**Dose Escalation with Over-dose and Under-dose Control**

**Using a quasi-continuous toxicity score in Phase I/II Clinical Trials**

Jieqi Tu1, 2, Zhengjia Chen1, 2

1 Division of Epidemiology and temporary medicine, especially for medical oncology. Therefore, we developed a Biostatistics, School of Public Health, Bayesian adaptive Phase I clinical trial design entitled escalation with overdoing University of Illinois at Chicago, Chicago, Illinois, USA

2 Biostatistics Shared Resource Core, University of Illinois Cancer Institute, Chicago, Illinois, USA

Corresponding author: Dr. Zhengjia Chen.

Address: 1603 West Taylor Street, SPHPI 947, Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, IL 60612

Email: znchen@uic.edu

Phone: 312-413-7059 (SPH), 312-996-6027 (Cancer Center)

Running Title: EWOUC-NETS design

**Abstract**

Escalation with overdose control (EWOC) is a Bayesian adaptive design for selecting dose levels in cancer phase I clinical trials and has been carried out for many years. However, the toxicity response was treated as binary indicator of dose limiting toxicity (DLT) and under-dose control was not considered in this design. Chen et al. (2010) proposed a novel toxicity scoring system to fully utilize patients’ toxicity information using a normalized equivalent toxicity score. Additionally, Chen and his colleagues in 2015 extended EWOC to Phase I/II clinical trials by controlling for under-dosing (EWOUC) to provide at least minimum efficacy of drug. In this paper, the EWOUC-NETS design was developed based on these two methods to combine their advantages [[1](#_ENREF_1)]. Moreover, we further extended this study by treating efficacy outcome as also continuous, and recommended dose (RD) was chosen based primarily on Bayesian method. The dose escalation decision rules were based on the posterior distribution of both toxicity and efficacy outcomes. We compared the operation characteristics of the proposed and existing methods through simulation studies under three scenarios. We found that EWOUC-NETS with continuous efficacy outcome effectively increased the accuracy in identifying the optimized utility dose (OUD) and provided better therapeutic effects.

1. **Introduction**

Anti-cancer drug development is a highly complicated, extremely expensive and time-consuming process. One of the most important steps in drug development is Phase I cancer clinical trials. Phase I trials are conducted to evaluate a new drug’s toxic effect on patients and to find the optimal dose, called the maximum tolerated dose (MTD) or Phase II dose, for subsequent studies. MTD is defined as the dose at which the probability of dose limiting toxicity (DLT) is equal or close to the target toxicity level (TTL, i.e., 33%). The main objective of a cancer Phase I clinical trial is to estimate MTD of a new drug or combination of drugs under safe administration and acceptable level of adverse events using toxicity responses. Cancer phase I trials are conducted sequentially, assigning dose levels to subjects based on the observed side effects from patients treated previously. These trials should be well-designed to avoid unacceptable toxic events and to ensure that patients are treated at an optimal dose as much as possible. Among a variety of existing Phase I clinical trial designs, escalation with over-dose control (EWOC), proposed by Babb, Rogatko and Zacks [[2](#_ENREF_2)], is one of the most popular Bayesian dose-finding method. EWOC directly addresses the ethical need to control the probability of overdosing. Numerous extensions of the EWOC have been proposed to improve the performance and to adapt for more complicated dose-finding problems. For example, most cancer Phase I trials treat DLT as a binary indicator of toxicity. DLT is defined as grade 3 or 4 non-hematologic and grade 4 hematologic toxicity, according to the National Cancer Institute (NCI) common termilogy criteria for adverse events [[3](#_ENREF_3)]. But patients experiencing multiple grade 2 toxicities would be classified the same as patients having one or more grade 1 toxicities. So, Chen, Krailo, Azen, and Tighiouart proposed a novel normalized equivalent toxicity scoring (NETS) system to fully utilize all toxicities of patients and toxicity variable is considered quasi-continuous in Phase I clinical trials [[4](#_ENREF_4)]. After that, Chen, Tiguiouart and Kowalski demonstrated that the integration of NETS with EWOC, called EWOC-NETS, can strongly increase the accuracy of MTD estimation and the trial efficiency [[5](#_ENREF_5)]. Nevertheless, the EWOC design and the aforementioned extensions of EWOC only take toxicity as the primary endpoint. Chen and his colleagues also proposed another extended version of EWOC – dose escalation with overdose and underdose control (EWOUC) in Phase I/II clinical trials. The EWOUC design considers toxicity and drug efficacy as dual primary endpoints to provide patients with at least minimal drug efficacy while controlling the probability of over-dosing.

In the EWOUC design, binary indicators of dose limiting toxicity (DLT) (whether DLT occurs during the observation window of one cycle of therapy) and efficacy event (whether a patient experiences an efficacy event during the observation window of one cycle of therapy) are used to describe the toxicity and efficacy outcomes. However, after incorporating efficacy into a dose-finding design, researchers might further utilize the efficacy data to reduce the loss of information. In some studies, the efficacy endpoint to evaluate antitumor activity is considered to be binary (e.g., response or non-response) on the basis of a threshold for tumor shrinkage as a continuous variable. However, categorization of continuous variables usually results in a considerable loss of information, which to some extent reduces the statistical efficiency. To address this issue, a new dose-finding approach for correlated continuous toxicity and efficacy outcomes in Phase I/II oncology trials is required. An increasing number of dose-finding methods using dual endpoints and continuous efficacy outcome have been proposed to address this issue. For instance, Tao and his colleagues proposed a dose-finding method for bivariate continuous toxicity and efficacy outcome, as well as bivariate mixed outcomes in up-and-down designs in Phase III trials [[6](#_ENREF_6)]. However, the up-and-down method of dose escalation has always shown inferior performance when compared with EWOC designs although it is the most used design in dose-finding cancer trials [[7](#_ENREF_7)]. Hirakawa proposed an adaptive dose-finding approach for treating correlated bivariate binary toxicity and continuous efficacy outcomes in designing Phase I oncology trials [[8](#_ENREF_8)]. In this method, a binary toxicity outcome and a continuous efficacy outcome are modeled by a factorization of the joint distribution of these outcomes using approaches introduced by Olkin and Tate [[9](#_ENREF_9)]. The joint distribution of the outcomes can be factorized into two formulations: (i) a marginal distribution for binary outcomes and a conditional distribution for continuous outcomes given the binary outcome, and (ii) a marginal distribution for continuous outcomes and a conditional distribution for binary outcomes given the continuous outcome. Hirakawa used the first formulation and proposed a dose-finding approach based on the Bayesian method [[8](#_ENREF_8)]. Our methodology is inspired by the studies mentioned above. Incorporating NETS and Hirakawa’s method to treat bivariate correlated model of continuous efficacy endpoint and binary toxicity endpoint, we propose a new Bayesian dose-finding approach in the framework of EWOC-NETS and EWOUC, called EWOUC-NETS, using correlated bivariate continuous toxicity and efficacy outcomes, and typically the toxicity outcome is described by NETS. Our design treats toxicity and efficacy as dual primary endpoints and both continuous.

Our methodology is established based on the EWOUC and NETS framework. The EWOUC design provides over-dose and under-dose control to protect patients from severe toxicity and ensure at least minimum efficacy at the same time. However, it does not incorporate NETS to better utilize all the information of toxicity and still treats efficacy outcome as a binary indicator. Hence, the considerable loss of information might affect the accuracy of identifying the OUD and therefore cannot provide the optimal therapeutic benefits to patients. This challenge motivated us to develop new designs with aims: (1) Use NETS to estimate toxicity outcome; (2) Use continuous efficacy outcome; (3) Use EWOUC trial design framework.

The remainder of the manuscript is organized as follows. In Section 2, we introduce the definition of NETS and describe the design EWOUC-NETS with probability models. In Section 3, extensive simulations are presented to evaluate the performance of the EWOUC-NETS including its operating characteristics. In Section 4, an application of the EWOUC-NETS to a real trial example is demonstrated. The article ends with discussion and conclusions.

1. **Methods**

In this section, we briefly review a normalized equivalent toxicity score NETS system introduced by Chen et al. [[4](#_ENREF_4)] for cancer Phase I trials design, using all grades of toxicities of patients in the trial. Then, we describe EWOUC-NETS, our new design for dose-finding based on EWOUC, proposed by Chen et al [[1](#_ENREF_1)], adapted to a quasi-continuous toxicity scoring system NETS and a continuous efficacy outcome, using a factorization of the joint distribution of the outcomes.

* 1. ***Introduction of NETS***

Phase I clinical trials usually have relatively small sample sizes and therefore limited information of toxicity and efficacy. Full utilization of all toxicity data is important to improve trial efficiency and accuracy of MTD estimation. Typically, patients in the trials can experience different types and grades of toxicities varying from 0 for no toxicity to 5 for death. According to the severities and types, toxicities can be categorized into five grades as follows:

Grade 0: no toxicity;

Grade 1: mild toxicity;

Grade 2: moderate toxicity;

Grade 3: severe toxicity;

Grade 4: life-threatening toxicity; and

Grade 5: death.

But such information is disregarded when the toxicity is treated as a binary indicator in EWOC and EWOUC clinical trial designs, which results in a considerable loss of valuable information and consequently reduces the statistical efficiency. Hence, here we use a normalized equivalent toxicity score (NETS), to replace the binary indicator and make full use of the information to get more accurate result.

The NETS is defined as below:

(1)

where denotes the adjusted grade of the th toxicity level of the th patient, and is the worst toxicity with the highest value of adjusted grade of all toxicities of the th patient. is defined to be 0 for grade 0 toxicity, 1 for grade 1 toxicity, 2 for grade 2 toxicity, 3 for grade 3 non DLT, 4 for grade 4 non DLT, 5 for grade 3 DLT, and 6 for grade 4 DLT, see Chen et al. [[4](#_ENREF_4)]. The parameter represents the weight for the correlation of the th toxicity with other toxicities for the *i*th patient. Chen and his colleagues discuss the is a slope parameter and values between 0.1 and 0.5. It represents the increasing “speed” of NETS due to additional toxicity besides the worst toxicity of the patient. Obviously, we can find that depends a lot on , the maximum toxicity level of patient *i*. The NETS, has been normalized with range 0 to 1 since the maximum possible value. for th patient is defined to be 0 if patient has no toxicity event and is equal to if patient has only one grade 1 toxicity.

A logistic function can be employed to model the relationship between dose level and the expectation of NETS for the th patient, denoted by , as:

(2)

can be considered as the average NETS (ANETS) at dose level . is a specified distribution function.

* 1. ***Outcome model and re-parameterization***

Let and denote the outcomes of NETS and a continuous efficacy variable. They are both assumed to be truncated normal distributed with range 0 to 1. So, the raw continuous efficacy outcome needs to be transformed by before formulating the model. Hirakawa proposed a novel adaptive dose-finding approach to handle correlated continuous efficacy outcome and binary toxicity outcome [[8](#_ENREF_8)]. The distribution of efficacy outcome in our method is established based on their framework, but fewer parameters need to be estimated.

We assume the marginal distribution of and follow the distributions below:

(3)

(4)

where is assumed to have a logistic relationship with dose levels:

(5)

It is also assumed that:

(6)

where and are unknown parameters. is the multiplicative heteroscedasticity proposed by Harvey [[10](#_ENREF_10)]. The parameter controls the degree of heteroscedasticity. The homoscedasticity is held when .

The correlation between toxicity and efficacy outcomes should also be taken into consideration. We still use NETS, denoted as , as the toxicity outcome. The model we are using is based on the factorization of the joint distribution of ():

(7)

Given , we assume that the distribution of is normal:

(8)

where is the parameter for the regression of on S. Large absolute values of indicate a strong correlation between the two outcomes. When = 0, the two outcomes are independent given the dose level of the agent in the model. The correlation based on this model is

(9)

However, in the above dose-toxicity and dose-efficacy relationships do not have clear clinical meanings. In order to provide clinical interpretation for dose-toxicity relationship, re-parameterization is needed, using parameters that are easy to be interpreted in a clinical way: MTD (), NETS () at the starting dose , target normalized equivalent toxicity score () as below:

(10)

(11)

Then, the two original parameters and can be re-written as follows:

(12)

(13)

Hence, if plugging in (12) and (13) into (2), we can re-write the dose-toxicity relationship as:

(14)

Similarly, the dose-efficacy relationship can also be re-parameterized, using MED (), efficacy outcome () at the starting dose and expected efficacy at MED ():

(15)

(16)

Correspondingly, and can be re-written as follows using (10) and (11):

(17)

(18)

Therefore, the dose-efficacy relationship (5) can be interpreted using parameters that have clear clinical meanings with (17) and (18):

(19)

Suppose that patients have been treated in the clinical trial, resulting in the NETS and continuous efficacy data . The likelihood function is given by (20) as:

(20)

where . Let denote the prior distribution of , then the posterior distribution of is given by (21) as:

(21)

In order to define , and are assigned to have uniform distributions in the interval and , where is a small positive value. Meanwhile, and are assigned to be non-informative uniform prior distributed in the interval , where is the maximum dose under investigation or determined by pre-clinical studies. When <, the trial will be terminated early because of high toxicity. When >, the trial will also be terminated early without a dose level being recommended, or new higher dose levels need to be added because no pre-specified dose levels meet the minimum efficacy requirement. When > and <, the trial will choose the dose with the highest utility.

* 1. ***Over-dose and under-dose control***

As Chen and his colleagues proposed for the EWOUC method [[1](#_ENREF_1)], in order to offer patients maximum benefit, we should prevent letting patients be exposed to overly toxic doses (i.e., doses above , the MTD) or low doses that cannot provide therapeutical effects (i.e., doses below , the MED). Hence, it is ideal that patients are treated at doses between and . Therefore, the trial should be terminated if is greater than and we cannot select any doses for further study as there is no doses that are both safe and efficacious. However, if there is an interval between the MED and MTD (i.e., ), investigational doses could be chosen within this interval to treat patients to ensure safety and minimum efficacy.

Suppose that patients have been treated in the clinical trial. The investigational dose for the incoming patient should satisfy both over-dose condition (22),

(22)

and the under-dose control condition (23),

(23)

where and are the feasibility bounds for toxicity and efficacy, respectively. These two conditions ensure that the probability of overly toxic dosing is less than , and the probability of under dosing is , given the currently observed clinical data. Additionally, the values of and usually vary as the trial proceeds, because information is limited at the beginning of the trial. Thus, using smaller value of and larger value of initially is preferred to estimate the MTD and the MED. As more data about the MED and MTD are collected during the trial, less conservative over-dose and under-dose controls can be utilized.  is set at 0.25 and its value is increased by a step size of 0.05 after each enrollment of new cohorts until the value of reaches 0.5. Chu, Lin and Shih [[11](#_ENREF_11)] showed that using their varying feasibility bound, , improves the speed of posterior estimation of the MTD to converge to the true MTD. Accordingly, is set to be decreased from 0.75 by a step size of 0.05 until it reaches 0.5. This under-dose control method has also been proved to be efficient in [[11](#_ENREF_11)].

In most cases, more than one investigational dose satisfies the over-dose and under-dose control conditions. These doses are called an acceptable dose set. To select one dose from the acceptable dose set for the incoming patient, the utility function is defined to evaluate each dose in the acceptable dose set:

(24)

where is the penalty term induced by toxicity. evaluates the tradeoff between the mean toxicity and the mean efficacy at a certain dose . Larger values of imply heavy penalty on toxicity, resulting in doses with low toxicity. The recommended dose chosen from the acceptable dose set would be the one with the highest utility (i.e., the highest value of ).

The dose-finding algorithm for our proposed design can be summarized as follows:

1. Treat the first cohort of patients at the minimum dose, .
2. Given the current cumulative clinical data , update the posterior distribution , and decide the recommended dose for the next cohort of patients that satisfies both over-dose and under-dose control conditions (22) and (23), with the highest posterior mean of utility, . However, if there are no doses fall within the acceptable dose set, the trial should be terminated.
3. Repeat steps 1) to 2), until the trial reaches the maximum sample size. Select the dose with the highest posterior mean to treat the next cohort of patients. This dose is the final recommended dose, called the optimized utility dose (OUD).
4. **Simulation**

In order to evaluate the performance of EWOUC-NETS, simulation studies are investigated comparing EWOUC and EWOC, under different scenarios.

***3.1 Simulation settings***

EWOUC-NETS is compared with EWOUC and EWOC. Three main aspects are compared between the designs: the accuracy of dose recommendation, therapeutic effects, and distribution of patients, under five different scenarios (Table 1). These scenarios simulate the different situations that could be encountered in a solid tumor trial. Scenario 1 (S1) is corresponded with an extremely good agent. Scenario 2 (S2), Scenario 3 (S3), Scenario 4 (S4) and Scenario 5 (S5) correspond to a good agent, a moderate agent, a bad agent and an extremely bad agent, respectively. Five dose levels (0.2, 0.4, 0.6, 0.8, and 1.0) are studied. Based on [[4](#_ENREF_4)], the targeted normalized equivalent score is set to be 0.476, and it is also the highest tolerated toxicity rate . Also, the minimal efficacy bound , is assumed to be 0.33. The MTDs and MEDs under each scenario are summarized in Figure 1. is set for all scenarios. It is assumed in the simulation settings that the toxicity outcome can be observed immediately, but it takes three months to assess the efficacy outcome of patients. We assign a uniform prior distribution (0, 0.333), and is assigned to have a uniform distribution (0, 0.5). Additionally, and both have independent uniform distribution (0.2, 1.2). is set to have a uniform distribution (-1, 1). The number of patients used for each trial is set to be 30 unless the trial is terminated in advance. The Metropolis-Hastings algorithm is used to sample from posterior distribution. The first 1000 iterations are treated as burn-in period and only the last 1000 iterations are utilized as the sample from posterior distribution. 1000 simulations are repeated for each scenario to evaluate the performance of designs.

* 1. ***Dose identification accuracy***

The accuracy of final dose recommended for the next patient is one of the most important factors to evaluate the performance of trial designs. We compare EWOUC-NETS with EWOC and EWOUC to evaluate the additional effect of using continuous toxicity and efficacy outcomes. EWOUC-NETS utilized the best utility to identify the recommended dose level for S1, S2 and S3. The OUD identified in these three scenarios are dose level 3, dose level 3 and dose level 4, respectively (Table 1). Under Scenario 1, EWOUC-NETS has a different recommended dose level than EWOC but agrees with EWOUC. However, the dose recommendation percentage is much higher than EWOUC (95.10% vs. 57.60%) while EWOC design prefers dose level 4 with recommendation percentage 50.30%. EWOC also has a relatively high recommendation percentage 45.40% at dose level 5. In addition, there is 0.3% of recommendation belongs to “None” when using EWOUC design, while all the recommendation percentage falls into five dose levels when using EWOUC-NETS. This also implies that EWOUC-NETS has almost no loss of information regarding identifying the OUD. Under S1, EWOC design selects dose level 4 instead of dose level 3, compared with EWOUC design and EWOUC-NETS design. Similarly, under S2, EWOUC-NETS has the same recommended dose levels with EWOUC, but the OUD is lower than the dose level selected by EWOC. EWOUC-NETS and EWOUC identify dose level 3 as the OUD with recommendation percentage 98.20% and 65.10%, respectively. EWOC chooses dose level 4 for the next cohort with recommendation percentage 64.80%. Under S3, both EWOUC-NETS and EWOUC are in favor of dose level 4 with recommendation percentage 77.40% and 76.10% respectively. In contrast, EWOC has the highest recommendation percentage (73.60%) at dose level 5. Although the recommendation rates are similar, EWOUC-NETS is still slightly higher than the other two designs. Since no targeted dose found in S4 and S5, the trial should be terminated early under S4 and S5. There is no recommendation percentage at each dose level for EWOUC-NETS design under both of S4 and S5. Nevertheless, for EWOUC, there is still small recommendation percentage falling into relatively lower dose levels under S4 (0.10% at dose level 1; 0.20% at dose level 2) and S5 (0.10% at dose level 1). EWOC still has high values of recommendation percentage at first three dose levels under S4 (53.00% at dose level 1; 43.30% at dose level 2; 3.70% at dose level 3) and S5 (53.40% at dose level 1; 42.70% at dose level 2; 3.90% at dose level 3). EWOC chooses a dose level by using the true MTD rather than the OUD. However, this could be more toxic for patients in the future clinical studies. The dose level selected by EWOUC and EWOUC-NETS designs are less toxic since they take into account the minimal efficacy and utilize the highest utility to select the final recommended dose level. Compared with EWOUC, EWOUC-NETS design provides a more accurate result under all five scenarios as the dose recommendation percentages are much more concentrated at the OUD and there is less loss of information when identifying the OUD.

* 1. ***Therapeutic effects***

To evaluate the performance of Phase I clinical trial design, therapeutic efficacy is undoubtedly a highly important dimension. One should consider therapeutic efficacy to be a composite of drug efficacy and toxicity. Furthermore, the toxicity of a dose usually increases as the drug efficacy increase. So, the utility function proposed in Section 2 could be a useful tool to measure the therapeutic effects.

The simulation results of average NETS and utility by different clinical trial designs under three scenarios are summarized in Table 2 and Table 3, respectively. The performance of EOUWC-NETS in the simulation studies is compared with previous designs. Although binary toxicity outcome has been used in EWOC and EWOUC, the binary toxicity outcome can be considered as a sketchy conversion of NETS. So, the DLT rates and average NETS can be compared in a rough sense. Under S1, S2 and S3, values of the average NETS of EWOUC-NETS are much lower than the DLT rates of EWOC and EWOUC designs. Under S1, the DLT rates for EWOC and EWOUC designs are 0.17 and 0.12 respectively while the average NETS for EWOUC-NETS is approximately 0.00. Also, they are 0.21, 0.16, and 0.04 under S2 and 0.16, 0.15 and 0.04 under S3. We notice that under S1, S2 and S3, EWOC design has the highest DLT rate/Average NETS and EWOUC-NETS design has the lowest. Under S1, S2 and S3, the DLT rates of EWOC design are the lowest among three designs, because EWOC treats most patients at nearly the MTD and the rate of DLT is relatively high at such dose levels. However, under S4 and S5, values of the average NETS of EWOUC-NETS design are relatively higher than DLT rates of the other two designs and are closer to the DLT rates of EWOUC design. Under S4, the DLT rates for EWOC and EWOUC designs are 0.22 and 0.31 but the average NETS for EWOUC-NETS is 0.37. Similarly, they are 0.21, 0.31 and 0.37 respectively under S5. Under each scenario, EWOUC-NETS design has higher expected utility than EWOUC design. For example, in S1, the expected utility increased 50% (from 0.44 to 0.66) under EWOUC-NETS design, compared with that of EWOUC design. More clearly, the expected utility increased from 0.33 to 0.79 under S2, and from 0.11 to 0.39 under S3. Under S4 and S5, the expected utilities of EWOUC-NETS design are both 0.22, while EWOUC has -0.57 and -0.59, respectively. As EWOUC-NETS uses NETS and continuous efficacy to describe patients’ toxicity and efficacy outcome, the expected utility calculated using EWOUC-NETS would be more precise to reflect the therapeutic effects on patients. The performance of EWOUC-NETS is also evaluated in the aspect of the percentage of patients treated at the OUD. We compare the EWOUC-NETS with EWOC and EWOUC under all five scenarios. The simulation results for patient distribution are summarized in Table 4. Under S1, S2 and S3, most patients are treated at the OUD under EWOUC and EWOUC-NETS, with 61.60% and 72.48% under S1, 63.29% and 75.99% under S2, and 51.98% and 61.46% under S3. However, the majority of patients are treated near MTD in EWOC design. For instance, 40.46% of patients are treated at dose level 4. Nevertheless, EWOUC and EWOUC-NETS have only 16.43% and 10.72% of patients at dose level 4. EWOUC-NETS has lower percentage of patients above the OUD and this can also lower the risk of over-dosing. Moreover, the performance of EWOUC-NETS is also investigated under scenarios that are not ideal. For instance, under S4 and S5, although no targeted dose level is identified, EWOUC-NETS design has lower percentage of patients treated above S2, compared with EWOC and EWOUC design. This implies that, no matter what the situation is, EWOUC-NETS has lower risk of over-dosing than EWOC and EWOUC designs. This is due to the utilization of both continuous toxicity and efficacy endpoints instead of binary indicators. Moreover, over-dose and under-dose control avoid large percentage of patients treated at dose level above the OUD or at dose level which cannot provide efficacious benefits. Although EWOUC and EWOUC-NETS select the exactly the same recommended dose level for the upcoming cohort under S1, S2 and S3, the percentages of recommendation and patients treated at such dose levels are obviously higher under EWOUC-NETS design. On the other hand, for EWOUC-NETS design, 83.34%, 83.31% and 65.90% of patients are treated above the MED under S1, S2 and S3, respectively. EWOC and EWOUC have relatively smaller portion of patients treated equal or above the MED (S1: EWOC-79.24% and EWOUC-78.03%; S2: EWOC-76.03% and EWOUC-76.43%; S3: EWOC-70.01% and EWOUC-54.38%). We notice that only under S3, EWOC design has higher portion of patients treated above the MED, but most of the patients are treated at the MTD. This could lead to higher DLT rate and less drug safety. This also implies that the under-dose control of EWOUC-NETS effectively lowers the probability of patients being treated at dose levels lower than the MED. Also, in S2, EWOC design treats patients mostly at dose level 5 (the MTD) while patients in EWOUC and EWOUC-NETS designs are mostly treated at dose level 4. Hence, the risk of over-dosing for patients is reduced by using EWOUC and EWOUC-NETS designs and the minimal efficacy is guaranteed in the meantime as the final doses recommended are above the MED. On the other hand, although EWOUC also accounts for over-dose and under-dose controls, the accuracy of identifying the OUD is lower than EWOUC-NETS, as most information is discarded if using binary dual endpoints. For instance, under S5, the MTD is set to be dose level 2, however, EWOUC design still has 14.57% patients treated above the MTD while EWOUC-NETS only has 1.29%. Thus, EWOUC-NETS is safer and more trustable for patients when dealing with extremely bad agents.

* 1. ***Sensitivity analysis***

We want to examine the robustness of the EWOUC-NETS design under relatively good scenarios (S1 and S2) using more combinations of () in extensive simulations, where is set to be 1, 2 and 3 in the joint model and is set to be 0.3 and 0.4.

We noticed that, under Scenario 1, for each combination of (), EWOUC-NETS design can detect the OUD correctly in at least approximately 83% of the cases with more than 54% of patients treated at the true OUD dose level. Similar results can be found under Scenario 2 (OUD can be detected correctly in more than 83% cases with 57% of patients treated at the true OUD dose level). The accuracy of the OUD detection and the therapeutic effects do not change much as the value of changes.

1. ***Application of EWOUC-NETS using a real trial example***

We used EWOUC-NETS to re-design the Phase I/II clinical trial described in Section 2. Thirty patients in 10 cohorts will be recruited and treated. Since the study is still in the design stage, real study data are not available. Hence, sample size n = 30 has been used for the real trial setting. The estimates are all generated from simulation studies for prospective planning purpose. The dose level 1 is set to be 20mg of BKM120 and 10mg of everolimus, and this starting dose has already been confirmed to be safe single agent doses. Other dose levels and results are summarized in Table 5. In this application, the targeted tolerated level (TTL), , is set to be 0.476. Meanwhile, the minimum efficacy bound, is selected as 0.33. Also, it is assumed that the MTD and the MED follow uniform distributions (20, 120). The parameter , which affects the correlation between the toxicity and efficacy endpoints for the joint factorization model, is set to be 0. Besides, the priors, and , at the starting dose level, are chosen as 0.01 and 0.02, respectively. Additionally, the parameter , is assumed to be 7. Furthermore, based on the cumulative clinical data **,** the posterior distribution will be updated and the toxicity and efficacy outcomes of the next cohort of patients will be assigned to the dose level with the highest posterior estimate of the utility  , from all the dose levels that satisfy both the over-dose and under-dose controls. If no dose levels are in the acceptable dose set, due to high toxicity or low efficacy, and , the trial will be terminated early. Otherwise, patients will be treated at the MTD for under-dose control, assuming the MTD is the most efficacious dose level. As we set the maximum sample size is 30, the trial will stop when the maximum sample size has been reached. The final OUD will be the dose level with the highest utility, and it will be the recommended dose level for the next cohort of patients in the clinical trial.

The dose level 3 is assumed to be the targeted dose level, and a total number of 1000 replicates of the trial has been generated. Also is chosen in the utility function. Other operating characteristics and details are summarized in Table 5. From the study result, we could see that EWOUC-NETS has the probability of approximately 91.70% to correctly identify the OUD. In the meantime, about 71.24% of patients are treated under the OUD, which is also the highest portion of patients that is treated under a certain dose level. The expected (mean) utility is 0.61. The average NETS of patients is around 0.0043 and approximately 78.10% of patients will benefit from efficacy events based on our 1000 replicates.

1. ***Discussion and Conclusion***

In this paper, we extend the EWOUC design by incorporating NETS and using continuous efficacy. The over-dose control and under-dose control ensures drug safety and improves the therapeutic effects for patients enrolled in the clinical trial.

In the EWOUC-NETS design, NETS [[4](#_ENREF_4)] is employed to model the toxicity outcome while the efficacy is also considered to be continuous by using the factorization model proposed by Hirakawa in 2012 [[8](#_ENREF_8)]. In this paper, we proposed a factorization model to describe the correlated joint probability of toxicity events and efficacy outcomes in Phase I cancer clinical trials. Using EWOUC-NETS, we can fully utilize the toxicity and efficacy information of patients in clinical trials and provide more accurate final dose estimation with higher safety and better therapeutic effects. We evaluate the performance of EWOUC-NETS mainly in two aspects: dose identification accuracy and therapeutic effects on patients, and we compare the performance in these two aspects with EWOC and EWOUC designs. Through extensive simulation studies, we have found that EWOUC-NETS functions well in identifying a final recommended dose between the MED and the MTD with the best utility. Under S1, S2 and S3, the average NETS of EWOUC-NETS designs are smaller than the DLT rates of EWOC and EWOUC designs. It might not be suitable to directly compare the average NETS and the average DLT rates, as the DLT rates cannot fully present and utilize the toxicity information of patients and are not accurate to represent the toxicity outcomes. Nonetheless, they can still serve as reference for clinicians to evaluate toxicity events in a rough sense.

The future research can involve the joint determination of multiple drug doses and personalized dose determination due to people’s different tolerance to the drugs, which is more accurate. So, this work can also be conducted in statistical genetics. Moreover, we can also extend the current EWOUC-NETS design from one single drug to two or multiple drugs, since most of the time, patients take more than one drug in a period, and there may exist drug interactions. Dose combination trials has drawn increasing attention nowadays in clinical trials. For example, Jimenez and his colleagues proposed a Bayesian adaptive design that uses drug combinations when a fraction of DLTs is attributable to one or more agents [[12](#_ENREF_12)]. This study only considers single endpoint, toxicity.

**ACKNOWLEDGEMENTS:**

Research reported in this publication was supported in part by the University of Illinois Cancer Center Biostatistics Shared Resource (BSR) core (ZC). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**CONFLICT OF INTEREST:**

The author(s) declared no potential conflict of interest with respect to the research, authorship, and/or publication of this article.

**REFERENCES:**

1. Chen, Z., et al., *Dose escalation with over-dose and under-dose controls in Phase I/II clinical trials.* Contemp Clin Trials, 2015. **43**: p. 133-41.

2. Babb, J., A. Rogatko, and S. Zacks, *Cancer phase I clinical trials: efficient dose escalation with overdose control.* Stat Med, 1998. **17**(10): p. 1103-20.

3. *Common Toxicity Criteria for Adverse Events v5.0*. 2021 [cited 2021; Available from: <https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm#ctc_50>.

4. Chen, Z., et al., *A novel toxicity scoring system treating toxicity response as a quasi-continuous variable in Phase I clinical trials.* Contemp Clin Trials, 2010. **31**(5): p. 473-82.

5. Chen, Z., M. Tighiouart, and J. Kowalski, *Dose escalation with overdose control using a quasi-continuous toxicity score in cancer Phase I clinical trials.* Contemp Clin Trials, 2012. **33**(5): p. 949-58.

6. Tao, Y., et al., *Dose-finding based on bivariate efficacy-toxicity outcome using Archimedean Copula.* PLoS One, 2013. **8**(11): p. e78805.

7. Rogatko, A., et al., *Escalation with Overdose Control is More Efficient and Safer than Accelerated Titration for Dose Finding.* Entropy (Basel), 2015. **17**(8): p. 5288-5303.

8. Hirakawa, A., *An adaptive dose-finding approach for correlated bivariate binary and continuous outcomes in phase I oncology trials.* Stat Med, 2012. **31**(6): p. 516-32.

9. Olkin, I. and R.F. Tate, *Multivariate Correlation Models with Mixed Discrete and Continuous Variables.* The Annals of Mathematical Statistics, 1961. **32**(2): p. 16.

10. Harvey, A., *Estimating regression models with multiplicative heteroscedasticity.* Econometrics., 1976. **44**: p. 4.

11. Chu, P., Y. Lin, and W. Shih, *Unifying CRM and EWOC designs for phase I cancer clinical trials.* Journal of Statistical Planning and Inference, 2009. **139**: p. 1146-1163.

12. Jimenez, J.L., M. Tighiouart, and M. Gasparini, *Cancer phase I trial design using drug combinations when a fraction of dose limiting toxicities is attributable to one or more agents.* Biom J, 2019. **61**(2): p. 319-332.