

Approaches to Phase 1 Clinical Trial Design Focused on Safety, Efficiency, and Selected Patient Populations: A Report from the Clinical Trial Design Task Force of the National Cancer Institute Investigational Drug Steering Committee

S. Percy Ivy¹, Lillian L. Siu², Elizabeth Garrett-Mayer³, and Larry Rubinstein¹

Abstract

The goals and objectives of phase 1 clinical trials are changing to include further evaluation of endpoints such as molecular targeted effects, in addition to dose-toxicity profile of the investigational agent. Because of these changes in focus, the National Cancer Institute and Investigational Drug Steering Committee's Task Force on Clinical Trial Design met to evaluate the most efficient ways to design and implement early clinical trials with novel therapeutics. Clinical approaches discussed included the conventional 3 + 3 cohort expansion phase 1 design, multi-institutional phase 1 studies, accelerated titration designs, continual reassessment methods, the study of specific target patient populations, and phase 0 studies. Each of these approaches uniquely contributes to some aspect of the phase 1 study, with all focused on dose and schedule determination, patient safety, and limited patient exposure to ineffective doses of investigational agent. The benefit of labor-intensive generation of preliminary biomarker evidence of target inhibition, as well as the value of molecular profiling of the study population, is considered. New drug development is expensive and the failure rate remains high. By identifying patient populations expected to respond to the study agent and tailoring the treatment with a novel drug, investigators will be one step closer to personalizing cancer treatment. The "fail early and fast" approach is acceptable if the appropriate patient population is evaluated in the phase 1 trial. The approaches outlined in this overview address the merits, advantages, disadvantages, and obstacles encountered during first in human studies.

Clin Cancer Res; 16(6); 1726–36. ©2010 AACR.

A workshop sponsored by The Clinical Trial Design Task Force of the Investigational Drug Steering Committee discussed the evolving role of the phase 1 clinical trial beyond the simple determination of dose, schedule, and adverse event profile. This report, generated following that workshop, provides a general overview of the designs, goals, and objectives for studies done for the first time in humans, focusing on traditional and adaptive designs, in addition to designs that limit the number of patients accrued at lower and presumably less effective dose levels.

Authors' Affiliations: ¹Investigational Drug Branch and Biometrics Research Branch, Cancer Therapy Evaluation Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, Maryland; ²Princess Margaret Hospital, Toronto, Canada; and ³Medical University of South Carolina, Charleston, South Carolina

Corresponding Author: S. Percy Ivy, Investigational Drug Branch, Cancer Therapy Evaluation Program, Division of Cancer Treatment and Diagnosis, 6130 Executive Blvd, Suite 7131, Rockville, MD 20852. Phone: 301-496-1196; Fax: 301-402-0428; E-mail: ivyp@ctep.nci.nih.gov.

doi: 10.1158/1078-0432.CCR-09-1961

©2010 American Association for Cancer Research.

Overview

The traditional first-in-human study or phase 1 investigation is used to determine the dose and schedule of an investigational agent and/or drug. This evaluation also provides the initial description of adverse events associated with agent administration in a dose-dependent fashion. The dose is determined using a variety of dose-escalation strategies that target a toxicity rate of 33% or less. This target is achieved by increasing the dose of the study drug until the toxicity rate reaches 33% (i.e., two of six patients). Investigators then drop back to the next lower dose level(s) to accrue additional patients at the recommended phase 2 dose (RP2D) and schedule [also called the maximally tolerated dose (MTD)] to further evaluate the adverse event profile of the study agent. Many studies have used a "3 + 3" cohort expansion design (Fig. 1) to reach the RP2D. This strategy has been successfully used for the determination of dose and schedule of cytotoxic agents used for patient treatment today; however, it does not necessarily suit the development of many molecularly targeted agents (MTA) and/or drugs in development over the last 10 years.

Phase 1 Cohort Expansion Design	
Number of Patients with DLT at a Given Dose Level	Escalation Decision Rule
0 out of 3	Enter 3 patients at the next dose level
≥ 2	Dose escalation will be stopped. <ul style="list-style-type: none"> • This dose level will be declared the maximally administered dose (highest dose administered). • Three (3) additional patients will be entered at the next lowest dose level if only 3 patients were treated previously at that dose.
1 out of 3	Enter at least 3 more patients at this dose level. <ul style="list-style-type: none"> • If 0 of these 3 patients experience DLT, proceed to the next dose level. • If 1 or more of this group suffer DLT, then dose escalation is stopped, and this dose is declared the maximally administered dose. • Three (3) additional patients will be entered at the next lowest dose level if only 3 patients were treated previously at that dose.
≤ 1 out of 6 at highest dose level below the maximally administered dose	This is considered the maximum tolerated dose and generally the recommended phase 2 dose. At least 6 patients must be entered at the recommended phase 2 dose.

Fig. 1. Standard phase 1 cohort expansion design used to determine dose based on toxicity rate.

Conventional phase 1 study design. The rationale for the 3 + 3 cohort expansion design is pragmatic with regard to determining toxicity-based dose escalation. One or two patients are insufficient to determine whether a 33% toxicity rate has been reached and dose-escalation should halt; therefore, three patients are required for the initial cohort size. Treatment may be escalated to the next higher dose level if no dose-limiting toxicity (DLT) occurs; however, if one drug-related DLT occurs in these three patients, the cohort is expanded to six patients to verify that the toxicity rate has reached 33% (i.e., two of six or fewer patients). When the toxicity rate reaches 33% in a cohort, the next lower dose level will be called the RP2D (so long as the toxicity rate is less than 33%) and the cohort will be expanded to 6 to 15 patients total to establish the preliminary safety profile of the study agent.

The “good,” and thus ethical, phase 1 study is governed by three guiding principles: safety, ethical conduct, and efficiency. Patient safety guides study design by minimizing the number of participants exposed to serious or life-threatening adverse events. Ethical conduct is structured on the concept that patients and their physicians have exhausted all other possible reasonable and standard therapeutic options; therefore, patients with various kinds of advanced malignancies that are either refractory to standard therapy or for which no curative therapy exists, are enrolled. The dose and schedule should be determined safely and efficiently with the least number of

patients exposed to subtherapeutic doses; the starting dose should be defined to minimize the risk to patients receiving initial treatment with a new agent. The RP2D should be generalizable for future disease-specific evaluations. When acceptable toxicity is used as a surrogate for activity and/or efficacy, most responses in phase 1 studies of cytotoxic agents occur within 80 to 120% of the RP2D (1). Efficiency dictates that the study will move forward into phase 2 evaluations in which therapeutic intent is most clearly defined in a disease-specific sense.

The starting dose for MTAs hinges on both safety and efficiency. The derivation of a starting dose for first-in-human phase 1 trials of MTAs in cancer patients is safe but is based on diverse practices using a variety of preclinical toxicologic parameters with both rodent- and nonrodent-based models being used (2, 3). These agents often do not produce DLT, such that the endpoint for the phase 1 investigation can be a maximally administered dose rather than a MTD. When nonrodent data are used for selection of the starting dose and toxicity is expected to be minimal, dose escalation may proceed using dose-doubling rather than a more classical modified Fibonacci dose escalation scheme. When the starting dose is determined using rodent nonclinical models, one tenth of the murine equivalent lethal dose in 10% of the animals, or the MELD10, is considered a safe starting dose in humans. This determination assumes that the toxicities observed interspecies are not different with regard to dose. Eisenhauer and colleagues (3) modeled

Downloaded from <http://aacrjournals.org/clinccancerres/article-pdf/16/6/1726/1995328/1726.pdf> by guest on 14 April 2024

percent MELD with regard to determination of the MTD and the number of dose levels to select a RP2D using trials of cytotoxic agents, and concluded that 20% of the MELD10 would be safe and reasonable to decrease the duration of phase 1 trials and to determine a safe dose. Jaap Verweij has looked at nonclinical toxicity studies and the role they have in determining a safe starting dose (3). He stated that the ratio of the MTD to starting dose was higher for murine and lower for nonmurine species; thus, nonmurine species may be more sensitive for determining MTD in human studies. Efficiency in the phase 1 investigation may be enhanced using dose-doubling for the starting dose if interspecies toxicity is limited and the dose escalation method is conservative.

The efficiency of phase 1 investigations also hinges on the number of patients per cohort, the aggressiveness of the dose escalation scheme, the number of participating centers, and the added value of inpatient dose escalation. Increasing the number of patients per cohort allows for a more accurate assessment of toxicity and diminishes the likelihood that serious adverse events are not identified. This approach increases the number of patients needed for dose and toxicity determination. Later in this manuscript a more extensive description of accelerated dose titration designs will be discussed; these designs limit the number of patients being treated at subtherapeutic levels, while optimizing trial efficiency and assuring patient safety. The concern that rapid dose escalation will compromise patient safety is addressed, using designs that minimize the risk of DLTs.

In the last 30 years the concept of the multi-institutional trial has emerged as an important strategy to increase efficiency. Multi-institutional phase 1 trials refer to those in which several institutions participate in the same trial such that patient slots at each cohort are either assigned by the sponsor or allocated on a first-come, first-serve basis, with the main intention to expedite trial completion. A recent review of phase 1 trials published between 1998 and 2006 in two major cancer journals was undertaken to evaluate the characteristics and efficiency of multi-institutional phase 1 trials. Of 463 trials reviewed, 55% were done in single institutions, whereas the others were conducted in 2 to 16 institutions. Among 30% of the trials that specified accrual time, there was no difference between trials involving one versus more institutions. Trials with one, two, or at least three participating institutions enrolled during a median of 21, 20, and 22 months, respectively. There was no association between the sponsor or the mechanism of drug action and the number of participating institutions. It should be noted, however, that concerns have been raised about multi-institutional phase 1 trials: that they can dilute the experience to recognize toxicity patterns for a novel drug or drug combination, as each investigator may only manage a limited number of trial patients (4). The conduct of multi-institutional phase 1 trials, especially those that involve more than two to three centers, should be discouraged unless justified by a low prevalence of the patient population under study.

The paradigm of toxicity-based dose escalation starts to unravel when novel, noncytotoxic MTAs are evaluated in the context of phase 1 studies. Determination of maximum acceptable toxicity may not be required if well-defined molecular-targeted effects are observed; the investigator is then challenged to define and develop a qualified assay for targeted effect to determine dose. In a review by Parulekar and Eisenhauer, 36 phase 1 studies used toxicity and 8 used pharmacological data as endpoints to select a RP2D. The evaluation of nontoxicity endpoints such as molecular-targeted effects in tumor or surrogate tissues, or functional imaging studies were not used, perhaps because they have not been widely validated (5). Further study and the development of reasonable and/or strict criteria are needed to accurately test the value of this approach to phase 1 development. In some situations, phase 0 investigations are one approach that can be used for the determination of validated molecular-targeted effects that can inform phase 1 endpoints and pharmacokinetics. This approach will be discussed later in this report.

This report will review in more detail the discussions that occurred as part of the Phase 1 Clinical Trial Design Meeting of the Clinical Trial Design Task Force of the Investigational Drug Steering Committee. The authors will discuss novel statistical designs including accelerated titration, continual reassessment methods, and phase 0 designs. In addition, the use of biomarkers and genetic profiling for patients selection in the phase 1 setting will be considered.

Accelerated titration designs. Investigators in oncology have had an ongoing interest in modifications to the standard phase 1 design to make it more efficient, treat fewer patients at nontoxic dose levels (which may be less efficacious), and increase the precision of phase 2 dose recommendations (3, 6). In response to this, Simon and colleagues (7) developed a family of accelerated titration designs (ATD) and proposed the use of an accompanying dose-toxicity model, on the basis of the work of Sheiner (8, 9). The main distinguishing features of these designs are (1) a rapid initial escalation phase; (2) inpatient dose escalation; and (3) the ability to analyze trial results using a dose-toxicity model that incorporates parameters for intra- and interpatient variation in toxicity and cumulative toxicity. All of the designs use 40% dose escalation steps, with dose escalation and/or de-escalation rules based on definitions of DLT and of "moderate" toxicity. The most recommended and popular of the designs (design 4) is carried out as follows.

Design 4 starts with an accelerated phase that uses single patient cohorts, with double dose steps (96–100% dose escalation) per dose level. When the first instance of DLT is observed or the second instance of moderate toxicity is observed (in any course), the cohort for the current dose level is expanded to three patients and the trial reverts to use of the standard phase 1 design for further cohorts (using 40% dose escalation steps).

To maximize each patient's chance to be treated at the potentially active dose, the ATD allows inpatient dose

escalation for a patient who remains on study and has no evidence of toxicity at the current dose. Specifically, the dose for the next course is escalated if less than moderate toxicity is observed for the patient during the current course. If moderate toxicity occurs, then the dose stays the same for the next course. If DLT occurs, the patient generally goes off study; however, if the patient remains on study, the dose is reduced.

Simon and colleagues (7) evaluated the performance of the ATDs by simulating phase 1 data on the basis of 20 sets of parameters estimated from 20 real trials. In the simulations, the average number of patients was much greater for the standard design than for any of the ATDs. With the standard design, an average of 23 patients experienced less than intermediate toxicity; these patients were undertreated. For design 4, with inpatient dose escalation, the number was less than five. The major reduction in the number of undertreated patients was achieved with very small increases in the average number of patients experiencing DLT or unacceptable toxicity.

The model proposed by Simon and colleagues (7) to facilitate analysis of the dose-toxicity relationship after the phase 1 trial is finished, accommodates inter- and inpatient variability, as well as cumulative toxicity. Sheiner and colleagues (8, 9) proposed the use of dose-toxicity models for phase 1 trials two decades ago. These models are rarely used, despite their potential for facilitating the definition of a phase 2 starting dose.

To assess the use and utility of the ATDs in the evaluation of novel oncology therapeutics, Dancey and colleagues (10) conducted a literature search and an analysis of 36 identified ATD trials. Approximately half of the studies used inpatient dose escalation. The ATDs, as used in these studies, rarely resulted in dose escalation beyond the RP2D. Among the 911 patients enrolled in these studies only one death from toxicity occurred during the accelerated titration phase. On the basis of its use in these selected studies, the ATD seems to provide an enhanced efficiency with acceptable safety.

ATDs can dramatically reduce the number of patients accrued to a phase 1 trial, as compared to the standard phase 1 design, and with inpatient dose escalation, ATDs also provide all patients entered in the trial a maximum opportunity to be treated at a therapeutic dose. Despite these benefits, we find that the designs are not widely used, likely owing to the conservativeness of investigators. When they are used, they are often used with an initial dose set at a much more conservative level than would be done for a standard design and without use of inpatient dose escalation, thus reducing their effectiveness. A recent comprehensive review of the risk-benefit relationship for phase 1 trials conducted over the past decade revealed an overall toxicity death rate of only 0.005 (11). An accompanying editorial (12) argued that such a low toxicity death rate, in the context of treatment for an often rapidly fatal disease, suggested that phase 1 trials are being conducted in an overly conservative fashion. Appropriate use of designs such as the ATDs might

increase the potential for benefit in phase 1 trials, with little increase in risk.

Overview of continual reassessment method and related designs. In 1990, O'Quigley and colleagues introduced the continual reassessment method (CRM), a Bayesian adaptive design for dose finding based on a binary toxicity outcome in which the goal is to find MTD of a single treatment agent (13). The novelty of the original CRM was determining the dose for the next patient on the basis of the toxicity outcomes for all patients previously treated on the trial using a mathematical model for the association between dose and toxicity. Current CRMs follow the same general principles as the original, although numerous modifications have been made over the years to add flexibility, extensions, and safety constraints.

Unlike the standard algorithmic phase 1 designs, such as the 3 + 3 and up-and-down approaches, both of which require a list of doses to be tried, the CRM requires the investigator to specify a number of design components in order to develop the trial design.

First, the investigators are required to determine the target level of toxicity, (i.e., the proportion of patients that will show a DLT at the target dose). This is generally chosen to be 0.20 to 0.33, and depends on the patient population and the nature of the expected DLTs. As in standard designs, the definition of DLT must be defined (one common definition of a DLT is any grade 3 or 4 toxicity, as defined by the Common Terminology Criteria for Adverse Events (CTCAE; ref. 14).

Second, the number of patients per cohort is selected. The original CRM included only one patient per cohort; however, more recent modifications suggest using two or three patients per cohort (15–17). The number selected depends on operating characteristics under different cohort sizes, patient accrual, and the desired DLT rate.

Third, to implement a CRM, a mathematical model of the relationship between dose and toxicity needs to be specified in which the outcome (probability of experiencing a DLT) is a function of dose. As part of the choice of mathematical model, the investigator must determine if the doses are from a continuous range or from a fixed set. Currently, there are number of standard approaches; the most popular being the "power" model that uses a discrete set of doses (18), and one- and two-parameter logistic models (16, 19).

Fourth, a stopping rule must be determined. Most CRMs simply predefine a total sample size and proceed with dose finding until the total sample size is reached. However, there have been proposals suggesting more efficient approaches; for example, when the sample size is adaptively determined on the basis of the precision with which the MTD is estimated, or until a certain number of patients are treated within a relatively narrow dose range, suggesting convergence to an optimal dose (19–21).

Implementation of the model will differ if a Bayesian approach is used versus other methods. Bayesian models specify a prior distribution which represents the *a priori* best guess at the dose-toxicity relationship. The dose for

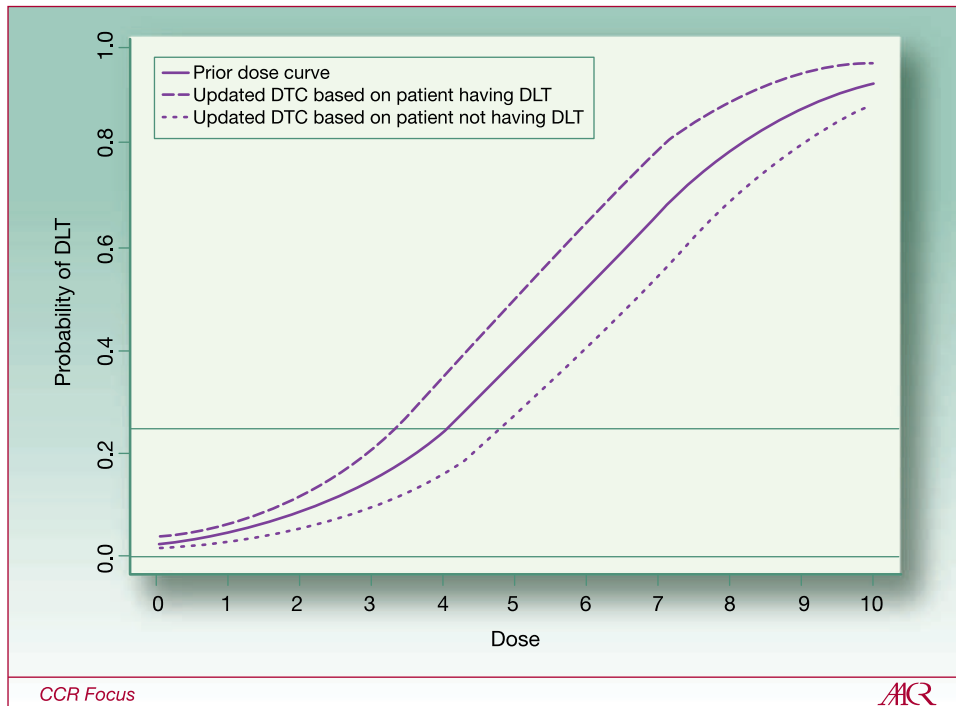


Fig. 2. Theoretical DTCs for continuous reassessment method with one patient per cohort. Solid line shows the prior DTC from which the dose for the first patient is selected. With a desired DLT rate of 0.25, the dose level for the first patient is level 4. The dashed line shows the estimated DTC after observing the first patient if the first patient experienced a DLT. If the first patient experienced a DLT at level 4, then patient 2 would receive dose level 3. The dotted line shows the estimated DTC if the first patient did not experience a DLT at level 4; patient 2 would receive dose level 5.

Downloaded from <http://aacrjournals.org/clinccancerres/article-pdf/16/6/1726/1995328/1726.pdf> by guest on 14 April 2024

the initial cohort is based on the prior dose toxicity curve. Specifically, the selected dose for the first cohort occurs at the dose that is associated with the desired DLT rate.

The initial and following steps for a CRM are illustrated in Fig. 2. If it is assumed that the desired level of toxicity is 0.25 (meaning that it would be acceptable if 25% of patients had a DLT), then, according to this dose-toxicity curve (DTC), dose level 4 would be the optimal starting dose. Assuming that the solid curve in Fig. 2 is the *a priori* curve, the first patient in the trial would receive dose level 4, whether he or she experienced a DLT is recorded and is combined with the prior DTC to get a better estimate of the true curve. When the curve is updated, a new estimated DTC is obtained. As the number of patients accumulates, the DTC estimate is based almost exclusively on the observed data and may look very little like the *a priori* curve. In Fig. 2, the dashed line shows the updated DTC if the first patient experienced a DLT: the updated DTC falls slightly to the left of the solid *a priori* curve and the dose for the next patient is decreased to level 3. However, if the first patient does not experience a DLT, the updated DTC (the dotted line) falls slightly to the right of the *a priori* curve and the dose for the next patient is increased to dose 5. The trial continues in this fashion: with the DTC reestimated after each patient is treated and the next patient being treated at the new estimate of the optimal dose.

As previously mentioned, a number of modifications of the CRM have occurred over the past 20 years. Some notable changes include using non-Bayesian approaches (16, 18, 19); for example, using predefined sets of doses and implementing a 3 + 3 design until a DLT is observed, and then to switching to a CRM to update dose. More flex-

ible models allow for escalation for multiple outcomes (22), escalation for multiple agents (23), and time-to-event toxicity outcomes for cases with delayed toxicities (24, 25). Additionally, other Bayesian designs have been introduced, such as the “escalation with overdose control” approach (26).

Despite the enthusiasm in the statistical community for novel dose-finding designs, these trial designs are not finding their way into clinical trial practice. Rogatko and colleagues reviewed the literature and found that, over a 15-year period, only 1.6% of dose-finding trials used novel designs, with the remaining trials using algorithmic designs (27). For the CRM and its variants, clinical colleagues have been slow or reticent to adopt these designs into practice. A number of factors contribute to the difficulties of practical implementation, including acceptance by protocol review committees, institutional review boards, and regulatory agencies [such as the U.S. Food and Drug Administration (FDA)]. Many of these bodies are charged with preserving patient safety and ethics, and members of the boards may not be convinced that novel designs are better for patients. In fact, they are; however, due to the more sophisticated nature of these designs, it can be challenging to help nonstatisticians understand how they behave. As a result, trials tend to be designed using more traditional methodologies even though they are maybe less efficient in a number of ways.

For more detail on CRM designs and for applications published in clinical journals, please see additional references (28, 29).

Moving beyond the primary goal of safety and dose selection in phase 1 trials. The conventional approach of anticancer

drug development proceeds stepwise from the evaluation of safety in phase 1 trials to the determination of activity in phase 2 trials, and ultimately, to confirmation of effectiveness in phase 3 trials. However, the high attrition rates in oncologic therapeutics have generated tremendous pressures to identify promising drug candidates early on and expedite their advancement, while abandoning those drugs with little hope of ever achieving regulatory approvals (30). In 2000, a new medicinal compound entering phase 1 evaluation only had a 5 to 8% chance of eventually reaching the market, with the leading causes of failure being lack of efficacy and safety issues (30, 31). Given these statistics, many have advocated for the need to establish new paradigms that can identify those compounds that do not hold promise earlier on in the drug development process, thus reducing time and resource investments.

The primary goal of phase 1 trials of novel anticancer agents is to establish a safe and reliable recommended dose and schedule for subsequent phase 2 and 3 testing. In the molecularly targeting era, many phase 1 trials have been conducted with the hope not only of achieving this primary goal, but also to provide a head start on the identification of specific target patient populations, and to generate preliminary biomarker evidence of target inhibition (Fig. 3). There are potential benefits and limitations in the attempt to execute these ancillary goals in phase 1 trials of new drugs. Whether these efforts have added value to the clinical development of MTAs or impeded their evaluation remains a subject of debate and controversy. In addition, the increased use of MTAs has given rise to the idea of small first-in-human studies to test an agent's ability to affect a molecular target (phase 0 studies) prior to initiation of phase 1 testing. These new approaches to the development of MTAs will be discussed in the next three sections.

Identification of specific target patient populations. The vast majority of phase 1 trials evaluate new agents for which target tumor types have not yet been identified, despite extensive preclinical research. As such, patients with various kinds of advanced malignancies that are either refractory to standard therapy or for which no curative therapy exists, are enrolled. When a new drug is combined with a known active regimen in phase 1b trials, patients with malignancies deemed appropriate for such combinations are often preferentially recruited, to optimize their chance of deriving benefit.

Phase 1 trials in which an MTA can be matched precisely with a population of patients whose tumors are driven predominantly by the target of interest, are uncommon. This situation is not surprising given that most advanced human malignancies have complex molecular compositions such that interrogation of a single or even a few of the relevant pathways in a small number of patients in a phase 1 trial without the presence of controls, is at best exploratory. Even for drugs that have been thoroughly tested from phase 1 to phase 3 trials, personalization so far has been the exception rather than the rule. Of 18 MTAs that have obtained regulatory approvals for solid tumors

by the FDA in the last decade (1998–2009), only one third had approvals that were predicated on the biomarker status of the patients' tumors (Table 1).

Recent examples of phase 1 trials that have successfully been enriched for specific target patient populations include the hedgehog inhibitor GDC-0449 in patients with advanced basal cell carcinoma (32, 33), the oral c-MET and ALK inhibitor PF-02341066 in patients with non-small cell lung cancer harboring the EML4-ALK rearrangement (34), and the B-RAF inhibitor PLX4032 in patients with malignant melanoma harboring the V600E B-RAF mutations (35). It is important to emphasize that the scientific basis and knowledge for target patient population selection in these trials are strong, with rationale extending beyond the commonly reported retrospective clinicopathological associations between target expression and clinical outcome with many drugs. Furthermore, even in these enriched phase 1 trials, patients with other advanced solid tumor types were enrolled, as it is essential not to detract from the primary goal of phase 1 trials to recommend a dose and schedule that is safe for subsequent disease-specific evaluations not limited to the target patient groups.

For the purpose of dose escalation, most contemporary phase 1 trials continue to recruit patients with advanced malignancies without enrichment when unique target populations are not clinically apparent. However, it has become increasingly popular for phase 1 trials of targeted agents to open enriched expansion cohorts once the RP2D has been reached, with the hopes to gain insight

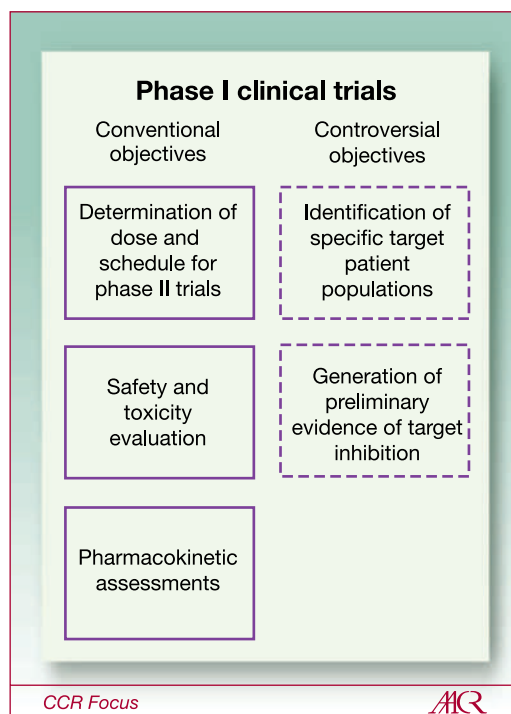


Fig. 3. Objectives of phase 1 clinical trials. Conventional objectives of phase 1 trials are listed in the left column and more controversial objectives of phase 1 trials are listed in the right column.

on the so-called proof-of-concept and to acquire early hints of antitumor activity in patients felt to possess the greatest chance of response. Molecular profiling of tumors using genotyping technologies for somatic mutations and gene amplifications, or selection of tumor types based on published frequencies of molecular aberrations, are examples of enrichment strategies that have been used at the end of phase 1 trials or in early phase 2 trials. The jury remains out on the merit of enrichment in early phase clinical trials; some investigators strongly believe that this approach will speed up the drug development path, whereas others fear that it is too limited without sufficient scientific justification (36). Regardless of the view, the reproducibility, reliability, and validity of the technological platforms must be assured if the results they generate are being applied to patient selection decisions in the clinical setting.

Generation of Preliminary Biomarker Evidence of Target Inhibition. Biomarkers are biological variables or characteristics that are measured by molecular, biochemical, or imaging techniques that can be evaluated as indicators of normal physiologic or pathological processes, or as

pharmacologic responses to a therapeutic intervention (37). Examples of biomarkers in phase 1 trials include variables measured at one time point, such as immunohistochemical expression of an activated protein in tumor tissues prior to study drug initiation, or serial measurements such as changes in serum concentrations of a growth factor, or changes in standardized uptake values of target lesions on fluorodeoxyglucose positron emission tomography (FDG-PET) before and after drug administration. The inclusion of biomarkers in phase 1 trials has significantly increased from 1991 to 2002 as reviewed by Goulart and colleagues (38), and undoubtedly, this upward trend has continued. In this review, the use of biomarkers to support dose and/or schedule selection for phase 2 trials was found in only 13% of published phase 1 trials; they were potentially useful for selecting a patient population in subsequent studies in 19% of trials; whereas a greater proportion of trials (39%) reported their role in providing evidence to support the proposed mechanism of action of the drug (38). For active drugs that are destined to succeed in their development, it is unclear how much this type of proof of mechanism by biomarkers in phase

Table 1. FDA approval of molecularly targeted drugs in solid tumors (1998–2009), according to whether approval status is predicated on molecular biomarker status of patients' tumors

Molecularly targeted agent	Class	Year of FDA approval	Approval status predicated on molecular biomarker status of patients' tumors
Trastuzumab	Anti-EGFR MAb	1998	Yes (HER2)
Imatinib	Multikinase TKI	2001	Yes (BCR-ABL)
Bortezomib	Proteasome inhibitor	2003	No
Gefitinib	Anti-EGFR TKI	2003	No*
Erlotinib	Anti-EGFR TKI	2004	No
Cetuximab	Anti-EGFR MAb	2004	Yes (EGFR)
		2009	Yes (K-RAS)
Bevacizumab	Anti-VEGF MAb	2004	No
Sorafenib	Multikinase TKI	2005	No
Sunitinib	Multikinase TKI	2006	No
Panitumumab	Anti-EGFR MAb	2006	Yes (EGFR)
		2009	Yes (K-RAS)
Vorinostat	HDAC inhibitor	2006	No
Dasatinib	Multikinase TKI	2006	Yes (BCR-ABL)
Decitabine	Hypomethylating agent	2006	No
Nilotinib	Multikinase TKI	2007	Yes (BCR-ABL)
Lapatinib	EGFR/HER2 TKI	2007	Yes (HER2)
Temsirolimus	mTOR inhibitor	2007	No
Everolimus	mTORi inhibitor	2009	No
Pazopanib	Multikinase TKI	2009	No

Abbreviations: EGFR, epidermal growth factor receptor; MAb, monoclonal antibody; TKI, tyrosine kinase inhibitor; VEGF, vascular endothelial growth factor; HDAC, histone deacetylase.

*Gefitinib has been approved for the first-line treatment of EGFR-positive non-small cell lung cancer by the European Medicines Agency (EMA) in 2009.

1 trials has enhanced their approval process over and above that achieved by radiological evidence of antitumor activity. Certainly, on the basis of the list of MTAs in Table 1, all of these agents were approved by the FDA because of radiological response or delay in tumor progression that has translated to an improvement in progression-free and/or overall survival. None of these active agents would have entered definitive phase 3 testing if biomarker evidence of target inhibition in early phase trials alone (e.g., laboratory evidence of inhibitor of downstream markers, reduction in apoptosis, or proliferation, etc.) was seen without objective radiological efficacy.

Perhaps a more informative role for biomarkers that are being done to show target inhibition in phase 1 trials is to help provide warning of agents with uncertain therapeutic indices. Early termination of the development of agents with limited signs of antitumor activity, or those with significant toxicity, in conjunction with minimal biological evidence of target inhibition, may help reduce the high failure rates of oncologic therapeutics. Another potentially valuable role for biomarkers of target inhibition is to shed information on novel first-in-class or best-in-class agents. Changes in target modulation by the drug in surrogate or tumor tissues of patients may yield new knowledge of unforeseen mechanisms of drug activity or resistance.

Phase 0 trials. The phase 0 trial is a new type of study, designed to be a first-in-human; the idea was initiated by an FDA guidance in 2006 (39). This type of study can be conducted to assess drug effect on a molecular target, by means of a pharmacodynamic assay in a very small number (i.e., 10–15) of patients, which promises to make phase 1 trials, and the early trial development process, in general, more efficient and effective. Currently only 5% of the investigational new drug (IND) applications to the FDA result in clinically approved agents (39–41). The fact that an increasing proportion of IND agents are molecularly targeted suggests that testing the agent for effectiveness against the target by means of a pharmacodynamic assay needs to be conducted very early in the drug development process. This early testing is particularly useful and important because preclinical tests of such effectiveness are often misleading. Phase 0 studies can be administered while the toxicology studies preparatory to filing a standard IND are being conducted. Phase 0 studies can also be used to prioritize among analogs or agents designed to have the same molecular target by means of comparing pharmacokinetic (for example, oral bioavailability) and/or pharmacodynamic characteristics, or, in imaging studies, to verify localization of the agent to the tumor. These studies represent an opportunity for developing and validating clinical pharmacodynamic assays very early in the drug development process. Typically, a phase 0 trial encompasses several escalating dose levels for the experimental agent. Therefore, they can contribute to better defining the appropriate dose range or administration schedule to take into phase 1 and phase 2.

Kummar and colleagues (40) and Murgo and colleagues (41) have described several phase 0 statistical designs. The first step in the trial design is to define what is meant by a pharmacodynamic “response” for each individual patient, which is analogous to defining what constitutes an objective tumor response for a patient in a phase 2 trial. In oncology, generally, the pharmacodynamic endpoint is assessed both in tumor tissue and in an easily assayed surrogate tissue such as blood [peripheral blood mononuclear cells (PBMC)]. The tumor tissue assay is considered to be more reliable with respect to reflecting the biological effect of the agent in what is generally the target tissue of interest (39, 42); however, the number of tumor biopsies usually is severely limited for ethical reasons (40, 41). The measure of treatment effect for the tumor pharmacodynamic assay is the difference between the pretreatment and post-treatment values. Defining a pharmacodynamic response, both for the tumor assay and for the PBMC assay, usually involves both a biologic criterion and a statistical criterion for what is significant. The biologic criterion generally depends upon characteristics of the biologic target of the agent. For example, in the National Cancer Institute (NCI) phase 0 trial of ABT-888 (42, 43), an inhibitor of the DNA repair enzyme poly (ADP-ribose) polymerase (PARP), the criterion chosen was that the reduction in the assay value had to be at least 50%. The statistical criterion may be either 90% confidence or 95% confidence (generally one-sided, because the anticipated treatment effect is generally in one direction) that the observed treatment effect is not a result of the sort of natural random variation in the assay, for an individual patient, that would be seen in the absence of a true treatment effect. Details about the definition of a pharmacodynamic response are illustrated in Fig. 4.

The second step in the trial design is to define what constitutes a promising observed pharmacodynamic response rate for each dose level; in other words, how many patients must show a pharmacodynamic response for the dose level to be declared biologically effective. This is analogous to setting a threshold for observed response rate in a phase 2 trial, which determines if the agent is deemed sufficiently promising for further testing (44). For a targeted pharmacodynamic response rate across patients, the power to declare the dose level effective for each of the two assays, as well as the false positive rate (in case the agent is ineffective), can be calculated. Figure 4 gives an example of a design to target a 60% pharmacodynamic response rates, across patients.

The NCI selected ABT-888 for the first-ever phase 0 trial for two reasons (42, 43). It was anticipated to have a wide margin of safety, and because elevated PARP levels are characteristic of tumors and can result in resistance to both chemotherapy and radiotherapy, PARP inhibitors hold promise of wide applicability as chemotherapy and radiotherapy sensitizers. The NCI trial showed statistically significant reduction in PARP levels (a surrogate for PARP inhibition) in both tumor and PBMCs (42, 43). It also gave an opportunity to explore the correlation between

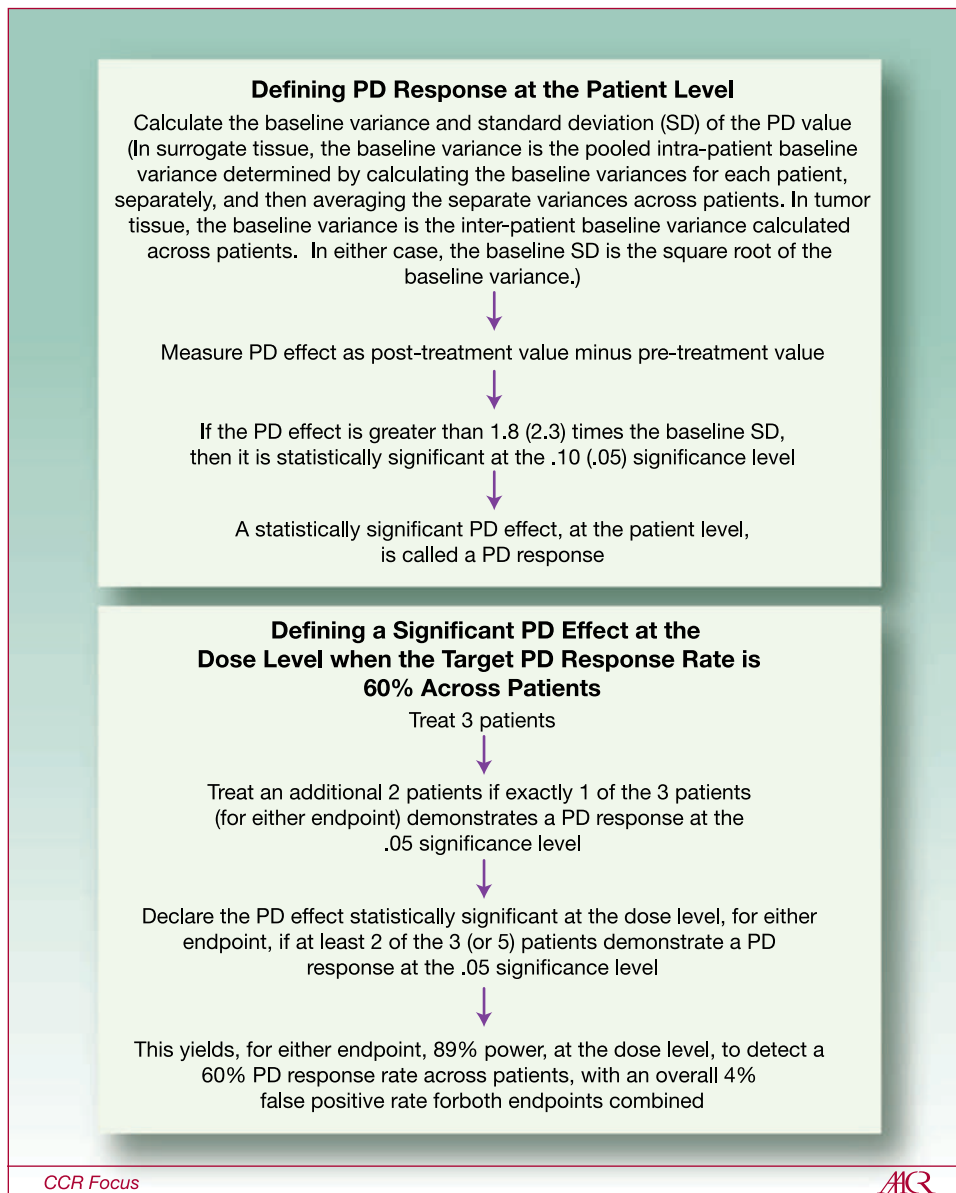


Fig. 4. (Top panel) Definition of pharmacodynamic (PD) response for an individual patient. Multipliers of the baseline standard deviation (SD) are derived from asymptotic normal distribution theory. Significance levels are one-sided. (Bottom panel) Definition of a promising observed response rate for a dose level with a two-stage design. The target (true) pharmacodynamic response rate, across patients, is 60%. Power and false positive rate are derived from the binomial distribution. Adapted with permission from Murgo et al. (41).

Downloaded from <http://aacrjournals.org/clinccancerres/article-pdf/16/6/1726/1995328/1726.pdf> by guest on 14 April 2024

blood and tissue marker levels, to determine to what extent blood levels could be used as a surrogate for the more difficult to obtain tissue levels.

To our knowledge, the NCI trial is the only phase 0 trial completed to date. Phase 0 trials do not replace phase 1 trials conducted in establishing DLTs and defining an RP2D. In contrast, data from phase 0 trials allow phase 1 studies to begin at a higher, potentially more efficacious dose, use a more limited and rationally focused schedule for pharmacodynamic sampling, and use a qualified pharmacodynamic analytic assay for assessing target modulation. Likewise, phase 0 trials, with pharmacodynamic endpoints, will not eliminate the need for phase 2 trials to establish the agent's ability to yield tumor response or clinical benefit; but, they will

allow for early termination of development of agents that fail to yield the anticipated biologic effect. Therefore, effort spent to conduct rationally designed phase 0 trials should conserve resources in the long run by improving the efficiency and success of subsequent clinical development (45, 46).

Summary

The design of phase 1 trials remains an open issue. The Clinical Trial Design Task Force initially addressed the merits, advantages, and disadvantages, of a variety of phase 1 approaches to drug development for first-in-human studies with investigational agents. The goals, objectives, and purposes for these studies continue to evolve rapidly and

are now further challenged by the addition of biomarker-based selection of patients to participate in these studies. The adage to “fail early and fail fast” is used to define drug development as more is asked of the phase 1 study. This discussion has addressed traditional, accelerated, and biomarker-driven trial designs. Other issues related to phase 1 trial designs, such as late-onset toxicities, cumulative toxicities, and multidrug combination are not covered by the scope of this review. The Clinical Trial Design Task Force of the Investigational Drug Steering Committee developed recommendations on phase 1 trial designs that are pragmatic and encourage the investigator to select a design that best suits the development of the agent under

study. The choice of the design that best suits the agent versus the agent that best suits the design remains fluid.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Jessica Cleck, PhD, Rochelle Coelho, MD, PhD, and Monica Chiaramonte, PhD, for their assistance in the preparation and editing of this manuscript.

Received 01/05/2010; revised 01/21/2010; accepted 01/25/2010; published OnlineFirst 03/09/2010.

References

- Von Hoff DD, Turner J. Response rates, duration of response, and dose response effects in phase I studies of antineoplastics. *Invest New Drugs* 1991;9:115–22.
- Le Tourneau C, Stathis A, Vidal L, Moore MJ, Siu LL. Choice of starting dose of molecularly targeted agents evaluated in first-in-human phase I cancer clinical trials. *J Clin Oncol*. 2010 Feb 1 [Epub ahead of print].
- Eisenhauer EA, O'Dwyer PJ, Christian M, et al. Phase I clinical trial design in cancer drug development. *J Clin Oncol* 2000;18:684–92.
- Dowlati A, Manda S, Gibbons J, et al. Multi-institutional phase I trials of anticancer agents. *J Clin Oncol* 2008;26:1926–31.
- 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;96:990–7.
- Arbuck SG. Workshop on phase I study design. Ninth NCI/EORTC New Drug Development Symposium, Amsterdam, March 12, 1996. *Ann Oncol* 1996;7:567–73.
- Simon R, Freidlin B, Rubinstein L, et al. Accelerated titration designs for phase I clinical trials in oncology. *J Natl Cancer Inst* 1997;89:1138–47.
- Sheiner LB. Implications of an alternative approach to dose-response trials. *J Acquir Immune Defic Syndr* 1990;3:S20–6.
- Sheiner LB, Beal SL, Sambol NC. Study designs for dose-ranging. *Clin Pharmacol Ther* 1989;46:63–77.
- Dancey J, Freidlin B, Rubinstein L. Accelerated titration designs. In: Chevret S, editor. *Statistical methods for dose-finding experiments*. Chichester, West Sussex (England); Hoboken (NJ): Wiley Press; 2006, p. 91–114.
- Horstmann E, McCabe MS, Grochow L, et al. Risks and benefits of phase 1 oncology trials, 1991 through 2002. *N Engl J Med* 2005;352:895–904.
- Kurzrock R, Benjamin RS. Risks and benefits of phase 1 oncology trials, revisited. *N Engl J Med* 2005;352:930–2.
- O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase 1 clinical trials in cancer. *Biometrics* 1990;46:33–48.
- Cancer Therapy Evaluation Program, Common Terminology Criteria for Adverse Events, Version 3.0, DCTD, NCI, NIH, DHHS, March 31, 2003.
- Faries D. Practical modifications of the continual reassessment method for phase I cancer clinical trials. *J Biopharm Stat* 1994;4:147–64.
- Goodman SN, Zahurak ML, Piantadosi S. Some practical improvements in the continual reassessment method for phase I studies. *Stat Med* 1995;14:1149–61.
- Moller S. An extension of the continual reassessment methods using a preliminary up-and-down design in a dose finding study in cancer patients, in order to investigate a greater range of doses. *Stat Med* 1995;14:911–22, discussion 923.
- O'Quigley J, Shen LZ. Continual reassessment method: a likelihood approach. *Biometrics* 1996;52:673–84.
- Piantadosi S, Fisher JD, Grossman S. Practical implementation of a modified continual reassessment method for dose-finding trials. *Cancer Chemother Pharmacol* 1998;41:429–36.
- Ishizuka N, Ohashi Y. The continual reassessment method and its applications: a Bayesian methodology for phase I cancer clinical trials. *Stat Med* 2001;20:2661–81.
- Zohar S, Chevret S. Phase I (or phase II) dose-ranging clinical trials: proposal of a two-stage Bayesian design. *J Biopharm Stat* 2003;13:87–101.
- Braun TM. The bivariate continual reassessment method. extending the CRM to phase I trials of two competing outcomes. *Control Clin Trials* 2002;23:240–56.
- Mandrekar SJ, Cui Y, Sargent DJ. An adaptive phase I design for identifying a biologically optimal dose for dual agent drug combinations. *Stat Med* 2007;26:2317–30.
- Cheung YK, Chappell R. Sequential designs for phase I clinical trials with late-onset toxicities. *Biometrics* 2000;56:1177–82.
- Braun TM. Generalizing the TITE-CRM to adapt for early- and late-onset toxicities. *Stat Med* 2006;25:2071–83.
- Babb J, Rogatko A, Zacks S. Cancer phase I clinical trials: efficient dose escalation with overdose control. *Stat Med* 1998;17:1103–20.
- Rogatko A, Schoeneck D, Jonas W, et al. Translation of innovative designs into phase I trials. *J Clin Oncol* 2007;25:4982–6.
- Garrett-Mayer E. The continual reassessment method for dose-finding studies: a tutorial. *Clin Trials* 2006;3:57–71.
- Loeb DM, Garrett-Mayer E, Hobbs RF, et al. Dose-finding study of 153Sm-EDTMP in patients with poor-prognosis osteosarcoma. *Cancer* 2009;115:2514–22.
- Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? *Nat Rev Drug Discov* 2004;3:711–5.
- Administration, UDoHaHS-FaD. Innovation or stagnation: Challenges and opportunity on the critical path to new medical products. 2004. <http://www.fda.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm077262.htm>.
- Von Hoff DD, LoRusso PM, Rudin CM, et al. Inhibition of the hedgehog pathway in advanced basal-cell carcinoma. *N Engl J Med* 2009;361:1164–72.
- LoRusso PM, Rudin CM, Borad MJ, et al. A first-in-human, first-in-class, phase (ph) I study of systemic Hedgehog (Hh) pathway antagonist, GDC-0449, in patients (pts) with advanced solid tumors. *ASCO Meeting Abstracts* 2008;26:3516.
- Kwak EL, Camidge DR, Clark J, et al. Clinical activity observed in a phase I dose escalation trial of an oral c-met and ALK inhibitor, PF-02341066. *ASCO Meeting Abstracts* 2009;27:3509.
- Flaherty K, Puzanov I, Sosman J, et al. Phase I study of PLX4032: Proof of concept for V600E BRAF mutation as a therapeutic target in human cancer. *ASCO Meeting Abstracts* 2009;27:9000.
- Garber K. Trial offers early test case for personalized medicine. *J Natl Cancer Inst* 2009;101:136–8.

37. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001;69:89–95.
38. Goulart BH, Clark JW, Pien HH, et al. Trends in the use and role of biomarkers in phase I oncology trials. *Clin Cancer Res* 2007;13:6719–26.
39. Kinders R, Parchment RE, Ji J, et al. Phase 0 clinical trials in cancer drug development: from FDA guidance to clinical practice. *Mol Interv* 2007;7:325–34.
40. Kummar S, Kinders R, Rubinstein L, et al. Compressing drug development timelines in oncology using phase '0' trials. *Nat Rev Cancer* 2007;7:131–9.
41. Murgo AJ, Kummar S, Rubinstein L, et al. Designing phase 0 cancer clinical trials. *Clin Cancer Res* 2008;14:3675–82.
42. Kummar S, Gutierrez M, Kinders R, et al. Phase 0 pharmacodynamic, pharmacokinetic study of ABT-888, an inhibitor of poly (ADP-ribose) polymerase (PARP), in patients with advanced malignancies. *Ann Oncol* 2008;19:iii12.
43. Kummar S, Kinders R, Gutierrez ME, et al. Phase 0 clinical trial of the poly (ADP-ribose) polymerase inhibitor ABT-888 in patients with advanced malignancies. *J Clin Oncol* 2009;27:2705–11.
44. Simon R. Optimal two-stage designs for phase II clinical trials. *Control Clin Trials* 1989;10:1–10.
45. Editors. Phase 0 trials: a platform for drug development? *Lancet* 2009;374:176.
46. LoRusso PM. Phase 0 clinical trials: an answer to drug development stagnation? *J Clin Oncol* 2009;27:2586–8.