Published in final edited form as:

Clin Trials. 2017 December ; 14(6): 611–620. doi:10.1177/1740774517723829.

Repeated measures dose-finding design with time trend detection in the presence of correlated toxicity data

Jun Yin¹, Xavier Paoletti², Daniel J Sargent^{1,†}, and Sumithra J Mandrekar¹

¹Department of Health Sciences Research, Mayo Clinic, Rochester, USA

²Biostatistics and Epidemiology Department, INSERM CESP OncoStat; Institut Gustave Roussy, Villejuif, France

Abstract

Background—Phase I trials are designed to determine the safety, tolerability and recommended phase 2 dose of therapeutic agents for subsequent testing. The dose finding paradigm has thus traditionally focused on identifying the maximum tolerable dose of an agent or combination therapy under the assumption that there is a non decreasing relationship between dose toxicity and dose efficacy. The dose is typically determined based on the probability of severe toxicity observed during the first treatment cycle. A novel endpoint, the total toxicity profile, was previously developed to account for the multiple toxicity types and grades experienced in the first cycle. More recently, this was extended to a repeated measures design based on the total toxicity profile to account for longitudinal toxicities over multiple treatment cycles in the absence of within patient correlation.

Methods—In this work, we propose to extend the design in the presence of within patient correlation. Furthermore, we provide a framework to detect a toxicity time trend (toxicity increasing, decreasing, or stable) over multiple treatment cycles. We utilize a linear mixed model in the Bayesian framework, with the addition of Bayesian risk functions for decision making in dose assignment.

Results—The performance of this design was evaluated using simulation studies and real data from a phase I trial. We demonstrated that using available toxicity data from all cycles of treatment improves the accuracy of maximum tolerated dose identification and allows for the detection of a time trend. The performance is consistent regardless of the strength of the within patient correlation. In addition, the use of a quasi-continuous total toxicity profile score significantly increased the power to detect time trends compared to when binary data only were used.

Conclusions—The increased interest in molecularly targeted agents and immunotherapies in oncology necessitates innovative phase I study designs. Our proposed framework provides a tool to tackle some of the challenges presented by these novel agents, specifically through the ability to understand patterns of toxicity over time, which is important in the cases of cumulative or late toxicities.

Corresponding Author: Sumithra J. Mandrekar, Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA, mandrekar.sumithra@mayo.edu.

[†]Deceased

Keywords

Dose finding; phase I clinical trials; maximum tolerated dose; adaptive; time trend; repeated measures; correlated toxicity; immunotherapy; molecularly targeted agents; continual reassessment method

Introduction

Phase I trials are designed to determine the safety, tolerability and recommended phase 2 dose of therapeutic agents for subsequent testing. The dose finding paradigm has traditionally focused on identifying the maximum tolerable dose (MTD) of an agent or combination of agents based on toxicity with the assumption that there is a positive, proportional relationship between dose and tumor cell kill (efficacy). In conventional dose finding designs, only the dose limiting toxicities (DLT) within the first cycle of treatment have been considered in the decision making; adverse events occurring in later treatment cycles are not formally considered. Furthermore, a binary variable indicating the occurrence of DLT based on grade 3–4 events is generally utilized, discarding mild or moderate toxicities or the risk that a patient suffers from multiple adverse events during the same treatment cycle.

Over the past decade, the emergence of molecularly targeted agents and immune checkpoint targeted antibodies has necessitated changes to the early drug development process.¹ These agents have milder toxicity profiles and thus the definition of DLT only taking account of grade 3–4 toxicities may not be appropriate. Among 13 phase 1 trials of immunotherapies,^{2–12} low rates of grade 3–4 toxicities were observed (<14%) in 12 trials, except for 1 trial¹⁰ where 36% grade 3–4 toxicities (10/28 patients) were observed, although 4 DLTs were late onset in that trial. Hence, most trials were not able to identify a MTD which was defined per-protocol to target a 17–33% DLT occurrence. Recent studies found that more than half of the 445 patients in the 36 clinical trials of molecular targeted agents developed their worse grade toxicity after the first cycle;¹³ for example immune related adverse events from ipilimumab were reported to occur 8–10 weeks after the treatment initiation.¹⁰ Therefore, it is clear that for these agents, both lower grade toxicity as well as toxicities observed beyond cycle 1 should be considered during the dose finding process.

A variety of toxicity scoring systems have been proposed to measure quantitatively and comprehensively the overall severity of multiple toxicities for a patient, such as the equivalent toxicity score¹⁴ or total toxicity burden.¹⁵ Suppose that there are $k=1 \dots K$ patients, let T_{ik} be the i th ($i=1 \dots K$) toxicity event of the k th patients. Equivalent toxicity score converts each toxicity event T_{ik} into a numerical value of adjusted grade G_{ik} predefined according to the severity of the toxicity event, while total toxicity burden is defined as the sum of the weights of all toxicities experienced by a patient, where weight reflects the relative clinical importance of each grade and type of toxicity.

Few dose-finding designs have been proposed that take into account toxicities from subsequent treatment cycles. Cheung and Chappell¹⁶ proposed a sequential dose-finding design using time to first DLT occurrence as the endpoint where interest is in measuring the

risk of cumulative toxicities over a given period.¹⁶ A dose-finding design using a mixed-effect proportional odds model for longitudinal ordinal toxicity in phase I oncology trials was proposed by Doussau et al.,¹⁷ which was suboptimal in the case of scarce data. We previously developed a Bayesian phase I design,¹⁸ incorporating a quasi-continuous toxicity endpoint, the total toxicity profile, from multiple treatment cycles. The total toxicity profile score captures a comprehensive toxicity profile including multiple types and grades of toxicities occurring during the first treatment cycle.¹⁹ The repeated measures design using longitudinal toxicity data from multiple cycles (assuming independence across cycles) improved the ability to identify the MTD compared to a design that utilized data from just one cycle.¹⁸ In this work, we further investigate the performance of the repeated measures design to detect toxicity trends over time both in the presence and absence of correlated toxicity data within a patient who receives multiple treatment cycles over time. The information about the time trend of the toxicity profile is important to guide the drug administration over time, especially in the setting of molecular targeted agents which are usually administered until disease progression or unacceptable toxicity.

In our previous work, we have focused exclusively on dose-finding methodology only. And in this article, the focuses are now towards time-trend detection methods and implementation. Different from our previous work, posterior estimates in toxicity from all cycles were used to make inference on time trend. In practice, it could help identify potential cumulative mild/moderate toxicities and late on-set toxicities associated with molecular targeted agents and immunotherapies, when a significant positive trend is present. A method to generate correlated multinomial data was utilized in this paper. Phase I data in reality is multi-dimensional; and therefore the process of generating complex but realistic phase I toxicity data is a challenge. Previously, we generated data in a simpler manner: independent longitudinal toxicity data using a non-parametric approach. In this article, we generated correlated longitudinal toxicity data, and demonstrated that the strength of within patient correlation has minimal impact on the performance of the model-based design. Furthermore, a different comparator is used in this paper: the quasi-likelihood continual reassessment method design was used in our previous work to assess dose-finding performance, and whereas the LMM-CRML was used here to compare the time trend detection capabilities of a continuous toxicity score to a DLT binary endpoint. In addition to the simulation tools we applied in the previous work, we further introduced a real life example using the individual patient data from a National Cancer Institute sponsored phase I trial in this paper. This is the first demonstration of total toxicity profile based toxicity over time in a real phase I trial.

The article is organized as follows: we first review the extension to the repeated measures design using the longitudinal total toxicity profile scores from multiple treatment cycles for dose-escalation. We then introduce an approach for generating correlated longitudinal toxicity data. Simulation results are presented for the performance of the repeated measures design in the presence of within-patient correlation (to account for the repeated toxicities from multiple cycles within a given patient). The main findings of this work and ideas for future research are presented in the Discussion section.

Motivating example

A recent phase I study (NCT00048334) explored a day 1, 3, and 5 schedule of romidepsin, a potent histone deacetylase inhibitor, in patients with advanced solid tumors,²⁰ using a 3+3 dose escalation design. DLT was defined in the first cycle (21 days) as a hematologic toxicity of grade 4 absolute granulocyte count for more than 5 days; grade 4 platelet count; grade 3+ non hematologic toxicity, excluding potassium, magnesium, calcium, phosphate, uric acid, nausea, and vomiting unless these events are scored as grade 4 or occur despite maximal prophylaxis. Twenty-eight patients were evaluated at six increasing doses of romidepsin ranging from 1 to 9 mg/m². Two patients experienced a DLT: one patient had a grade 3 hypoxia and atrial fibrillation at dose level 2 (2 mg/m²) and the other patient had difficult-to-treat grade 3 nausea and anorexia at dose level 6 (9mg/m²). While no other patient had a DLT at dose 9mg/m² (DLT probability = 11%, 95% CI: 0.3–48%), most patients did not tolerate prolonged dosing at 9mg/m². Late toxicities were observed and prompted dose reduction in 7 out of the 9 patients enrolled at 9mg/m². The 7mg/m² dose was therefore the recommended phase II dose with the option to increase to 9mg/m² if tolerable by the patient. A total of 100 cycles were delivered to 28 patients; 11 patients received six cycles of treatment. One hundred and fourteen cases of grade 3 toxicity and five cases of grade 4 toxicity (in four patients) were recorded, including 93 grade 3 and 4 toxicities after the first cycle.

Methods

Repeated measures design model and estimation

Yin et al.¹⁸ previously developed a Bayesian phase I design accounting for comprehensive toxicity profiles from multiple treatment cycles for dose-finding. Patient-level toxicity profiles were quantified by a quasi-continuous toxicity score, the Total Toxicity Profile (*TTP*), first proposed by Ezzalfani et al.¹⁹ To compute the *TTP* score, the specific toxicities have to be identified and their possible outcomes graded by the Common Terminology Criteria for Adverse Events,²¹ which ranges from grade 0 (no toxicity), to grade 4 (life-threatening). Note that we limit the toxicity outcome to only grade 0–4, because the occurrence of a grade 5 event, corresponding to death, would require direct interaction between the entire study team and potentially regulatory bodies to determine whether to continue or terminate the trial. Pre-specified weights are then assigned to the toxicities to characterize the clinical importance of each grade of each type of toxicity. Given a patient's observed toxicity profile, the *TTP* score is computed as the Euclidean norm of the weight matrix given the observed toxicity.

To be consistent with Ezzalfani et al.,¹⁹ let us consider 3 types of toxicities: renal, neurologic and hematologic toxicities. Table 1 lists the possible toxicity outcomes, severity weight for each outcome, and whether its occurrence is considered a DLT. A patient with grade 2 renal, grade 0 neurologic, and grade 4 hematologic toxicities will have $TTP = \sqrt{0.75^2 + 0 + 1^2} = 1.25$. The *TTP* is further divided by a normalization constant v , so that the resulting *nTTP* score is within the range of (0,1): $nTTP = TTP/v$ (more details on normalization are available in article by Ezzalfani et al.¹⁹). Yin et al.¹⁸ proposed an adaptive

dose-finding design based on the $nTTP$ using a repeated measures model for longitudinal toxicity data. Due to the small sample size in phase I trials, parsimonious parametric models have been preferred as a working model to guide dose escalation;^{22–24} therefore, a linear model with a random intercept to model the longitudinal toxicity data was used. Assume that n patients are to be sequentially enrolled with K dose levels: $x_1, \dots, x_K$ representing the dose at K levels; and further assume that a patient i is treated at the same dose throughout the trial at all the cycles $j=1, \dots, J$. Let y_{ij} denote the observed $nTTP$ value for the patient i at the cycle j . The linear mixed model can be described as:

$$y_{ij} = \beta_0 + \beta_1 x_i + \beta_2 t_j + \gamma_i + \varepsilon_{ij} \quad (1)$$

where β_0 is the effect of the intercept; β_1 captures the dose toxicity relationship (we assume a monotonically increasing relationship where $\beta_1 > 0$); β_2 captures the time trend of toxicity where $\beta_2 > 0$ implies an increasing toxicity over time and $\beta_2 < 0$ implies a decreasing toxicity rate over time; ε_{ij} is the measurement error $\varepsilon_{ij} \sim N(0, \sigma^2)$; and γ_i is the random intercept for each patient where $\gamma_i \sim N(0, \sigma_\gamma^2)$, and σ_γ^2 is the variance of the random intercept. Assuming that each patient has a different overall tolerance to treatment, the random intercept γ_i is introduced to account for the correlation within the same patient for toxicities observed over repeated cycles of treatments, as well as contributing to the inter patient variability. The MTD is defined as the dose with the $nTTP$ score at cycle 1 closest to the pre specified target π_1 in the Bayesian risk function: $\argmin_{x_k} E|nTTP_1(x_k) - \pi_1|$, where $nTTP_1(x_k)$ represents the $nTTP$ score at cycle 1 given dose x_k . The Bayesian risk is defined as the posterior expected value of the distance between $nTTP_1(x_k)$ and π_1 . Although the decision rule is only based on the target toxicity at cycle 1, the accumulated toxicity data from multiple treatment cycles is utilized in the parameter estimation. A Bayesian inference paradigm coupled with a Markov chain Monte Carlo algorithm is adopted here similar to the original continual reassessment method design.²⁵ We have also run additional simulations that defining MTD using multiple cycles of data upon completion of the trial. More details are included in the online appendix.

Repeated measures design dose escalation procedure

Following the principle of adaptive dose-finding designs, patients are sequentially enrolled starting with the lowest dose. Estimates of the toxicity score at dose level k at cycle 1, $nTTP_1(x_k)$, are obtained using Model (1); the decision criteria for dose allocation is based on the selection of the dose that minimizes $E|nTTP_1(x_k) - \pi_1|$. A new cohort of patients is enrolled at the selected dose only after the previous cohort completes at least one cycle of treatment. The algorithm can be summarized as below:

1. The 3+3 design is used for the first two patient cohorts as a run in, to provide data for initial parameter estimation before switching to Model (1).
2. After run in, before each new cohort of patients are treated:
 - a. Fit the linear mixed model using all collected data from all patients treated thus far in the trial.

- b. Evaluate the decision criteria and identify the recommended dose that minimizes the Bayesian risk between the estimated $nTTP$ score and the target $nTTP$ score at cycle 1.
 - c. Treat the new cohort with the recommended dose, or the next higher dose if a decision is made that dose levels may not be skipped during dose escalation.
 - d. No intra patient dose escalation is permitted: patients are treated with the same dose until they are removed from the trial due to excessive toxicity or disease progression.
3. Terminate the trial when the pre specified maximum number of patients has been reached.

Upon completion of the trial, the trend in the toxicity over time (from multiple cycles of treatment) can be estimated using the complete longitudinal data. The 95% credible interval of β_2 from Model (1) is computed: if the 95% credible interval does not include 0, it indicates the presence of a time effect (increasing or decreasing); otherwise, we conclude that the evidence does not support the presence of a time effect.

Simulation studies

Simulation studies were used to assess the performance of the repeated measures design in terms of recommending the correct dose and time trend detection. We included a fixed sample size of 36 patients, and considered six dose levels. Every patient was administered a maximum of six treatment cycles, and treatment was stopped anytime if the patient experienced a DLT, generating missing data at subsequent cycles.

Data generation

We extended the approach of Ezzalfani et al.,¹⁹ of generating $nTTP$ data for the first cycle to generate correlated toxicity data over multiple cycles. In Ezzalfani et al., three toxicity types (renal, neurological, and hematological) were considered for illustration purposes and they were assumed independent; the associated weight matrix as well as indicator of DLT are shown in Table 1. For each toxicity type, the probability matrix of observing grade 0 to 4 at each dose level (from dose level 1 to 6) was pre-defined (details can be found in the study by Ezzalfani et al.¹⁹). We used the pre-defined toxicity probabilities in the article by Ezzalfani et al.¹⁹ as the marginal toxicity probabilities for cycle 1, and then extended these toxicity probabilities for subsequent cycles assuming an increasing, flat, and decreasing time trend, as detailed below.

Denote the marginal toxicity probability at cycle j ($j=1, \dots, J$) as $p_{j, I, h}(x)$ of observing a grade h ($h=0, 1, 2, 3, 4$) for toxicity type I ($I=1, 2, 3$) at dose x_k . The marginal toxicity probabilities capture the mean structure of the ordinal toxicity data, including the marginal dose toxicity relationship as well as the marginal dose cycle relationship. The ordinal toxicity data generated from the marginal toxicities are considered independent across and within the cycles (see supplemental online Appendix for details). To induce correlation among the repeated measures over cycles within a patient, we employed a latent random effect.

Specifically, if we denote the individual toxicity probability at cycle as j as $p_{j,ih}(x)$ for patient i , and, w_{ih} as the random effect that captures the within patient correlation across multiple cycles, a Baseline Categorical Logit model incorporating this can be denoted as below:²⁶

$$\log \left(\frac{p_{ij,lh}(x_k)}{p_{ij,l0}(x_k)} \right) = \log \left(\frac{p_{j,lh}(x_k)}{p_{j,l0}(x_k)} \right) + w_{ilh}, h=0, 1, 2, 3, 4. \quad (2)$$

In this model, w_{ih} is the random intercept, following a Gaussian distribution $w_{ih} \sim N(0, \sigma_w^2)$ with an exchangeable correlation structure, and the parameter σ_w^2 controls the degree of correlation between cycles. The sum of the probability of observing a toxicity grade from 0 to 4 equals to 1 for any toxicity type: $\sum_{h=0}^4 p_{ij,lh}(x_k) = 1$ for any i, j, l and k .

Simulation scenarios

We considered 12 clinical scenarios in total, with 4 different MTD locations (dose 2, dose 3, dose 4, and dose 5) and 3 different time trends (no time trend, decreasing time trend, and increasing time trend). A total of 500 trials were simulated for each scenario. The MTD is defined as the dose level where the $nTTP$ equals the target of 0.28 at cycle 1; which corresponds to a probability of DLT of 33% (please refer to the study by Ezzalfani et al.¹⁹ for details on the target $nTTP$). The scenarios where the true MTD is respectively dose level 2, 3 or 4, and 5 represent severe, moderate and mild dose-toxicity relationships. In addition, we also conducted simulations with patient dropout due to disease progression. More details are included in the online appendix demonstrating the consistent performance of our model in these scenarios. Additional results for true MTD dose = 1 are also available in the appendix.

Performance evaluation criteria

For the dose finding evaluation, we evaluated the percentage of correct selection which is the proportion of the trials recommending the true target dose. We also reported the percentage of patients assigned to each dose level to evaluate the safety and efficiency of repeated measures design. The quasi-likelihood continual reassessment method¹⁹ is used as a benchmark for comparison. The quasi-likelihood continual reassessment method is an extension of the continual reassessment method approach where $nTTP$ scores instead of DLTs from cycle 1 are used as the toxicity outcome. In the quasi-likelihood continual reassessment method, $nTTP$ scores were assumed to follow quasi-Bernoulli distributions and were modeled using logistic regression models. As in conventional phase I trials, the quasi likelihood continual reassessment method only uses toxicity from cycle 1 for dose escalation, and hence is not applicable for time trend detection.

For time trend detection, we reported the percentage of simulations where a time trend was detected at the 5% level. Tests were carried out at the end of each trial by comparing the 95% credible interval of β_2 , which is the estimate of time effect, to 0. A significant time trend was declared when the 95% credible interval of β_2 excluded 0.

Results

Dose-finding

Within patient correlated toxicity data over cycles were generated by the above Baseline-Categorical Logit model (Eq.2); the parameter value of σ_w^2 was set to 10. Additional simulation scenarios of correlation parameters are available in the online appendix. Distributions of dose assignment and final dose recommendation are displayed in Table 2. Since the quasi-likelihood continual reassessment method only uses cycle 1 data, we first fit the repeated measures design model to cycle 1 data only (denoted as “RMD, cycle 1” in Table 2) for a fair comparison. Repeated measures design and quasi-likelihood continual reassessment method are comparable when only cycle 1 toxicity data was used. We then applied repeated measures design to longitudinal toxicity data from multiple cycles. In the absence of any time effect (denoted as “RMD, constant trend” in), the repeated measures design model of longitudinal toxicity data improved the identification of the target dose compared to the quasi-likelihood continual reassessment method and the repeated measures design using just cycle 1 data (4–9% improvement in percentage of correct selection) in all scenarios except when MTD was dose 2; the mean number of cycles was 2.0–3.1 per patient. Therefore, using toxicity data from multiple cycles is advantageous in the presence of no detectable time trend.

In case of an underlying true negative time effect (denoted as “RMD, decreasing trend” in Table 2), the repeated measures design model of longitudinal toxicity data consistently improved the identification of the target dose in all scenarios (an improvement of 3–6% in percentage of correct selection, compared to quasi-likelihood continual reassessment method); the mean number of cycles was 2.6 to 3.4 per patient. In case of an underlying true positive time effect (denoted as “RMD, increasing trend” in), repeated measures design performed slightly worse in severe dose-toxicity scenarios when the true MTD was dose levels 2 or 3 (4% less in percentage of correct selection than the quasi-likelihood continual reassessment method), but better in mild dose-toxicity scenarios when the true MTD was dose levels 4 or 5 (5–7% gain in percentage of correct selection, compared to the quasi-likelihood continual reassessment method); the mean number of cycles was 1.6 to 2.7 per patient.

Time-trend detection

Table 3 reports the percentage of trials where a time trend was detected at the completion of the trial using the longitudinal analysis at the 5% level. The false positive rate in the absence of a time trend is less than 13% in all scenarios, indicating an acceptable type I error control. In scenarios when a positive trend was present, the test for trend was significant in 56–67% of the trials, indicating promising power for detecting a positive trend. In scenarios when a negative trend was present, a considerable loss of power was observed (9–19% power) with repeated measures design.

The loss of power was mainly due to the missing data from DLT dropout. Within the treatment paradigm for each patient, treatment was stopped anytime a DLT was observed, leading to patient dropouts in later cycles for toxic scenarios (see Figure S1 in the Appendix

for more details). Hence in the presence of a negative trend, the data to demonstrate the lower toxicity in subsequent cycles was frequently absent. To have a better idea of the bias, although unrealistic, we also considered an analysis with complete longitudinal data, where all patients completed their 6 cycles of treatment, regardless of DLT occurrence. As expected, the power was above 98% for all the scenarios in this instance.

We further investigated the performance of repeated measures design when a stronger time effect was present in the data: the slope of time effect was increased by 95–126% for positive trends, and 72–110% for negative trends (please see the online Appendix for details). In the simulated data with a stronger time trend, the power of detecting a positive trend increased to above 88%, and the power of detecting a negative trend increased to 56–74% (Table 3).

Application

The data from the romidepsin trial previously introduced were re-analyzed using the continuous $nTTP$ scores. Only the final estimates of the time effect using the complete data are provided here, as the retrospective analysis of dose allocation is not directly feasible. For illustration purpose, we considered the three frequent types of toxicities: gastrointestinal (23.4%), liver (9.7%), and hematological (32.3%) adverse events, and their severity weights were assigned as in the following weight matrix (column represents grade 0–4, and row represents toxicity type in the order of gastrointestinal, liver and hematological toxicity):

$$w = \begin{pmatrix} 0 & 0.25 & 0.5 & 1 & 1.5 \\ 0 & 0.25 & 0.5 & 1 & 1.5 \\ 0 & 0 & 0 & 0.75 & 1 \end{pmatrix}$$

If a patient had multiple toxicities for a given type during the same cycle, only the maximum grade for that toxicity type was considered. The numbers of patients remaining in the study decreased over time (approximately half of the patients dropped out after cycle 2), grade 2 toxicities were experienced in later cycles and grade 3/4 hematological toxicities occurred in cycles 5 and 6 (Table 4).

Figure 2 shows the observed average $nTTP$ per cycle (dashed line) and the predicted $nTTP$ (solid line) from the repeated measures design model, demonstrating a good fit to the data despite the few patients per cycle. Analysis with repeated measures design yielded an estimate of time effect of 0.03 (95% CrI: 0.01–0.04), which equals to a 12% increase in the toxicity score ($12\% = 0.03/0.25$) from cycle 1 to cycle 2, if the observed $nTTP = 0.25$ in cycle 1. A caveat to this analysis is that only the planned dose was available instead of the actual administrated dose, especially for the seven patients who had a dose reduction from 9 to 7 mg/m² in the subsequent cycles due to toxicities.

Discussion

Recent development of cancer immunotherapies and molecularly targeted agents require innovative approaches beyond the traditional phase I designs commonly used for cytotoxic agents. Specifically, immunotherapies and molecular targeted agents often demonstrate

reduced rates of severe toxic reactions, however their adverse effects may persist or even accumulate over multiple treatment cycles, as opposed to the acute toxicities experienced with cytotoxic agents often observed in the first treatment cycle. Since prolonged minor to moderate toxicities can have a significant impact on a patient's quality life,²⁷ treatment compliance^{28–30} or even be life threatening if not properly managed,^{31–33} it is important to assess the trend of toxicity events across multiple treatment cycles.

In this study, we demonstrated through simulation studies that using available toxicity data from all cycles of treatment improves the accuracy of MTD identification and allows the detection of a time trend in many instances. In the absence of within-patient correlation, the repeated measures design consistently improved the estimation of the MTD; when within-patient correlation was present, the repeated measures design performed better than the quasi-likelihood continual reassessment method³⁴ in most cases. In the presence of a strong within-patient correlation, similar results were observed between the repeated measures design and the quasi-likelihood continual reassessment method (Table S1). In addition, the use of a quasi-continuous toxicity score (TTP) significantly increased the power to detect time trends compared to when binary DLT data were used.

The repeated measures design methods require input from the physicians involved in the trial. We note that this is no different than any other Phase I trial where the study team has to identify unacceptable grades and types of toxicities a priori. Initially, the physicians need to specify toxicities to be monitored and their possible clinical outcomes graded by the Common Terminology Criteria for Adverse Events. As an example, we use three types of toxicities: renal, neurological and hematological toxicities, for illustration purposes. The physicians are then asked to assign a severity weight for each grade of each toxicity within some numerical range. As the method is invariant to the range, the range only needs to be positive. Grade 0 corresponds to no toxicity and a null weight. We demonstrated one way of specifying the severity weights in the example. Based on the severity weights, the most severe toxicity profile (grade 4 for all three toxicity types) corresponds to a maximum TTP score of 2.34. In this case, the normalization constant was chosen as 2.5, just so it is slightly larger than the maximum TTP. By dividing TTP scores by the normalization constant, the $nTTP$ scores are constrained between 0 and 1. The assumption of normality of the toxicity score, $nTTP$ (either the raw as used in the simulations or transformed scores, in case of skewed data), is necessary for the linear mixed effect model assumptions. A natural extension would be to explore non-linear models with other link functions.

An interesting finding of this study is the lack of power to detect negative trends due to missing data from patient drop out in the later cycles. In the presence of a negative trend in toxicity over time (i.e., decreasing relationship between toxicity and cycle within each dose level), when a patient stops treatment due to a severe toxicity, the ability to observe the decreasing toxicity trend at subsequent cycles is compromised. The loss of power was even more remarkable when only a dichotomized toxicity endpoint (i.e., DLT) was used³⁵ (details in online Appendix). In practice, the existence of a decreasing trend may be considered less critical, as a dose that causes excessive early toxicity generally would not be considered appropriate even if the rate of toxicity could be lower in subsequent cycles. A study of 54 phase I cancer clinical trials of molecular targeted agents from 1999 to 2013 found that

approximately 20% of the patients had to be dose reduced after cycle 1.³⁶ Recently, ceritinib was approved by the Food and Drug Administration at the MTD determined based on toxicity from cycle 1; however, it was associated with late toxicity resulting in dose reduction for half of the patients in the trial.³⁷

Sensitivity analyses were conducted to investigate the performance of repeated measures design under varying strength of within patient correlation. The parameter σ_w^2 in the Baseline-Categorical Logit model (Eq.2) controls the degree of correlation in the longitudinal data: larger values of σ_w^2 result in stronger correlation in the data. A minimal reduction (<6%) in the percentage of correct selection was observed in some scenarios when stronger within-patient correlation was present (see online Appendix for details). No impact was observed in terms of time trend detection.

To conclude, the novelty of this approach includes three key features: (1) we extended our previously published repeated measures design to allow for time trend detection upon the completion of the trial. Our prior work focused only on dose-finding. We have demonstrated that the use of a continuous toxicity endpoint significantly increases the power to detect time trends compared to a binary endpoint; (2) we introduced a method to generate correlated multinomial data, and identified that the strength of within-patient correlation has minimal impact on the performance of this model based design; (3) we utilized individual patient data from a National Cancer Institute sponsored phase I trial, and applied our methods to the trial. Our proposed design provides a potential tool to tackle challenges presented by molecular targeted agents and immunotherapies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by National Institutes of Health Grants R01CA174779. Dr. Xavier Paoletti was partly supported by the European Community program ITN Marie Curie IDEAS and by the French NCI (INCA Optidose program).

Appendix

Alternative Approach for Time-trend Detection

The LMM-CRML¹ is an extension of continual reassessment method, based on a logistic mixed effect model for binary toxicity data from multiple treatment cycles. In this design, a random effect logistic relationship link the logit of the risk of dose limiting toxicities (DLT) at each cycle to the dose and the cycle (Table S2). Cycle is modeled as an ordinal variables and only log-linear time trend are investigated. Parameter estimates obtained from the multiple cycles of treatment provide in turn an estimate of the time trend effects in toxicity upon trial completion to match the design using *nTTP*.

Definition of MTD Using Multiple Cycles of Data

We have run additional simulations and considered defining maximum tolerated dose (MTD) using multiple cycles of data upon completion of the trial as the reviewer suggested. We defined decision rule 1 as the same as in the article: the MTD is defined as the dose with the nTTP score at cycle 1 closest to the pre-specified target 0.28. Decision rule 2 is defined as the dose with the average nTTP score over all cycles closest to the pre-specified target 0.28. Decision rule 3 defines MTD as the dose with the average nTTP closest to 0.28, given that all nTTP across all cycles at the dose = 0.28. The tables below show the results corresponding to the scenarios when the true MTD is dose level 2, with increasing, constant and decreasing trend, respectively. For results for other scenarios (true MTD = dose 3, 4, and 5), please refer to Table S3.

As shown in Table S4, when there is an increasing trend, rule 1 ignored the trend and still recommended dose 2 (whose mean nTTP was closest to the target 0.28) as the MTD (82%), however, rule 2 recommended dose 1 (97%) whose average nTTP over cycles was closest to the target, while rule 3 recommended dose 1 with only 61% times since the mean nTTP at dose 1 cycle 6 is 0.31. With a constant trend, there was no difference in the MTD identification. And with a decreasing trend, rule 1 and 3 still recommended dose 2, while rule 2 recommended a higher dose (dose 3) whose average nTTP was closest to 0.28. Similar results for true MTD dose = 3, 4 and 5 are available in Table S3.

Simulations with Patient Dropout due to Disease Progression

We have conducted extensive simulations to evaluate the time trend detection and target dose recommendation assuming a 10% and 20% dropout due to patient death; this is independent of DLT dropout (Table S5). With higher percentage of patient death and more patient dropout, some reduction in the accuracy of MTD identification was observed but a high percentage of correct selection (77–88%) was still maintained overall, indicating limited impact on the dose-finding. For time trend detection, the results are consistent as before: the power to detect either a positive or negative trend was even more compromised with more missing data.

Generating Marginal Toxicity Probabilities for Subsequent Cycles

We borrowed the pre-defined toxicity probabilities from Ezzalfani et al.² as our marginal toxicity probabilities for cycle 1, and then extend them for subsequent cycles in the following fashion. Denote, at cycle 1, the marginal toxicity probability of observing a grade h ($h=0,1,2,3,4$) for toxicity type I ($I=1,2,3$) at dose k as $p_{1,Ih}(x_k)$. Given the toxicity type I and the dose level k , we assumed that $p_{1,Ih}(x_k)$ was generated from a standard normal distribution such that $p_{1,I0}(x_k) = \Phi(C_{I,k,0})$, $p_{1,I1}(x_k) = \Phi(C_{I,k,1}) - \Phi(C_{I,k,0})$, $p_{1,I2}(x_k) = \Phi(C_{I,k,2}) - \Phi(C_{I,k,1})$, $p_{1,I3}(x_k) = \Phi(C_{I,k,3}) - \Phi(C_{I,k,2})$, and $p_{1,I4}(x_k) = 1 - \Phi(C_{I,k,3})$ where $C_{I,k,0}$, $C_{I,k,1}$, $C_{I,k,2}$, and $C_{I,k,3}$ are a set of fixed percentiles. Shifting the mean of the standard normal distribution to the right increases the probability of observing higher grades, which corresponds to a higher nTTP value; while shifting the mean to the left increases the probability of observing lower grades and hence corresponds to a lower nTTP value. Therefore, to generate marginal

toxicity probabilities with an increasing trend, we shifted the standard normal distribution to the right by 0.1 for cycle 2, by 0.2 for cycle 3, 0.3 for cycle 4, and so on and so forth. In the same manner, to generate marginal toxicity probabilities with a decreasing trend, we shifted the standard normal distribution to the left by 0.1 for cycle 2, by 0.2 for cycle 3, 0.3 for cycle 4, and so on and so forth. To generate a stronger trend as shown in Table 3, we increased the effect size of trend effect, by moving the standard normal by a larger step (0.2 instead of 0.1 per cycle).

References

1. Paoletti X, Doussau A, Ezzalfani M, et al. Dose finding with longitudinal data: simpler models, richer outcomes. *Stat Med.* 2015; 34:2983–2998. [PubMed: 26109523]
2. Ezzalfani M, Zohar S, Qin R, et al. Dose-finding designs using a novel quasi-continuous endpoint for multiple toxicities. *Stat Med.* 2013; 32:2728–2746. [PubMed: 23335156]

References

1. Postel-Vinay S, Aspeslagh S, Lanoy E, et al. Challenges of phase 1 clinical trials evaluating immune checkpoint-targeted antibodies. *Ann Oncol.* 2016; 27:214–224. [PubMed: 26578728]
2. Brahmer JR, Drake CG, Wollner I, et al. Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. *J Clin Oncol.* 2010; 28:3167–3175. [PubMed: 20516446]
3. Weber JS, O'Day S, Urba W, et al. Phase I/II study of ipilimumab for patients with metastatic melanoma. *J Clin Oncol.* 2008; 26:5950–5956. [PubMed: 19018089]
4. Hodi FS, Mihm MC, Soiffer RJ, et al. Biologic activity of cytotoxic T lymphocyte-associated antigen 4 antibody blockade in previously vaccinated metastatic melanoma and ovarian carcinoma patients. *Proc Natl Acad Sci U S A.* 2003; 100:4712–4717. [PubMed: 12682289]
5. Hamid O, Robert C, Daud A, et al. Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. *N Engl J Med.* 2013; 369:134–144. [PubMed: 23724846]
6. Maker AV, Yang JC, Sherry RM, et al. Inpatient dose escalation of anti-CTLA-4 antibody in patients with metastatic melanoma. *J Immunother.* 2006; 29:455–463. [PubMed: 16799341]
7. Powles T, Eder JP, Fine GD, et al. MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature.* 2014; 515:558–562. [PubMed: 25428503]
8. Herbst RS, Soria JC, Kowanzetz M, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature.* 2014; 515:563–567. [PubMed: 25428504]
9. Berger R, Rotem-Yehudar R, Slama G, et al. Phase I safety and pharmacokinetic study of CT-011, a humanized antibody interacting with PD-1, in patients with advanced hematologic malignancies. *Clin Cancer Res.* 2008; 14:3044–3051. [PubMed: 18483370]
10. Patnaik A, Kang SP, Rasco D, et al. Phase I study of pembrolizumab (MK-3475; anti-PD-1 monoclonal antibody) in patients with advanced solid tumors. *Clin Cancer Res.* 2015; 21:4286–4293. [PubMed: 25977344]
11. Camacho LH, Antonia S, Sosman J, et al. Phase I/II trial of tremelimumab in patients with metastatic melanoma. *J Clin Oncol.* 2009; 27:1075–1081. [PubMed: 19139427]
12. Ribas A, Camacho LH, Lopez-Berestein G, et al. Antitumor activity in melanoma and anti-self responses in a phase I trial with the anti-cytotoxic T lymphocyte-associated antigen 4 monoclonal antibody CP 675,206. *J Clin Oncol.* 2005; 23:8968–8977. [PubMed: 16204013]
13. Postel-Vinay S, Gomez-Roca C, Molife LR, et al. Phase I trials of molecularly targeted agents: should we pay more attention to late toxicities? *J Clin Oncol.* 2011; 29:1728–1735. [PubMed: 21444876]
14. Chen Z, Krailo MD, Azen SP, et al. A novel toxicity scoring system treating toxicity response as a quasi-continuous variable in phase I clinical trials. *Contemp Clin Trials.* 2010; 31:473–482. [PubMed: 20609419]

15. Bekele BN, Thall PF. Dose-finding based on multiple toxicities in a soft tissue sarcoma trial. *J Am Stat Assoc.* 2004; 99:26–35.
16. Cheung YK, Chappell R. Sequential designs for phase I clinical trials with late-onset toxicities. *Biometrics.* 2000; 56:1177–1182. [PubMed: 11129476]
17. Doussau A, Asselain B, Le Deley MC, et al. Dose-finding designs in pediatric phase I clinical trials: comparison by simulations in a realistic timeline framework. *Contemp Clin Trials.* 2012; 33:657–665. [PubMed: 22521954]
18. Yin J, Qin R, Ezzalfani M, et al. A Bayesian dose-finding design incorporating toxicity data from multiple treatment cycles. *Stat Med.* 2017; 36:67–80. [PubMed: 27633877]
19. Ezzalfani M, Zohar S, Qin R, et al. Dose finding designs using a novel quasi-continuous endpoint for multiple toxicities. *Stat Med.* 2013; 32:2728–2746. [PubMed: 23335156]
20. Amiri-Kordestani L, Luchenko V, Peer CJ, et al. Phase I trial of a new schedule of romidepsin in patients with advanced cancers. *Clin Cancer Res.* 2013; 19:4499–4507. [PubMed: 23757352]
21. Trotti A, Colevas AD, Setser A, et al. CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment. *Semin Radiat Oncol.* 2003; 13:176–181. [PubMed: 12903007]
22. Shen LZ, O'Quigley J. Consistency of continual reassessment method under model misspecification. *Biometrika.* 1996; 83:395–405.
23. Cheung YK, Chappell R. A simple technique to evaluate model sensitivity in the continual reassessment method. *Biometrics.* 2002; 58:671–674. [PubMed: 12230003]
24. Paoletti X, Kramar A. A comparison of model choices for the continual reassessment method in phase I cancer trials. *Stat Med.* 2009; 28:3012–3028. [PubMed: 19672839]
25. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase I clinical trials in cancer. *Biometrics.* 1990; 46:33–48. [PubMed: 2350571]
26. Agresti A. Logit models for multinomial responses. *Categorical data analysis (Second).* 2002:267–313.
27. Duran I, Hotte SJ, Hirte H, et al. Phase I targeted combination trial of sorafenib and erlotinib in patients with advanced solid tumors. *Clin Cancer Res.* 2007; 13:4849–4857. [PubMed: 17699864]
28. Sleijfer S, Wiemer E, Seynaeve C, et al. Improved insight into resistance mechanisms to imatinib in gastrointestinal stromal tumors: a basis for novel approaches and individualization of treatment. *Oncologist.* 2007; 12:719–726. [PubMed: 17602061]
29. Noens L, van Lierde MA, De Bock R, et al. Prevalence, determinants, and outcomes of nonadherence to imatinib therapy in patients with chronic myeloid leukemia: the ADAGIO study. *Blood.* 2009; 113:5401–5411. [PubMed: 19349618]
30. Chirgwin JH, Giobbie-Hurder A, Coates AS, et al. Treatment Adherence and Its Impact on Disease-Free Survival in the Breast International Group 1–98 Trial of Tamoxifen and Letrozole, Alone and in Sequence. *J Clin Oncol.* 2016; 34:2452–2459. [PubMed: 27217455]
31. Topalian SL, Hodi FS, Brahmer JR, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med.* 2012; 366:2443–2454. [PubMed: 22658127]
32. Gao J, He Q, Subudhi S, et al. Review of immune-related adverse events in prostate cancer patients treated with ipilimumab: MD Anderson experience. *Oncogene.* 2015; 34:5411–5417. [PubMed: 25659583]
33. Fecher LA, Agarwala SS, Hodi FS, et al. Ipilimumab and its toxicities: a multidisciplinary approach. *Oncologist.* 2013; 18:733–743. [PubMed: 23774827]
34. Ezzalfani M, Zohar S, Qin R, et al. Dose-finding designs using a novel quasi-continuous endpoint for multiple toxicities. *Stat Med.* 2013; 32:2728–2746. [PubMed: 23335156]
35. Paoletti X, Doussau A, Ezzalfani M, et al. Dose finding with longitudinal data: simpler models, richer outcomes. *Stat Med.* 2015; 34:2983–2998. [PubMed: 26109523]
36. Postel-Vinay S, Collette L, Paoletti X, et al. Towards new methods for the determination of dose limiting toxicities and the assessment of the recommended dose for further studies of molecularly targeted agents—Dose-Limiting Toxicity and Toxicity Assessment Recommendation Group for Early Trials of Targeted therapies, an European Organisation for Research and Treatment of Cancer-led study. *Eur J Cancer.* 2014; 50:2040–2049. [PubMed: 24880774]

37. Shaw AT, Kim DW, Mehra R, et al. Ceritinib in ALK-rearranged non-small-cell lung cancer. *N Engl J Med*. 2014; 370:1189–1197. [PubMed: 24670165]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Expected $nTTP$ at each dose level at each cycle when the true MTD is d_2 , d_3 , d_4 , and d_5 . The top panel shows dose toxicity curves with increasing trends; the bottom panel shows dose-toxicity curves with decreasing trends. The dashed lines represent the target $nTTP$ 0.28.

Figure 2.

Observed nTTP (dashed line) vs. predicted nTTP from the repeated measures design model (solid line) per cycle. The observed nTTP per dose per cycle were marked with colors (red represents dose level 1, yellow dose level 2, green dose level 3, light blue dose level 4, dark blue dose level 5, and pink dose level six).

Table 1

Toxicities and severity weights in the example

Toxicity Type	Grade	Severity	DLT
Renal	1	0.5	No
	2	0.75	No
	3	1	Yes
	4	1.5	Yes
Neurological	1	0.5	No
	2	0.75	No
	3	1	Yes
	4	1.5	Yes
Hematological	3	0.5	No
	4	1	Yes

DLT: dose limiting toxicities

Table 2

Dose recommendation and allocation percentage, with within patient correlated toxicity data over cycles.

	Dose Recommendation						Dose Allocation						# Cycles (Mean, SD)
	d ₁	d ₂	d ₃	d ₄	d ₅	d ₆	d ₁	d ₂	d ₃	d ₄	d ₅	d ₆	
True MTD = d₁													
QLCRM	0.90	0.10	0	0	0	0	0.70	0.25	0.05	0	0	0	NA
RMD, cycle 1	0.90	0.10	0	0	0	0	0.71	0.26	0.03	0	0	0	NA
RMD, decreasing trend	0.94	0.06	0	0	0	0	0.79	0.15	0.06	0	0	0	2.1(1.5)
RMD, constant trend	0.90	0.10	0	0	0	0	0.76	0.18	0.06	0	0	0	15(1.1)
RMD, increasing trend	0.90	0.10	0	0	0	0	0.74	0.19	0.07	0	0	0	1.2(0.8)
True MTD = d₂													
QLCRM	0.03	0.86	0.11	0	0	0	0.19	0.63	0.17	0	0	0	NA
RMD, cycle 1	0.03	0.86	0.11	0	0	0	0.30	0.67	0.03	0	0	0	NA
RMD, decreasing trend	0.08	0.89	0.03	0	0	0	0.25	0.73	0.02	0	0	0	2.7(2.0)
RMD, constant trend	0.13	0.82	0.05	0	0	0	0.28	0.69	0.04	0	0	0	2.0(1.8)
RMD, increasing trend	0.11	0.82	0.07	0	0	0	0.27	0.69	0.04	0	0	0	1.6(1.4)
True MTD = d₃													
QLCRM	0	0.03	0.84	0.13	0	0	0.09	0.15	0.57	0.18	0	0	NA
RMD, cycle 1	0	0.08	0.84	0.08	0	0	0.11	0.46	0.39	0.04	0	0	NA
RMD, decreasing trend	0	0.05	0.90	0.04	0	0	0.13	0.32	0.50	0.06	0	0	2.6(2.0)
RMD, constant trend	0	0.08	0.88	0.06	0	0	0.11	0.36	0.50	0.03	0	0	2.3(2.2)
RMD, increasing trend	0	0.19	0.80	0.01	0	0	0.14	0.40	0.44	0.03	0	0	2.1(1.9)
True MTD = d₄													
QLCRM	0	0	0.02	0.81	0.17	0	0.08	0.09	0.13	0.51	0.19	0	NA
RMD, cycle 1	0	0	0.02	0.81	0.17	0	0.09	0.15	0.19	0.48	0.09	0	NA
RMD, decreasing trend	0	0	0.08	0.87	0.05	0	0.09	0.13	0.22	0.47	0.08	0	2.9(2.2)
RMD, constant trend	0	0	0.07	0.90	0.03	0	0.09	0.14	0.22	0.48	0.07	0	2.7(2.1)
RMD, increasing trend	0	0	0.10	0.86	0.04	0	0.10	0.14	0.23	0.49	0.05	0	2.3(2.0)
True MTD = d₅													

	Dose Recommendation						Dose Allocation						# Cycles (Mean, SD)
	d ₁	d ₂	d ₃	d ₄	d ₅	d ₆	d ₁	d ₂	d ₃	d ₄	d ₅	d ₆	
QLCRM	0	0	0	0.03	0.80	0.18	0.08	0.08	0.08	0.12	0.45	0.18	NA
RMD, cycle 1	0	0	0	0.03	0.80	0.18	0.09	0.1	0.1	0.22	0.38	0.11	NA
RMD, decreasing trend	0	0	0	0.06	0.83	0.11	0.10	0.10	0.10	0.16	0.40	0.15	3.4(2.0)
RMD, constant trend	0	0	0	0.07	0.87	0.06	0.09	0.11	0.10	0.17	0.43	0.09	3.1(2.2)
RMD, increasing trend	0	0	0	0.07	0.87	0.06	0.10	0.10	0.10	0.18	0.43	0.09	2.7(2.1)
True MTD = d₆													
QLCRM	0	0	0	0.06	0.16	0.78	0.06	0.06	0.06	0.12	0.26	0.44	NA
RMD, cycle 1	0	0	0	0.05	0.17	0.78	0.06	0.06	0.06	0.16	0.26	0.40	NA
RMD, decreasing trend	0	0	0	0.04	0.15	0.81	0.08	0.08	0.10	0.16	0.20	0.38	3.6(2.3)
RMD, constant trend	0	0	0	0.01	0.13	0.86	0.06	0.06	0.10	0.14	0.22	0.42	3.3(2.2)
RMD, increasing trend	0	0	0	0.02	0.13	0.85	0.08	0.08	0.08	0.12	0.22	0.42	2.9(2.1)

Results at the target dose are in bold. d_k represents dose level k ($k=1,\dots,6$).

QLCRM: quasi-likelihood continual reassessment method, RMD: repeated measures design, MTD: maximum tolerated dose.

Table 3

Percentage of trials where a time effect was detected, with correlated longitudinal toxicity data.

Scenario	Correlated data ^a	
	RMD	RMD with a stronger time effect ^b
True MTD = d₂		
No time effect	13	13
Positive effect	67	89
Negative effect	14	74
True MTD = d₃		
No time effect	13	13
Positive effect	65	94
Negative effect	19	70
True MTD = d₄		
No time effect	7	7
Positive effect	64	88
Negative effect	12	56
True MTD = d₅		
No time effect	6	6
Positive effect	56	91
Negative effect	9	72

Data are presented in percentage (%). RMD: repeated measures design, MTD: maximum tolerated dose, d_k represents dose level k ($k=1,\dots,6$).

^aWithin-patient correlated toxicity data over cycles were generated by the Baseline-Categorical Logit model (Eq.2) with the parameter value of σ_w^2 set equal to 10.

^bThe slope of time effect was increased by 95–126% for positive trend, and 72–110% for negative trend. Please see the online Appendix for details on how to generate ordinal toxicity data with a stronger time effect.

Observed gastrointestinal, liver and hematological toxicities, and nTTP scores per cycle in the romidepsin trial.

Table 4

Cycle	1	2	3	4	5	6
Patients (N)	28	24	14	12	11	11
GI, liver, heme tox						
# of G2	38	33	26	14	16	20
# of G3+ non-heme	10	1	0	5	1	1
# of G4 heme	0	0	1	0	2	1
Observed nTTP	0.25	0.28	0.25	0.36	0.37	0.35

GI: gastrointestinal