 48.9 | **65.9** |  |  |
| TSNP | Selection (%) | 0.2 | 0.6 | 5.8 | 12.6 | **80.7** |  |  |
|  | Patients | 3.3 | 4.1 | 7.3 | 10.9 | **34.4** | 52.1 | 18.7 |
| P1 | Selection (%) | 1.3 | 0.4 | 6.6 | 32.2 | **59.5** |  |  |
|  | Patients | 3.6 | 4.0 | 7.9 | 16.1 | **28.4** | 48.6 | 17.5 |
| P2 | Selection (%) | 0.3 | 0.7 | 5.7 | 23.3 | **70.0** |  |  |
|  | Patients | 3.3 | 4.2 | 7.1 | 11.0 | **34.4** | 52.1 | 18.8 |

In Table 3, we compared the TSNP with a recently proposed Bayesian model-assistant design named U-BOIN. The U-BOIN design used the Gumbel copula model to generate correlated toxicity-efficacy outcomes. To make a fair comparison, the data in Table 3 were also generated from the Gumbel copula with the association parameter set as 0.2. Similar to the TSNP design, the U-BOIN design makes no parametric assumption on the dose-response curves and uses the Dirichlet-multinomial distribution to model the joint toxicity-efficacy response outcomes. In Table 3, we allowed the trial to early terminate for either toxicity or futility and specified the nominal levels as for toxicity and for futility. The results show that the TSNP design and U-BOIN have comparable operating characteristics. Specifically, the U-BOIN design is slightly better than the TSNP design in scenarios 1 and 2, and The TSNP design outperforms the U-BOIN design in the other two scenarios. Considering that TSNP is a frequentist design and requires no additional knowledge in Bayesian statistics to understand the quantitative background and conduct the trial, it provides a plausible option for the clinicians to run simple yet well-performing phase I/I clinical trials with basic statistical knowledge on MLE and p-value.

***Table 2: The operating characteristics of the TSNP, P1 and P2 designs. The highest acceptable DLT rate is 0.2. The nominal level for toxicity is .***

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | *Dose Level* | | | | | % of % of | |
| Design |  | 1 | 2 | 3 | 4 | 5 | TR | DLT |
| *Scenario 1.* |  |  |  |  |  |  |  |  |
|  | DLT rate | 0.01 | 0.05 | 0.3 | 0.4 | 0.5 |  |  |
|  | TR rate | 0.1 | 0.3 | 0.4 | 0.5 | 0.6 |  |  |
|  | Utility value | 32.2 | **45.7** | 43.6 | 45.9 | 48.4 |  |  |
| TSNP | Selection (%) | 8.2 | **69.8** | 18.6 | 2.8 | 0.6 |  |  |
|  | Patients | 6.8 | **29.4** | 15.3 | 5.4 | 3.1 | 33.7 | 16.4 |
| P1 | Selection (%) | 5.8 | **56.0** | 31.7 | 6.1 | 0.4 |  |  |
|  | Patients | 5.2 | **24.6** | 18.5 | 7.7 | 4.0 | 36.2 | 20.3 |
| P2 | Selection (%) | 10.0 | **61.7** | 23.6 | 4.1 | 0.6 |  |  |
|  | Patients | 6.6 | **28.4** | 15.7 | 6.1 | 3.2 | 33.8 | 17.0 |
| *Scenario 2.* |  |  |  |  |  |  |  |  |
|  | DLT rate | 0.01 | 0.05 | 0.1 | 0.5 | 0.6 |  |  |
|  | TR rate | 0.05 | 0.1 | 0.4 | 0.2 | 0.1 |  |  |
|  | Utility value | 28.5 | 31.0 | **51.1** | 24.0 | 15.0 |  |  |
| TSNP | Selection (%) | 2.7 | 4.2 | **93.0** | 0.0 | 0.0 |  |  |
|  | Patients | 5.2 | 7.7 | **39.9** | 5.7 | 1.6 | 30.5 | 13.7 |
| P1 | Selection (%) | 0.9 | 10.5 | **87.4** | 1.1 | 0.1 |  |  |
|  | Patients | 4.1 | 9.1 | **38.7** | 6.5 | 1.7 | 29.9 | 14.4 |
| P2 | Selection (%) | 1.9 | 6.1 | **90.8** | 1.2 | 0 |  |  |
|  | Patients | 4.9 | 7.4 | **40.5** | 5.6 | 1.6 | 30.7 | 13.7 |
| *Scenario 3.* |  |  |  |  |  |  |  |  |
|  | DLT rate | 0.01 | 0.05 | 0.2 | 0.5 | 0.6 |  |  |
|  | TR rate | 0.1 | 0.4 | 0.4 | 0.4 | 0.4 |  |  |
|  | Utility value | 32.2 | 53.0 | 47.3 | 35.9 | 32.5 |  |  |
| TSNP | Selection (%) | 2.2 | **67.2** | 30.1 | 0.4 | 0.0 |  |  |
|  | Patients | 5.1 | **28.0** | 20.1 | 4.7 | 2.1 | 37.3 | 15.1 |
| P1 | Selection (%) | 1.9 | **25.9** | 70.8 | 1.4 | 0.0 |  |  |
|  | Patients | 4.2 | **14.8** | 31.8 | 6.5 | 2.7 | 38.1 | 20.3 |
| P2 | Selection (%) | 2.7 | **62.7** | 34.1 | 0.3 | 0.2 |  |  |
|  | Patients | 5.3 | **26.9** | 21.0 | 4.7 | 2.1 | 37.3 | 15.4 |
|  |  |  |  |  |  |  |  |  |
| *Scenario 4.* |  |  |  |  |  |  |  |  |
|  | DLT rate | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 |  |  |
|  | TR rate | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 |  |  |
|  | Utility value | 21.5 | 24.3 | 27.2 | 29.2 | 31.5 |  |  |
| TSNP | Selection (%) | 3.5 | 0.1 | 0.0 | 0.0 | 0.0 |  |  |
|  | Patients | 13.9 | 2.4 | 1.2 | 0.8 | 0.5 | 12.4 | 52.3 |
| P1 | Selection (%) | 6.3 | 0.4 | 0.0 | 0.0 | 0.0 |  |  |
|  | Patients | 14.2 | 3.3 | 1.9 | 1.1 | 0.6 | 13.1 | 52.8 |
| P2 | Selection (%) | 3.9 | 0.0 | 0.0 | 0.0 | 0.0 |  |  |
|  | Patients | 14.1 | 2.4 | 1.3 | 0.8 | 0.5 | 12.2 | 52.8 |

***Table 3: The operating characteristics of the TSNP design and U-BOIN design. The highest acceptable DLT rate is φ=0.3, and the lowest acceptable TR rate is λ=0.2. The nominal level for toxicity and futility is 0.1 and 0.2.***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | *Dose Level* | | | | |
| Design |  | 1 | 2 | 3 | 4 | 5 |
| *Scenario 1.* |  |  |  |  |  |  |
|  | DLT rate | 0.05 | 0.1 | 0.15 | 0.4 | 0.5 |
|  | TR rate | 0.1 | 0.3 | 0.5 | 0.6 | 0.7 |
|  | Utility value | 31.1 | 44.2 | **56.8** | 53.9 | 56.1 |
| TSNP | Selection (%) | 1.1 | 12.2 | **61.7** | 19.7 | 5.0 |
|  | Patients | 4.3 | 8.8 | **25.0** | 15.1 | 6.6 |
| U-BOIN | Selection (%) | 1.1 | 14.4 | **63.7** | 16.7 | 2.8 |
|  | Patients | 5.8 | 12.0 | **27.4** | 11.8 | 2.4 |
| *Scenario 2.* |  |  |  |  |  |  |
|  | DLT rate | 0.05 | 0.1 | 0.4 | 0.5 | 0.6 |
|  | TR rate | 0.3 | 0.65 | 0.65 | 0.65 | 0.65 |
|  | Utility value | 45.9 | **69.6** | 57.1 | 53.0 | 48.9 |
| TSNP | Selection (%) | 2.6 | **84.2** | 11.2 | 1.6 | 0.4 |
|  | Patients | 5.4 | **33.9** | 13.7 | 4.7 | 2.3 |
| U-BOIN | Selection (%) | 2.4 | **87.5** | 9.1 | 0.9 | 0.1 |
|  | Patients | 6.4 | **39.4** | 11.8 | 2.1 | 0.2 |
| *Scenario 3.* |  |  |  |  |  |  |
|  | DLT rate | 0.01 | 0.05 | 0.1 | 0.2 | 0.4 |
|  | TR rate | 0.1 | 0.2 | 0.3 | 0.65 | 0.4 |
|  | Utility value | 32.2 | 38.5 | 44.2 | **65.4** | 40.9 |
| TSNP | Selection (%) | 0.4 | 2.2 | 6.3 | **89.1** | 1.8 |
|  | Patients | 3.5 | 5.1 | 8.0 | **34.6** | 8.7 |
| U-BOIN | Selection (%) | 0.8 | 4.9 | 8.0 | **85.1** | 1.1 |
| *Scenario 4.* | Patients | 4.4 | 7.1 | 10.4 | **32.4** | 5.8 |
|  | DLT rate | 0.05 | 0.1 | 0.5 | 0.6 | 0.7 |
|  | TR rate | 0.01 | 0.05 | 0.1 | 0.4 | 0.5 |
|  | Utility value | 24.5 | 26.1 | 18.7 | 33.9 | 36.1 |
| TSNP | Selection (%) | 2.1 | 1.9 | 4.2 | 1.6 | 0.0 |
|  | Patients | 9.4 | 14.2 | 10.7 | 4.0 | 1.6 |
| U-BOIN | Selection (%) | 2.5 | 2.9 | 6.4 | 0.5 | 0.1 |
|  | Patients | 12.4 | 16.7 | 10.7 | 1.4 | 0.1 |

***Figure 3. The correct OBD selection percentage of the TSNP designs under different sample sizes. All the scenarios are selected from Table 1.***

Figure 3 was dedicated to the sample size justification. In Table 1, we used a sample size of N=60 to run the simulation. The underlying criteria is that we require a sample size achieving at least 50% correct OBD selection percentage across all the scenarios considered. To further investigate this issue, in Figure 3, we depicted the correct OBD selection percentages of scenarios 4, 5, 7, and 8, which represents the lowest four correct OBD selection percentages in Table 1. We considered different sample sizes of 40, 50, and 60 with m=8, 10, and 12, respectively. Let ϕ be the correct OBD selection percentage, we summarized the results as follows:

* Sce. 4: ϕ=44.6% for N=40, ϕ=45.4% for N=50, ϕ=49.3% for N=60;
* Sce. 5: ϕ=49.0% for N=40, ϕ=52.6% for N=50, ϕ=58.6% for N=60;
* Sce. 7: ϕ=50.7% for N=40, ϕ=52.5% for N=50, ϕ=53.7% for N=60;
* Sce. 8: ϕ=48.2% for N=40, ϕ=53.8% for N=50, ϕ=57.0% for N=60.

Based on this sensitivity analysis, we select 60 as the sample size for simulation studies.

*An example trial*

To illustrate the application of the proposed TSNP design, we carry out a hypothetical phase I/II clinical trial with the highest acceptable DLT rate of 0.3, 5 pre-specified doses, a maximum sample size of 30 with m=6. The same desirability scores as those used in Table 1 is used here. The trial can only be early terminated for toxicity with a nominal level Figure 4 shown the trial's process.

** ***Figure 4. A hypothetical phase I/II clinical trial using the proposed TSNP design.***

The trial starts by assigning the first cohort of patients at the lowest dose. None of the patients reports either DLT or TR. We then escalate the dose and treat the next cohort of patients at dose

2. Again, none of the patients reports either DLT or TR. We further escalate to dose 3. Two of the patients assigned to dose 3 experience neither DLT nor TR, and the other one experiences TR only, so we escalate to dose 4. One patient assigned to dose 4 experiences neither DLT nor TR, and the other two report TR only. Then, we escalate to the highest dose 5. For three patients treated with dose 5, two of them report DLT only, and the other one experiences neither DLT nor TR, which yields an isotonic estimate of the DLT rate of 2/3 and is higher than 0.3+0.1=0.4. As a result, we de-escalate to dose 4. For the three patients allocated to dose 4, one of them experiences neither DLT nor TR, but the other two report both DLT and TR. Because there are already 6 patients allocated to dose 4, we switch to the second dose-finding stage after treating six cohorts of patients.

To start stage II and assign the seventh cohort of patients, we first determine the admissible set. The p-values based on Fisher's exact test indicate that all the doses are safe with the lowest p-value of 0.22. Then, we need to calculate the utility values at each dose, and we take dose 4 as an example. Because the proportional estimates of the toxicity rates from dose 1 to 5 using the first six cohorts of patients in stage I follow the monotonically increasing pattern, the isotonic estimate is simply the proportional estimate, and we get . After plugging in formula (3) with we get the estimates of the joint toxicity-efficacy response rates as and . Then, we plug in formula (2) and estimate the utility value at dose 4 as , which is the highest utility values among all the doses. Consequently, we assign the seventh cohort of patients at dose 4. The trial continues, as shown in Figure 4. At the end of the trial, based on all the accumulated data, we find all the doses are safe and dose 4 yields the highest utility value of 46.67. As a result, we recommend dose 4 as the final OBD identified from this trial.

*Software*

To facilitate the widespread of the proposed TSNP design, we have developed easy-to-use R codes. In particular, crossratio () and Gumbel() generate the joint toxicity-efficacy response probabilities for each dose level based on the cross-ratio model and Gumbel copula model. TSNP.OC() generates the operating characteristics of the TSNP design through simulated trials, like those provided in Tables 1-3. TSNP.I() and TSNP.II() provide the dose assignment and OBD selection for practical trial implementation, as shown in Figure 4. We have provided more details in the Appendix. All the codes are freely available at <https://github.com/yongzang2020/TSNP>.

***4. Discussion***

In this article, we have proposed a two-stage nonparametric (TSNP) design to find the optimal biologic dose for immunotherapy in a phase I/II trial. The proposed design is purely nonparametric; the isotonic regression method is used to compute the DLT rate. The closed-form estimates for the joint toxicity-efficacy response rate are derived through a conditional likelihood function. Therefore, no computational intensive “black box” method, such as numerical optimization or MCMC approach is required, making the proposed design simple and transparent to the clinical community. Simulation studies show that compared with existing computational intensive Bayesian parametric designs (P1 and P2 in Table 1 and Table 2), the TSNP design yields the highest correct OBD selection percentages and, in general, assigns more patients to the OBD. When all the doses are overly toxic and/or less efficacious, the TSNP design has about 90% to terminate the trial early. The TSNP design yields similar operating characteristics as the recently proposed Bayesian model-assistant U-BOIN design. However, unlike the U-BOIN design which relies on the prior and posterior probabilities to conduct the trial, the TSNP only needs the MLE and p-value from Fisher's exact test to run the trial. More importantly, user-friendly software for simulation and real trial implementation is freely available for download. The easy-to-use Shiny app, which will provide a graphical user interface, is also under development.

We define the utility function as a weighted average of the joint toxicity-efficacy response rates. We are also aware that the form of the utility function is not unique. In practice, people may define a variety of the optimal biological dose through different utility function structure. However, as long as the utility function can be defined as a function of the joint response rates, the proposed design can be applied with little modification.

The primary interest of this paper is to developing phase I/II trial design for immunotherapy. Like immunotherapy, the dose-efficacy curve for a molecularly targeted agent may not follow the monotonic pattern either. Therefore, The TSNP can also be used for the molecularly targeted agent. Moreover, because the TSNP is a nonparametric design with no model assumption on the dose-response curve, it can also be applied for conventional therapies such as radiotherapy or chemotherapy if the OBD is defined as a toxicity-efficacy tradeoff.

The proposed design is appropriate for trials in which the response outcome is binary and observable shortly after the treatment initiation. It cannot be applied directly to cases in which the response outcome requires a long follow-up time to be assessed. A possible strategy is to treat the delayed efficacy outcome as a missing data problem and use missing data methodology to resolve the problem (24-27). The proposed designs can also be extended to find the optimal dynamic treatment regime that can adapt the same patient dose assignment among different treatment cycles. Lastly, In addition to single-agent trials, immunotherapy is often used together with other agents or other kinds. It is of intense interest to extend the TSNP design to identify the optimal dose combination for phase I/II drug-combination trials. Future research in this direction is warranted.

***Appendix: R codes illustration and examples***

We have developed R codes to implement the TSNP design. The first function crossratio (ttox, teff, gamma) will generate correlated joint toxicity-efficacy response rates on each dose level based on the cross-ratio model. The arguments of cross-ratio () are:

*ttox: marginal toxicity probability ()*

*teff: marginal efficacy probability ()*

*gamma: cross ratio between ttox and teff, gamma>1 indicates a positive correlation ()*

The second function Gumbel(ttox, teff, c) will generate correlated joint toxicity-efficacy response rates on each dose level based on the Gumbel copula model. c is defined as:

c: association parameter between tox and eff, c>0 indicates a positive correlation

Both crossratio () and Gumbel() will create a probability matrix, with rows presenting different responses and columns representing different doses. The probability matrix can be expressed as

.

The third function TSNP.OC(pi,cohortsize,ncohort,tau,m,phi.t,c.t,phi.e,c.e,omega,start,ntrial) will generate operating characteristics of TSNP design through simulated trials. The arguments of TSNP.OC() are:

*pi: the joint toxicity-efficacy probability matrix*

*cohortsize: the size (number of patients) for each cohort*

*ncohort: the number of cohort for a phase I/II trial*

*tau: the meaningful difference of toxicity rate used in the CCD design ()*

*m: if m patients have been treated at the current dose, switch the trial to the second stage*

*phi.t: the maximum tolerable toxicity rate ()*

*c.t: nominal level of toxicity for admissible set ()*

*phi.e: the minimum tolerable efficacy rate (λ)*

*c.e: nominal level of futility for admissible set ()*

*omega: the desirability scores ()*

*start: the starting dose level*

*ntrial: the number of simulated trials*

To obtain the simulation result of scenario 1 in Table 1, we first use crossratio () to specify the joint toxicity-efficacy response rates as

*pi=crossratio(ttox=c(0.05,0.1,0.15,0.25,0.4), teff=c(0.15,0.2,0.4,0.3,0.2), gamma=c(1,2,2,3,3))*

Then, the following R codes will provide the simulation results.

*TSNP.OC(pi=pi,cohortsize=3,ncohort=20,tau=0.1,m=12,phi.t=0.3, c.t=0.1, phi.e=0.4, c.e=0, omega=c(25,100,0,50), start=1,ntrial=10000)*

.

The last two functions TSNP.I(y00,y01,y10,y11,tau,m,phi.t,c.t,current) and TSNP.II(y00,y01,y10,y11,omega,phi.t,c.t,phi.e,c.e,ss) will provide the dose assignment for coming cohort of patients and OBD selection for practical trial implementation. TSNP.I() is for the first stage and TSNP.II() is for the second stage. The new arguments in these two functions are:

*y00: number of patients with tox=0, eff=0 at different doses*

*y01: number of patients with tox=0, eff=1 at different doses*

*y10: number of patients with tox=1, eff=0 at different doses*

*y11: number of patients with tox=1, eff=1 at different doses*

*current: current dose level for the latest cohort of patients in the trial*

*ss: maximum sample size*

Finally, the following R codes will provide the dose assignment as shown in Figure 4

*## cohort 1*

*data=matrix( c(3,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0) ,nrow=4 )*

*TSNP.I(y00=data[1,],y01=data[2,],y10=data[3,],y11=data[4,],tau=0.1,m=6,phi.t=0.3,c.t=0.1,current=1)*

*.*

*.*

*.*

*## cohort 5*

*data=matrix( c(3,0,0,0,3,0,0,0,2,1,0,0,1,2,0,0,1,0,2,0) ,nrow=4 )*

*TSNP.I(y00=data[1,],y01=data[2,],y10=data[3,],y11=data[4,],tau=0.1,m=6,phi.t=0.3,c.t=0.1,current=5)*

*## cohort 6*

*data=matrix( c(3,0,0,0,3,0,0,0,2,1,0,0,2,2,0,2,1,0,2,0) ,nrow=4 )*

*TSNP.II(y00=data[1,],y01=data[2,],y10=data[3,],y11=data[4,],omega=c(25,100,0,50),phi.t=0.3,c.t=0.1,phi.e=0.2,c.e=0,ss=30)*

*.*

*.*

*.*

*## cohort 10*

*data=matrix( c(3,0,0,0,3,0,0,0,3,1,0,2,6,4,2,3,1,0,2,0) ,nrow=4 )*

*TSNP.II(y00=data[1,],y01=data[2,],y10=data[3,],y11=data[4,],omega=c(25,100,0,50),phi.t=0.3,c.t=0.1,phi.e=0.2,c.e=0,ss=30)*

***Acknowledgment***

The authors thank the two reviewers for their helpful comments, which improve the presentation of this paper. The research of Yan Han, Hao Liu, and Yong Zang is partially supported by NIH P30 grant CA082709. The research of Sha Cao is partially supported by Indiana University Precision Health Initiative.

***Data Availability statement***

No real data is used in this study. The R codes to implement the proposed research are freely available at <https://github.com/yongzang2020/TSNP>.

.

***References***

1. Syn NL, Teng MWL, Mok TSK, Soo RA. De-novo and acquired resistance to immune checkpoint targeting. Lancet Oncol. 2017; 18: 731-741.

2. Topalian SL, Weiner GJ, Pardoll, DM. Cancer immunotherapy comes of age. Journal of Clinical Oncology. 2011; 23: 4828-4836.

3. Couzin-Frankel J. Cancer immunotherapy. Science. 2013; 324: 1432-1433.

4. Makkouk A, Weiner GJ. Cancer immunotherapy and breaking immune tolerance: new approaches to an old challenge. Cancer Research. 2015; 75: 5-10.

5. Khalil DN, Smith EL, Brentjens RJ, Wolchok JD. The future of cancer treatment: immunomodulation, CARs and combination immunotherapy. Nature Reviews Clinical Oncology. 2016; 13: 273–290.

6. Yuan Y, Nguyen HQ, Thall PF. Bayesian Designs for Phase I-II Clinical Trials. Chapman & Hall/CRC Biostatistics Series. 2016.

7. Braun, TM. The bivariate continual reassessment method: extending the CRM to phase I

trials of two competing outcomes. Controlled clinical trials 2002; 23: 240-256.

8. Thall PF, Cook J. Dose-finding based on efficacy-toxicity tradeoffs. Biometrics. 2004; 60: 684-693.

9. Yin G, Li Y, Ji Y. Bayesian dose-finding in phase I/II clinical trials using toxicity

and efficacy odds ratios. Biometrics. 2006; 62: 777-784.

10. Zhang W, Sargent DJ, Mandrekar SJ. An adaptive dose-finding design incorporating both toxicity and efficacy. Statistics in Medicine 2006; 25: 2365-2383.

11. Zang Y, Lee JJ, Yuan Y. Adaptive designs for identifying optimal biological dose for molecularly targeted agents. Clinical Trials. 2014; 11: 319-327.

12. Zang Y, Lee JJ. A robust two‐stage design identifying the optimal biological dose for phase I/II clinical trials. Statistics in Medicine 2017; 36: 27-42.

13. Liu S, Johnson, VE. A robust Bayesian dose-finding design for phase I/II clinical trials. Biostatistics, 2016; 17: 249-263.

14. Liu, S., Guo, B. and Yuan, Y. A Bayesian Phase I/II Design for Immunotherapy Trials. Journal of the American Statistical Association, 2018; 113: 1016-1027.

15. Zhou Y, Lee JJ, Yuan Y. A utility-based Bayesian optimal interval (U-BOIN) phase I/II design to identify the optimal biological dose for targeted and immune therapies. Statistics in Medicine 2019; 38: 5299-5316.

16. Houede N, Thall PF, Nguyen H, et al. Utility-based optimization of combination

therapy using ordinal toxicity and efficacy in phase I/II trials. Biometrics, 2010; 66: 532-540.

17. Thall PF, Nguyen HQ, Braun TM, et al. Using joint utilities of the times to response

and toxicity to adaptively optimize schedule-dose regimes. Biometrics, 2013; 69: 673-682.

18. Thall PF, Nguyen HQ, Zohar S Maton P. Optimizing sedative dose in preterm

infants undergoing treatment for respiratory distress syndrome. Journal of the American

Statistical Association, 2014; 109: 931-943.

19. Guo B, Yuan Y. Bayesian phase I/II biomarker-based dose finding for precision

medicine with molecularly targeted agents. Journal of the American Statistical Association, 2017; 112: 508-520.

20. Murray TA, Thall PF, Yuan, Y, McAvoy S Gomez DR. Robust treatment comparison based on utilities of semi-competing risks in non-small-cell lung cancer. Journal of the American Statistical Association, 2017; 112: 11-23.

21. Ivanova A, Flournoy N. Comparison of Isotonic Designs for Dose-Finding. Stat Biopharm Res, 2009; 1:101.

22. Ivanova A, Flournoy N, Chung Y. Cumulative cohort design for dose-finding. Journal of Statistical Planning and Inference 2007; 137: 2316 – 2327.

23. Dale JR. Global cross-ratio models for bivariate, discrete, ordered responses. Biometrics 1986; 42:909–917.

24. Cheung YK, Chappell R. Sequential Designs for Phase I Clinical Trials with Late-onset Toxicities. Biometrics 2000: 1177-1182.

25. Yuan Y, Yin G. Robust EM Continual Reassessment Method in Oncology Dose Finding. Journal of the American Statistical Association 2011; 106: 818-831.

26. Liu S, Yin G, Yuan Y. Bayesian Data Augmentation Dose Finding with Continual Reassessment Method and Delayed Toxicity. Annals of Applied Statistics 2013; 4: 2138-2156.

27. Jin IH, Liu S, Thall P, Yuan Y. Using Data Augmentation to Facilitate Conduct of Phase I/II Clinical Trials with Delayed Outcomes. Journal of the American Statistical Association 2014; 109: 525-536.

1. *The OBD is not necessarily the dose with the highest utility value among all the doses. For*

   *example, in scenario 3, dose 2 gives the highest utility value but also yields an overly toxic toxicity rate of 0.4, which is excluded from the admissible set. Dose 1 is in bold font because it is the OBD with an acceptable toxicity rate.* [↑](#footnote-ref-1)