Nautilus Spiral_4592. Photograph courtesy of Henry Domke, MD. www.henrydomke.com

Innovative trial designs addressing the limitations of traditional dose-escalation methods have yet to establish their clinical superiority in the phase 1 trial setting.

Phase 1 Trial Design: Is 3 + 3 the Best?

Aaron R. Hansen, MBBS, Donna M. Graham, MBBCh, Gregory R. Pond, PhD, and Lillian L. Siu, MD

Background: *Concerns have been recognized about the operating characteristics of the standard 3 + 3 dose-escalation design. Various innovative phase 1 trial designs have been proposed to address the issues and new challenges posed by molecularly targeted agents. However, in spite of these proposals, the conventional design is still the most widely utilized.*

Methods: *A review of the literature of phase 1 trials and relevant statistical studies was performed.*

Results: *Beyond statistical simulations, sparse clinical data exist to support or refute many of the shortcomings ascribed to the 3 + 3 rule method. Data from phase 1 trials demonstrate that traditional designs identified the correct dose and relevant toxicities with an acceptable level of precision in some instances; however, no single escalation method was proven superior in all circumstances.*

Conclusions: *Design selection should be guided by the principle of slow escalation in the face of toxicity and rapid dose increases in the setting of minimal or no adverse events. When the toxicity of a drug is uncertain or a narrow therapeutic window is suggested from preclinical testing, then a conservative 3 + 3 method is generally appropriate. However, if the therapeutic window is wide and the expected toxicity is low, then rapid escalation with a novel rule- or model-based design should be employed.*

Introduction

The primary objective of a phase 1 oncology trial is to define the recommended phase 2 dose (RP2D) of a new drug or multiagent combination in the sched-

From the Princess Margaret Cancer Centre, University Health Network, Division of Medical Oncology and Hematology (ARH, DMG, LLS) and the Department of Medicine, University of Toronto, Ontario, Canada (ARH, DMG, LLS), the Department of Oncology, McMaster University (GRP), and the Ontario Clinical Oncology Group, Hamilton, Ontario, Canada (GRP).

Submitted December 18, 2013; accepted January 24, 2014.

Address correspondence to Lillian L. Siu, MD, Drug Development Program, Princess Margaret Cancer Centre, 610 University Avenue, Suite 5-718, Toronto, Ontario, Canada M5G 2M9. E-mail: lillian.siu@uhn.ca

No significant relationships exist between the authors and the companies/organizations whose products or services may be referenced in this article.

ule tested. Although many crucial components make up a phase 1 study, this article will focus on various dose-escalation methods that can be incorporated into such trials. The conventional 3 + 3 design, as described by Storer[1] in 1989, was originally introduced in the 1940s[2] and was among the earliest dose-escalation and de-escalation schemes utilized. However, several concerns have been raised about the quality of the operating characteristics of the 3 + 3 design. Statistical simulations have demonstrated that a trial using the 3 + 3 design identifies the maximum tolerated dose (MTD) in as few as 30% of trials.[3] Furthermore, some argue that this method of dose escalation may result in a high proportion of patients being treated at subtherapeutic doses.[4]

Innovative trial designs that offer potentially more superior operating characteristics have been

proposed. However, despite these advances, the 3 + 3 dose-escalation method remains the most popular method employed by researchers of phase 1 trials. In a review of more than 1,200 phase 1 studies from 1991 to 2006, more than 98% of trials utilized the 3 + 3 dose escalation scheme.[5] This figure was confirmed in a review of 181 phase 1 trials from January 2007 to December 2008, with more than 96% of trials using the traditional 3 + 3 design or a variation.[6]

Given the time frame of these reviews, most of the agents included in these 2 reports were cytotoxic chemotherapeutic agents,[5,6] with molecularly targeted agents (MTAs) comprising only 18% in the latter report.[6] MTAs are defined as anticancer agents that selectively target molecular pathways, as opposed to DNA, tubulin, or cell division machinery.[7] The recent surge in the development of MTAs has challenged early-phase trial design. Because MTAs were purported to have a more specific therapeutic index on tumor tissue while sparing normal tissue, it was believed that studies of MTAs would have resulted in a shift toward newer methods with less conservative dose escalations. However, in a review of 155 MTA phase 1 trials published between 2000 and 2010, more than 60% of them incorporated the conventional 3 + 3 dose escalation.[8] Logistical simplicity of the 3 + 3 design and clinician familiarity with the escalation rules may explain this observation and preclude the transition to novel designs on a larger scale. Moreover, it is not clear whether the advantages offered by newer methodologies are negated by their complexity and difficulty to implement. This article will review the current phase 1 trial design landscape and evaluate which dose-escalation methods are optimal for determining dose and safety in an efficient manner, in addition to addressing several challenges faced by modern phase 1 trials.

Dose-Escalation Designs

Phase 1 trials must prioritize safety while attempting to maintain efficiency. A typical dose-escalation phase 1 study selects a safe starting dose based on preclinical data from in vitro and in vivo testing of the drug. Incremental dose increases for assigned patient cohorts occur until a prespecified end point is reached, which, in general, is the incidence of dose-limiting toxicities (DLTs). DLTs are severe but ideally reversible adverse events that occur within a protocol-defined period, usually the first cycle. Geographical variations in nomenclature for the final dose levels in phase 1 trials can be confusing. In Europe, the MTD is defined as the dose level in which an allowable frequency of DLTs has been exceeded, and the dose level immediately below is usually expanded to confirm that its incidence of DLTs is within an acceptable threshold.[9] If such is the case, then that penultimate dose level is considered

the RP2D. In North America, the highest dose level reached, in most instances due to an unacceptable incidence of DLTs, is referred to as the maximum administered dose (MAD). The MTD is typically defined as the dose level immediately below the MAD and corresponds with the RP2D.[9] In the event that the MTD is not reached, other non–toxicity-based end points can be considered in order to recommend a dose, such as a pharmacodynamic (PD) marker of sufficient pathway inhibition of a putative molecular target.

Phase 1 trial designs are broadly divided into rule- or model-based methods. Rule-based methods utilize prespecified rules based on actual observations of target events (eg, DLT) from the clinical data to assign patients to dose levels and determine the MTD or RP2D. Model-based designs use a statistical estimation of the target toxicity level by depicting the dose-toxicity relationship. The model is used to assign patients to dose levels and define the MTD or RP2D. Safeguards are established in most model-based designs to limit escalation above the MTD and patient exposure to excessively toxic treatment doses.[4] The Table[2,4,10-18] outlines selected examples of these phase 1 trial designs.

The most commonly used rule-based design is the traditional 3 + 3 design, which guides "up-and-down" decisions, using the modified Fibonacci mathematical series to determine the amount of dose increase for cohorts of sequentially enrolled patients. The Fibonacci series ensures that dose increases are initially large but increments are smaller at higher dose levels (Fig 1A).[2] Newer trial designs have explored modifications of the process of dose escalation using statistical and empirical methods.[7]

In an effort to expedite the dose-escalation process to more rapidly and efficiently determine the RP2D, accelerated titration designs have been proposed (Fig 1B).[4] Although they are not commonly used in clinical practice, these designs use single-patient cohorts for initial dose escalation until DLT or 2 moderate toxicities occur during the first treatment cycle. Following the initial acceleration phase, the traditional 3 + 3 design is applied, which theoretically reduces the number of patients treated at subtherapeutic doses. In this design, intrapatient dose escalation is permitted if no or minimal toxicities are observed.[4] Accelerated titration designs are still rule based, but they can incorporate model-based designs following the initial single-patient cohorts to create hybrid methods.

Interests in exploring nontoxicity end points have been fuelled by the observation that MTAs do not have the classical monotonic dose-toxicity curve as seen with cytotoxic chemotherapy. Therefore, escalating doses solely based on toxicity may not be the most appropriate strategy, and using alternative end points such as pharmacokinetic (PK) or PD data may be more appropriate for determining the optimal dose of these

agents. Robust preclinical data to support the use of these end points are imperative if such methodology is to be used. Pharmacokinetically guided dose escalation (PGDE) involves regular PK assessments to determine subsequent dose modifications (Fig 1C).[10] The main practical limiting factors of PGDE include interpatient variability and the need for timely PK analysis.[10] Where a PD end point of target inhibition is used, there must be a predefined target, available tumor tissue for testing, and a validated assay to determine and quantitate the extent of target inhibition. These are difficult criteria to meet, and the ability to reliably correlate PD readouts with clinical outcomes based on limited data further challenges the use of this strategy.

Model-based designs were developed to improve precision in estimating the RP2D as well as efficiency during dose escalation. Model-based designs establish a dose-toxicity curve prior to patient enrollment and then use toxicity data from enrolled patients and Bayesian statistical methods to modify and update this curve as the study proceeds. The first model-based method was the continual reassessment method (CRM).[2] Modifications to this design have been developed with the aim of improving safety and avoiding an overestimation of the MTD. These include modified CRM,[11,12] escalation with overdose control (Fig 1D),[13] and the time-to-event continual reassessment method (TITE-CRM).[14] It has been suggested that model-

Table. — Selected Examples of Trial Escalation Designs

Trial Design	Method	Comments
Rule-Based Designs		
Key Points		
No prior assumptions about dose-toxicity curve		
Decision to escalate based on toxicity results from first course administration of current level		
3 + 3[2] (Fig 1A)	3 patients treated per dose level If no DLT, dose is escalated for the next cohort of 3 patients If 1 DLT, 3 additional patients are treated at this level with dose escalation only if no additional DLTs If ≥ 2 DLTs, prior dose level is defined as MTD MTD decided when 6 patients are treated at a dose level with < 2 DLTs	Potentially more patients are treated at subtherapeutic doses Statistical simulations suggest RP2D often at lower doses than other designs

Accelerated titration[4] (Fig 1B)	A series of designs have been proposed. All have fixed increments for dose escalation:
	  Design 1 is as for 3 + 3 design but with 40% dose increments
	  Design 2 has single patient cohorts during accelerated phase. When a first-course DLT or second first-course intermediate toxicity occurs, cohort expands and reverts to design 1
	  Design 3 has single patient cohorts with double-dose escalation steps (80% dose increments). Revert to design 1 with same trigger as design 2
	  Design 4 is per design 3 but trigger to revert to design 1 is any course DLT or second instance of any course intermediate toxicity
		Acceleration and escalation phase in one design
		Intrapatient dose escalation, where permitted, may mask delayed or cumulative toxicity

PGDE		
Key Points		
Requires real-time PK measurement and assessment for dose modification		
Assumes DLT can be predicted by plasma drug concentration		

PGDE[10] (Fig 1C)	Starting dose is determined by animal data as standard
	First cohort treated as standard with average AUC measured for this cohort
	Dose escalation occurs according to distance to target AUC either:
	  Initially by a factor equal to the square root of the ratio of the target AUC to the AUC associated with the initial dose and subsequently follow a Fibonacci scheme; OR
	  Initially by a factor of 2 until AUC is 40% of target AUC and subsequently follows Fibonacci scheme
		Interpatient variability may limit dose escalation
		Differences between species used for dose estimation and humans may affect utility of method

Table. — Selected Examples of Trial Escalation Designs (continued)

Trial Design	Method	Comments
Model-Based Designs		
Key Points Establishes dose-toxicity curve prior to patient enrollment Uses toxicity data from enrolled patients to modify curve as study proceeds Requires good biostatistical support for constructing and updating dose-toxicity estimates		
CRM[17]	A target level of toxicity defined at baseline. Increased dose levels defined with initial expectations of the probability of DLT at these doses to construct a statistical dose-toxicity model Single patient cohorts; fixed sample size With treatment of successive patients the statistical model is recalculated using Bayesian principles to update estimated probability of a DLT and increase certainty associated with dose-toxicity relationship Dose associated with the target DLT rate according to the final dose-toxicity model at trial completion is defined as the MTD If a patient experiences no toxicity, dose may be escalated to the next dose level for subsequent cycles	May overestimate dose for MTD Uncertainty about toxicity of investigational agent may be reflected in initial dose-toxicity model
Modified CRM[11,12]	Conservative starting dose, as with 3 + 3 design Dose escalation may only occur by a single dose level per patient cohort Following DLT, the dose for the next patient may not be escalated Cohorts may be larger than single patient Stopping rule defined rather than fixed sample size	Safety improved compared with CRM
EWOC[13] (Fig 1D)	Dose-toxicity curve modeled to minimize the probability a patient will be treated at an unacceptably high dose, ie, a dose where the probability of a DLT is greater than some value	Dose-toxicity curve constantly remodeled requiring additional statistical support
TITE-CRM[14]	Data from all treated patients, including partial data, incorporated into dose-toxicity curve and subsequent dose calculations Patients experiencing DLT are fully weighted Patients not experiencing toxicity are weighted by the proportion of time observed on study	Allows toxicity information of patients to contribute to dose recommendation before all patients are fully followed
mTPI[18]	Bayesian framework used to calculate posterior probabilities of intervals, reflecting relative distance between toxicity rate of each dose level to the target probability with a fixed sample size Dose-escalation decisions depend on category of toxicity rate: categorization as underdosing results in dose escalation and overdosing results in de-escalation When toxicity rate is categorized as proper dose, dose is not modified	Software provided online Fewer patients treated at doses above MTD
Mixed-effect POM[15]	Repeated measurements of graded toxicities are used to generate per-cycle toxicity estimates Modified definition of RP2D: dose associated with a predefined probability of severe toxicity per cycle	May be more useful for agents with chronic toxicities (eg, MTAs)
Fractional dose-finding methods[16]	May be applied to 3 + 3 design or CRM Time to toxicity modeled using Kaplan-Meier estimator. Mass of each censored observation redistributed to the right Censored observation is taken if patient has not experienced a DLT at the time of observation. Fractional contribution of each patient used in dose-escalation decisions	May shorten duration of trial while maintaining accurate determination of RP2D

AUC = area under the curve, CRM = continuous reassessment method, DLT = dose-limiting toxicity, EWOC = escalation with overdose control, MTA = molecularly targeted agent, MTD = maximum tolerated dose, mTPI = modified toxicity probability interval, PGDE = pharmacologically guided dose escalation, PK = pharmacokinetic, POM = proportional odds model, RP2D = recommended phase 2 dose, TITE = time to event.

based designs achieve better estimations of the target probability of a DLT at the RP2D while minimizing suboptimal dosing.[19] These designs require biostatistics expertise and adequate software for modeling.

By contrast to cytotoxic chemotherapy, which is generally delivered intermittently and for a predefined duration, MTAs are often continuously dosed until the development of resistance and, thus, chronic or late toxicity can emerge. Recently, model-based methods have been employed to incorporate these factors into dose-escalation decisions. The modified toxicity probability interval,[20] proportional odds model (POM),[21] mixed-effect POM,[15] and fractional dose-finding methods[16] utilize toxicity information from treated patients, including those who have not experienced a DLT at the time of observation, to accurately reflect the ongoing effects of the agents under investigation.

Phase 1 Design Elements

Critical aspects of a phase 1 study include the accurate determination of RP2D, a comprehensive assessment of toxicity, and efficiency in time to study completion, patient accrual, and logistical requirements. Study designs that can precisely determine RP2D and drug toxicity profile, shorten trial duration, subject fewer patients to subtherapeutic or overly toxic doses, and place fewer demands on clinical and trial resources are preferred.

Determining the Recommended Dose

Beyond statistical simulations, sparse clinical data exist that compare the accuracy of different dose-escalation schemes in predicting the correct RP2D. In a review from 1990 to 2012 of registration trials of cancer drugs approved by the US Food and
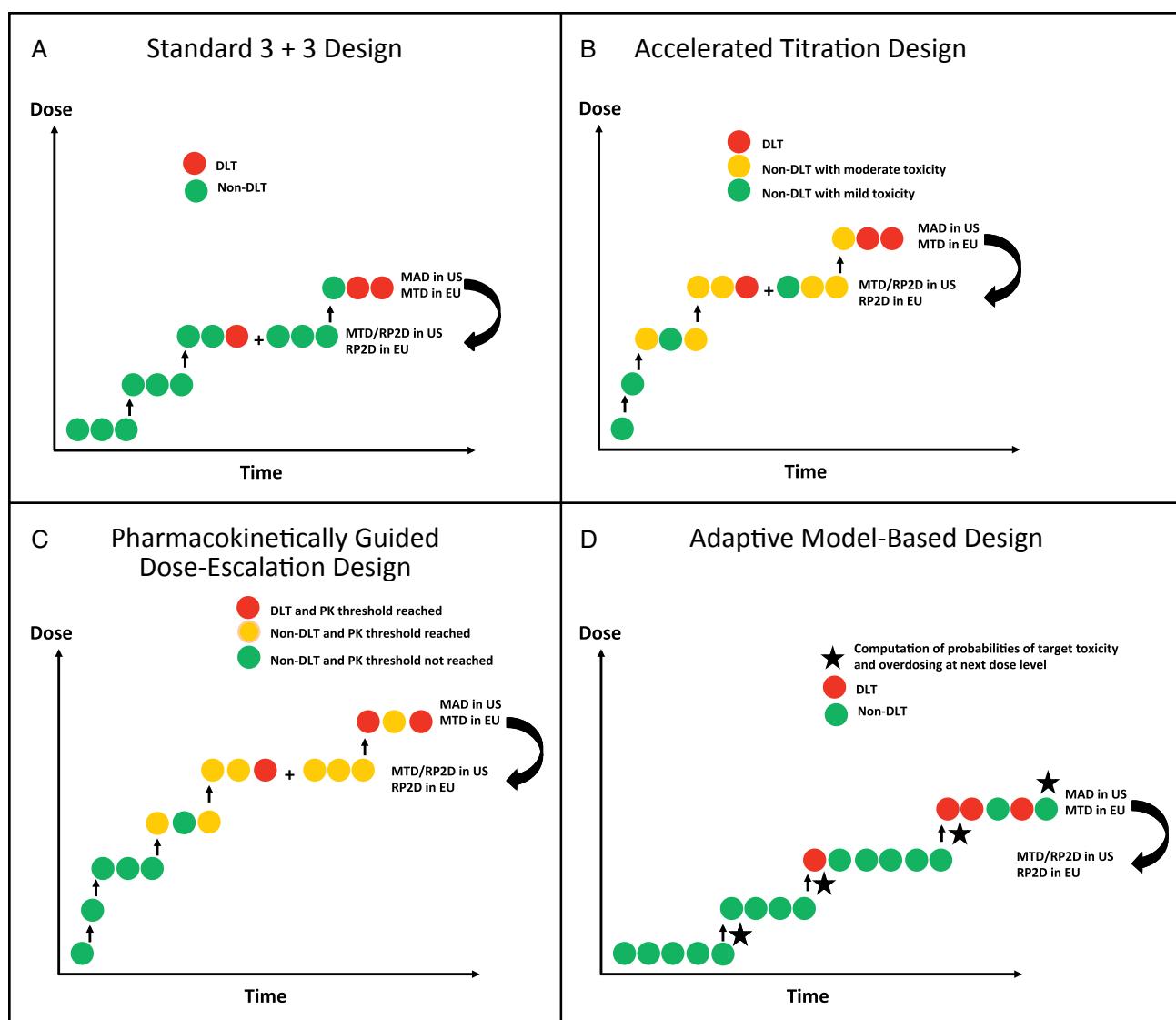

Fig 1. — (A) Schematic of the standard 3 + 3 design. (B) Schematic of the rule-based accelerated titration design. (C) Schematic of the pharmacokinetically guided dose-escalation design. (D) Schematic of an adaptive model-based design (eg, escalation with overdose control). DLT = dose-limiting toxicity, EU = European Union, MAD = maximum administered dose, MTD = maximum tolerated dose, PK = pharmacokinetics, RP2D = recommended phase 2 dose, US = United States.

Drug Administration and their corresponding phase 1 studies, the registered doses of the drugs were within 20% of the RP2D in 73% (62 of 85) of the matched phase 1 trials.[22] An important caveat reported in this study was that phase 1 trials of MTAs performed poorly in predicting the registered dose to within 20% of the RP2D when compared with phase 1 trials of cytotoxic treatments (odds ratio = 0.2 [0.03–0.8], P = .025).[22] The type of dose escalation employed in these phase 1 studies was not reported, and uncertainty remains as to whether the registered dose is the optimal dose of drug. Furthermore, drugs that failed to achieve registration were not included; thus, their impact on the results is unknown.

The broad implication of this study is that, overall, phase 1 trials are reasonably accurate in predicting the correct RP2D in most — but not all — cases. Given that most phase 1 trials utilize the 3 + 3 design, it is plausible to extrapolate that these figures reflect the accuracy of this type of methodology in most instances. However, the lack of trial design details from this review renders it difficult to draw definitive conclusions. These empirical data contrast with the many statistical simulations that report the inaccuracy of the 3 + 3 method compared with novel rule- or model-based strategies.[11,17,23-26] In a recent comparative simulation study of the modified toxicity probability interval and the 3 + 3 designs using matched sample sizes, the former was found to reliably select the true MTD in 32 out of 42 scenarios.[18] The reason for the discrepancy between empirical and statistical simulation data is unclear, but many factors can influence the RP2D that may not be incorporated into the simulation; in addition, computational modeling has its own set of limitations.

Rather than ascribing a particular escalation scheme as the most appropriate for all situations, it is more informative for clinical trialists to determine the contexts in which specific methods perform better or are more fit-for-purpose. A key principle of any phase 1 study is to escalate slowly in the face of toxicity and rapidly in the setting of minimal or no adverse events. In situations in which the toxicity of a drug is uncertain or preclinical testing suggests that the therapeutic window is narrow, then a conservative 3 + 3 escalation may be reasonable. By contrast, rapid escalation designs using either novel rule- or model-based schemes would be appropriate if the expected toxicity was low or high-quality preclinical data indicated a wide therapeutic index. Failing to accurately define the true RP2D of a drug may compromise further drug development by testing a subtherapeutic dose and schedule in phase 2 or 3 trials. Agents not successfully registered may be either intrinsically ineffective or not evaluated at the correct RP2D.

Determining the Relevant Toxicities

A critical assumption with the early-phase drug development of cytotoxic agents is that toxicity and efficacy monotonically increase with dose. In the era of MTAs, this assumption is challenged by the premise of differential target saturation of the drug in tumor versus normal tissue, whereby target blockade can become saturated in tumor, thus leading to a plateau of antitumor effect. However, toxicity may continue to rise with an increasing drug dose due to unsaturated targets in normal tissue. In addition, off-target effects may occur such that dosing to toxicity for MTAs may not be appropriate. Despite this, determining the toxicity level remains a critical goal in all phase 1 trials, regardless of the dose-escalation method. It has been reported that around 70% of all clinically significant adverse events are identified in the phase 1 study.[22]

The probability of observing an adverse event related to a study drug is increased by using a larger sample size. A trial has a 90% chance of detecting an adverse event occurring in 5% of patients with a sample size of 57 patients, a rate that increases to 99% with 82 patients.[27] Increasing the sample size of a trial can be achieved by using prespecified expansion cohorts after the dose-escalation component of the trial is completed. Often referred to as a phase 1b or dose-expansion phase study, this part of the trial typically tests a particular dose and schedule, which may be restricted to a population of patients with specific tumor types, molecular aberrations, or some other criteria.[28]

Study Size, Timing, and Other Logistics

Several factors can affect the efficiency of a phase 1 trial (eg, number of patients enrolled and the fraction of those treated at subtherapeutic doses, trial duration, number of dose levels tested). Statistical simulations have attempted to assess the efficiency of various dose-escalation designs.[28,29] In a comparative simulation of 3 + 3 and CRMs, the latter treated fewer patients at subtherapeutic doses and, depending on the true MTD, required a smaller sample size than the 3 + 3 design.[31] However, real-world experience does not appear to demonstrate substantial differences in efficiency across dose-escalation methods, possibly because of neutralization by other factors. In a retrospective review of 81 trials evaluating 60 MTAs, the ratio of MTD or MAD to starting dose was calculated for each trial and compared with the number of dose levels tested.[32] Trials with a high MTD or MAD to starting dose ratio (presumably representing trials of agents with wider therapeutic indices) versus those with a low ratio (presumably representing trials of agents with narrower therapeutic indices) had similar numbers of tested dose levels. This phenomenon was not explained by the use of different dose-escalation

methods. Instead, it appears that higher increments between dose levels were selected for agents deemed to have wide therapeutic indices based on preclinical data, regardless of the dose-escalation method used.[32] Furthermore, in a review of 84 phase 1 trials of MTAs from 2000 to 2010, trial size and the number of patients treated at doses below MTD were similar across 3 + 3, accelerated titration, and CRM designs.[33]

One criticism of the conventional 3 + 3 design is that the escalation is unnecessarily slow, thus prolonging the duration of the trial.[4] Similar complaints have been leveled at the CRM, because results from the last patient are required to dose subsequent patients.[11] Modifications to CRM and accelerated titration designs have sought to address these concerns, and simulations have demonstrated that they speed up the completion of a phase 1 study but also provide more information.[4] However, clinical data to substantiate these simulations are sparse.

It is important to recognize that the 3 + 3 method does not require statistical software support or modeling, and it is widely acknowledged as being uncomplicated and safe to implement.[6] Processes underlying model-based designs are often perceived by clinicians as being nontransparent.[34] Furthermore, the software that these systems utilize is often considered difficult for those who are not technically inclined.[34] Although adaptations to these models have been suggested, such logistical barriers may limit their widespread implementation in phase 1 studies.

Challenges to Phase 1 Designs
Delayed or Chronic Toxicities
Many MTAs or biological agents are chronically administered as either continuous oral or frequent intravenous doses, thus predisposing patients to delayed or chronic toxicities. These cumulative adverse events can affect the tolerability of the drug and should be considered when recommending a phase 2 dose. Assessing for DLTs in the initial cycle or first few weeks of treatment will not account for late or cumulative adverse events. In a retrospective review of 36 phase 1 trials of MTAs, approximately one-half of severe toxicities occur after cycle 1.[35] In an initiative to redefine criteria for determining the DLT, adverse events from more than 2,000 patients, representing 54 monotherapy MTA studies, concluded that 49.5% of patients experienced their first grade 3 or higher toxicity after cycle 1, and a significant proportion of patients required a dose reduction for selected grade 2 or lower events as early as cycle 1.[36]

Several designs have been proposed to account for chronic adverse events, including TITE-CRM, mixed-effect POM, and fractional dose-finding methods.[15,16] These latter designs have been tested in statistical simulations but have not been widely applied in

clinical trials. In studies of drugs with delayed toxicities, the TITE-CRM and 3 + 3 design were compared utilizing the Monte Carlo simulation.[37] The former was reported to result in trials of shorter length and a higher number of patients treated at or around the MTD. When the TITE-CRM was incorporated into 2 phase 1 studies of concurrent chemoradiotherapy for pancreatic cancer, these trials failed to reproduce the predicted methodological benefits, supporting the notion that all designs must be implemented into actual trials before their alleged advantages can be accepted.[38,39]

Nontoxicity End Points and Optimal Biological Dosing
Some MTAs have different dose–effect relationships than traditional cytotoxics; therefore, dosing these agents to toxicity may not be appropriate. In a review of 24 phase 1 trials of MTAs, 683 patients were retrospectively assigned to a cohort based on a comparative dose to the MTD (low [≤ 25% MTD], medium [25%–75% MTD], or high [≥ 75% MTD]), and their outcomes were compared.[40] No significant difference in response rate or survival was seen across the cohorts, implying that these agents have a nonmonotonic dose-efficacy curve. Under this assumption, escalating doses until predefined PK or PD parameters demonstrate when a target is saturated, or when a biological pathway is maximally altered, would be an alternative to the traditional toxicity end point. This type of strategy has been termed optimal biological dosing (OBD) and is defined as the dose that produces the most favorable prespecified effect on a biomarker. However, toxicity remains the predominant end point as evidenced by several reviews of phase 1 trials, which have reported that the proportion of studies using nontoxicity end points ranges from 24% to 48%.[6,7,22]

Because the 3 + 3 design typically uses toxicity to guide dose escalation, it is unlikely that such a design would be optimized to ascertain the OBD. The use of real-time PK monitoring in the PGDE method makes it an appealing strategy to determine the OBD. Several adaptive dose-finding designs have been proposed to identify the OBD. The nonparametric and semi-parametric methods have been proven to have the best operating characteristics via computational testing and are recommended for use.[41] However, to date, no trial reporting an OBD has utilized either adaptive methodology. Given the enthusiasm surrounding OBD, it is anticipated that more trials utilizing novel trial designs will emerge, providing a deeper understanding of which escalation methods are the most appropriate.[42]

Combination Therapy
The development of treatment resistance has long been recognized with cytotoxic therapies. To circumvent this phenomenon, combination chemotherapy regimens

with non–cross-reactive compounds, administered either concurrently or sequentially, were tested.[43] This rationale also forms the basis for testing MTA combinations. Several unbiased genome sequencing studies have shown that many tumors have a complex molecular profile of low frequency mutations.[44] Single-agent MTAs have a limited benefit in patients with advanced cancer, with progressive disease eventually developing in all patients, even among those with significant or prolonged responses.[45] Complex tumor biology implicates the involvement of multiple interconnected signalling networks that drive cancer initiation and metastasis.[46,47] Thus, rationally designed combination MTA therapy is an attractive strategy to prevent — or, at the very least, delay — the emergence of treatment resistance by perturbing multiple cellular networks, potentially thwarting the development of molecular escape pathways.[45] Major concerns in administering these combination regimens include the cumulative impact of overlapping toxicities, the potentiation of adverse events due to drug–drug interactions, and the occurrence of unexpected adverse events.[48,49]

As with delayed toxicities and OBD, no single dose-escalation method has proven to be superior for evaluating combination regimens in clinical testing. Combination studies are more complex because a number of schedules may be recommended, and different schedules can produce the same level of toxicity. These types of trials often have RP2D and toxicity information from the monotherapy phase 1 trials of each agent that can be utilized by either rule- or model-based designs, possibly improving the efficiency of the trial. However, preclinical data describing PK or PD interactions or potential synergy with regard to antitumor activity or toxicity may be lacking. Given that the MTD of each drug in the combination is known, it is unlikely that either rule- or model-based designs will escalate patients to overly toxic dose levels. The challenge facing both types of designs is the optimal method by which to escalate each drug. Careful selection of starting doses and subsequent dose levels is crucial, regardless of the escalation method. In rule-based methods, the agents can be escalated to prespecified doses either sequentially, in parallel, or with one drug fixed at either a high or low dose while the other is increased toward the RP2D. The zone method illustrates a rule-based escalation method in which consecutive dosing zones are comprised of sequentially increasing dose combinations and patients are randomized to these preset doses within an individual zone and escalation decisions are made using a modified 3 + 3 method.[50] These dose levels are then compared and the combinations that are either too toxic or unlikely to be efficacious are eliminated. Statistical simulations have reported that this design requires a smaller sample size, has better

power, and patients are more likely to be treated at efficacious doses compared with conventional escalation methods.[50] Various Bayesian parametric models have been designed and they purport to recommend a more accurate dose and schedule for combination regimens.[51-53] As with monotherapy trials, the escalation method selected for these types of studies should be individualized and guided by the anticipated level of toxicity from the combination.

Conclusion

This article is neither a defense nor homage to the standard 3 + 3 design, but rather it is a review of the clinical evidence. The conventional escalation method is slow, inaccurate in recommending doses, and enrolls a significant proportion of patients at subtherapeutic doses. However, sparse empirical data exist to support or refute the assertion that any one trial design is superior to another. The evidence from this review suggests that the 3 + 3 design identifies the recommended phase 2 dose and toxicities with an acceptable level of precision in some circumstances, while being straightforward to operate with few logistical demands. Newer escalation methods have had limited incorporation into the clinical trial setting and, consequently, have had little opportunity to demonstrate favorable clinical application beyond statistical simulations. Although such designs may have clinical utility, simulations performed under ideal circumstances may not reflect the true clinical reality in which, for example, protocol deviations for dose escalations are permitted due to specific patient-, institutional-, or treatment-related factors. Thus, novel trial designs demonstrating superiority over the 3 + 3 method in statistical simulations without corroborating clinical evidence are of theoretical value alone.

Following the explosion of molecularly targeted agents under investigation, delayed toxicities, optimal dosing, and combination treatments have presented challenges to the design of phase 1 trials. Time-to-event designs may be the most suitable for testing agents with delayed or chronic toxicities; however, recently proposed designs, such as the mixed-effect proportional odds model, may prove appropriate with clinical experience. The pharmacokinetically guided dose-escalation scheme is a strategy that may account for optimal biological dosing, although the ability to confidently define optimal biological activity in the phase 1 setting remains elusive. A study of an investigational drug or combination with an uncertain or high probability of producing serious adverse events should adopt a conservative escalation scheme such as the 3 + 3 design. By contrast, studies of agents predicted to have less toxicity can be quickly escalated with either an accelerated titration- or model-based design until an acceptable toxicity range is reached. Regardless of

the type of agent or setting, the guiding principles of safe starting dose selection, minimizing the number of patients treated at subtherapeutic doses, rapid escalation in the absence of toxicity, and slow escalation in the presence of adverse events will ensure the sound design of any phase 1 trial. Therefore, although the 3 + 3 design may not be the best in all circumstances, in this context it should not be abandoned and may still have a place in the design of phase 1 studies.

References

1. Storer BE. Design and analysis of phase I clinical trials. *Biometrics.* 1989;45(3):925-937.
2. Dixon WJ, Mood AM. A method for obtaining and analyzing sensitivity data. *J Am Stat Assoc.* 1948;43(241):109-126.
3. Reiner E, Paoletti X, O'Quigley J. Operating characteristics of the standard phase I clinical trial design. *Comput Stat Data Anal.* 1999;30(3):303-315.
4. Simon R, Rubinstein L, Arbuck SG, et al. Accelerated titration designs for phase I clinical trials in oncology. *J Natl Cancer Inst.* 1997;89(15):1138-1147.
5. Rogatko A, Schoeneck S, Jonas W, et al. Translation of innovative designs into phase I trials. *J Clin Oncol.* 2007 2007;25(31):4982-4986.
6. Le Tourneau C, Lee JJ, Siu LL. Dose escalation methods in phase I cancer clinical trials. *J Natl Cancer Inst.* 2009;101(10):708-720.
7. Parulekar WR, Eisenhauer EA. Phase I trial design for solid tumor studies of targeted, non-cytotoxic agents: theory and practice. *J Natl Cancer Inst.* 2004;96(13):990-997.
8. Le Tourneau C, Razak AR, Gan HK, et al. Heterogeneity in the definition of dose-limiting toxicity in phase I cancer clinical trials of molecularly targeted agents: a review of the literature. *Eur J Cancer.* 2011;47(10):1468-1475.
9. Eisenhauer EA, Twelves C, Buyse ME. *Phase 1 Cancer Clinical Trials: A Practical Guide.* New York: Oxford University Press; 2006.
10. Collins JM, Zaharko DS, Dedrick RL, et al. Potential roles for preclinical pharmacology in phase I clinical trials. *Cancer Treat Rep.* 1986;70(1):73-80.
11. Goodman SN, Zahurak ML, Piantadosi S. Some practical improvements in the continual reassessment method for phase I studies. *Stat Med.* 1995;14(11):1149-1161.
12. Faries D. Practical modifications of the continual reassessment method for phase I cancer clinical trials. *J Biopharm Stat.* 1994;4(2):147-164.
13. Babb J, Rogatko A, Zacks S. Cancer phase I clinical trials: efficient dose escalation with overdose control. *Stat Med.* 1998;17(10):1103-1120.
14. Cheung YK, Chappell R. Sequential designs for phase I clinical trials with late-onset toxicities. *Biometrics.* 2000;56(4):1177-1182.
15. Doussau A, Thiébaut R, Paoletti X. Dose-finding design using mixed-effect proportional odds model for longitudinal graded toxicity data in phase I oncology clinical trials. *Stat Med.* 2013;32(30):5430-5447.
16. Yin G, Zheng S, Xu J. Fractional dose-finding methods with late-onset toxicity in phase I clinical trials. *J Biopharm Stat.* 2013;23(4):856-870.
17. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase I clinical trials in cancer. *Biometrics.* 1990;46(1):33-48.
18. Ji Y, Wang S-J. Modified toxicity probability interval design: a safer and more reliable method than the 3 + 3 design for practical phase I trials. *J Clin Oncol.* 2013;31(14):1785-1791.
19. Mick R, Ratain MJ. Model-guided determination of maximum tolerated dose in phase I clinical trials: evidence for increased precision. *J Natl Cancer Inst.* 1993;85(3):217-223.
20. Westblade LF, Jennemann R, Branda JA, et al. Multicenter study evaluating the Vitek MS system for identification of medically important yeasts. *J Clin Microbiol.* 2013;51(7):2267-2272.
21. Van Meter EM, Garrett-Mayer E, Bandyopadhyay D. Proportional odds model for dose-finding clinical trial designs with ordinal toxicity grading. *Stat Med.* 2011;30(17):2070-2080.
22. Jardim DL, Hess KR, Lorusso PM, et al. Predictive value of phase I trials for safety in later trials and final approved dose: analysis of 61 approved cancer drugs. *Clin Cancer Res.* 2014;20(2):281-288.
23. O'Quigley J. Another look at two phase I clinical trial designs. *Stat Med.* 1999;18(20):2683-2692.
24. Korn EL, Midthune D, Chen TT, et al. A comparison of two phase I trial designs. *Stat Med.* 1994;13(18):1799-1806.
25. Neuenschwander B, Branson M, Gsponer T. Critical aspects of the Bayesian approach to phase I cancer trials. *Stat Med.* 2008;27(13):2420-2439.
26. Berry SM, Carlin BP, Lee JJ, et al. *Bayesian Adaptive Methods for Clinical Trials.* Boca Raton, FL: CRC Press; 2010.
27. DeMichele A, Berry DA, Zujewski J, et al. Developing safety criteria for introducing new agents into neoadjuvant trials. *Clin Cancer Res.* 2013;19(11):2817-2823.
28. Manji A, Brana I, Amir E, et al. Evolution of clinical trial design in early drug development: systematic review of expansion cohort use in single-agent phase I cancer trials. *J Clin Oncol.* 2013;31(33):4260-4267.

29. Eckhardt S, Siu L, Clark G, et al. The continual reassessment method (CRM) for dose escalation in phase I trials in San Antonio does not result in more rapid study completion. Paper presented at the Annual Meeting of the American Society for Clininical Oncology, Chicago, Illinois, 1999. Abstract 627.
30. Walling J, Zervos P, McCarthy S, et al. Dose escalation methodology in phase I clinical trials: a comparison of the modified continual reassessment method (MCRM) and a traditional method. Experience with the multitargeted antifolate (MTA). Paper presented at the Annual Meeting of the American Society for Clininical Oncology, Chicago, Illinois, 1997.
31. Iasonos A, Wilton AS, Riedel ER, et al. A comprehensive comparison of the continual reassessment method to the standard 3 + 3 dose escalation scheme in phase I dose-finding studies. *Clin Trials.* 2008;5(5):465-477.
32. Le Tourneau C, Stathis A, Vidal L, et al. Choice of starting dose for molecularly targeted agents evaluated in first-in-human phase I cancer clinical trials. *J Clin Oncol.* 2010;28(8):1401-1407.
33. Le Tourneau C, Gan HK, Razak AR, et al. Efficiency of new dose escalation designs in dose-finding phase I trials of molecularly targeted agents. *PloS One.* 2012;7(12):e51039.
34. Sweeting M, Mander A, Sabin T. Bcrm: Bayesian continual reassessment method designs for phase I dose-finding trials. *J Stat Software.* 2013;54(13). http://www.jstatsoft.org/v54/i13/paper. Accessed April 11, 2014.
35. 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(13):1728-1735.
36. Postel-Vinay S, Le Tourneau C, Olmos D, et al. Towards new methods for the determination of dose limiting toxicities and recommended dose of molecularly targeted agents. Presented at the European Cancer Congress (ECCO-ESMO-ESTRO), Amsterdam, The Netherlands, 2013.
37. Normolle D, Lawrence T. Designing dose-escalation trials with late-onset toxicities using the time-to-event continual reassessment method. *J Clin Oncol.* 2006;24(27):4426-4433.
38. Muler JH, McGinn CJ, Normolle D, et al. Phase I trial using a time-to-event continual reassessment strategy for dose escalation of cisplatin combined with gemcitabine and radiation therapy in pancreatic cancer. *J Clin Oncol.* 2004;22(2):238-243.
39. Desai SP, Ben-Josef E, Normolle DP, et al. Phase I study of oxaliplatin, full-dose gemcitabine, and concurrent radiation therapy in pancreatic cancer. *J Clin Oncol.* 2007;25(29):4587-4592.
40. Jain RK, Lee JJ, Hong D, et al. Phase I oncology studies: evidence that in the era of targeted therapies patients on lower doses do not fare worse. *Clin Cancer Res.* 2010;16(4):1289-1297.
41. Zang Y, Lee JJ, Yuan Y. Adaptive designs for identifying optimal biological dose for molecularly targeted agents. http://odin.mdacc.tmc.edu/~yyuan/Software_release/TargetAgent/targetAgentDF.pdf. Accessed April 16, 2014.
42. Kummar S, Gutierrez M, Doroshow JH, Murgo AJ. Drug development in oncology: classical cytotoxics and molecularly targeted agents. *Br J Clin Pharmacol.* 2006;62(1):15-26.
43. Goldie JH, Coldman AJ. A mathematic model for relating the drug sensitivity of tumors to their spontaneous mutation rate. *Cancer Treat Rep.* 1979;63(11-12):1727-1733.
44. Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer genome landscapes. *Science.* 2013;339(6127):1546-1558.
45. Dancey JE, Chen HX. Strategies for optimizing combinations of molecularly targeted anticancer agents. *Nat Rev Drug Discov.* 2006;5(8):649-659.
46. Fidler IJ, Gersten DM, Hart IR. The biology of cancer invasion and metastasis. *Adv Cancer Res.* 1978;28:149-250.
47. Hanahan D1, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646-674.
48. Dy GK, Adjei AA. Understanding, recognizing, and managing toxicities of targeted anticancer therapies. *CA Cancer J Clin.* 2013;63(4):249-279.
49. Park SR, Davis M, Doroshow JH, et al. Safety and feasibility of targeted agent combinations in solid tumours. *Nat Rev Clin Oncol.* 2013;10(3):154-168.
50. Huang X, Biswas S, Oki Y, et al. A parallel phase I/II clinical trial design for combination therapies. *Biometrics.* 2007;63(2):429-436.
51. Yin G, Li Y, Ji Y. Bayesian dose-finding in phase I/II clinical trials using toxicity and efficacy odds ratios. *Biometrics.* 2006;62(3):777-787.
52. Thall PF, Millikan RE, Mueller P, et al. Dose-finding with two agents in phase I oncology trials. *Biometrics.* 2003;59(3):487-496.
53. Yuan Y, Yin G. Sequential continual reassessment method for two-dimensional dose finding. *Stat Med.* 2008;27(27):5664-5678.