Commentary

Workshop on phase I study design*
Ninth NCI/EORTC New Drug Development Symposium, Amsterdam, March 12, 1996

S. G. Arbuck
Investigational Drug Branch, Cancer Therapy Evaluation Program, Division of Cancer Treatment, Diagnosis and Centers, National Cancer Institute, Rockville, MD, USA

Summary

Many of the methods currently used in phase I trial design date back to the 1970's. Recently, some investigators have considered whether it might be possible to perform more efficient phase I trials that minimize the number of patients treated at biologically ineffective doses, maximize the precision of recommended phase II doses, and maintain patient safety. The objective of this Workshop was to examine aspects of phase I trial design, with consideration of both standard and novel approaches. Topics included choice of starting dose level, definition of dose-limiting toxicity, development and incorporation of non-toxicity endpoints, use of alternate dose escalation schemes, and definitions of tumor response and clinical benefit, including consideration of whether these determinations from phase I studies should alter the decision to proceed to phase II drug evaluation. For each topic, this summary includes a description of the standard approach, a summary of the speaker's presentation, and commentary. The Workshop initiated discussion and reassessment that are expected to lead to testing of some of the suggested alternate approaches.

Key words: cancer drug development, dose, dose escalation, dose-limiting toxicity, phase I trials, response, toxicity, toxicology

Introduction

Many of the guidelines established for phase I trial execution were developed during the 1970's. Elizabeth Eisenhauer introduced this Workshop by defining questions requiring reconsideration. She noted that the adequacy of phase I trials is best assessed by determining whether recommended doses prove adequate in phase II trials; however, there has been no systematic evaluation to address this question. The definition of dose-limiting toxicity determines the recommended phase II dose, and an accepted incidence of toxicity in phase II studies is necessary to assess how suitable the recommended dose was. During this workshop Dr. Stan Kaye proposed criteria for dose-limiting toxicity.

Recently, attention has been directed toward the objective of treating as many phase I patients as possible close to the maximally tolerated dose (MTD); yet, assurance of patient safety remains an important requirement. Choice of both starting dose and dose escalation scheme determines distribution of patients at biologically active and inactive dose levels, and affects safety. Starting dose was discussed by Dr. Jaap Verweij, and Dr. Michaele Christian addressed alternate dose escalation schemes. Dr. Will Steward discussed the importance of non-toxicity phase I endpoints for development of novel agents, with varying molecular targets and mechanisms of action. Although response has not been considered a primary objective of phase I trials, Dr. Dan Von Hoff considered whether it should affect the decision to continue development of a new agent.

Starting dose levels for phase I studies

Guidelines for determining phase I starting doses for chemotherapeutic agents were developed following retrospective reviews that compared specific toxic dose levels in animals and in humans. As safety is the primary concern, the starting dose is usually calculated as the highest fraction of a specific toxic dose level, in a particular animal species, that did not exceed the MTD for any of the drugs reviewed retrospectively. Reviews of data for many drugs from any single animal species revealed that the ratio of that specific toxic dose level to the human MTD is highly variable, ranging over about 2 logs. Frequently toxicology results from more than one animal species, most often a rodent and non-rodent species, are used to determine an appropriate starting dose. In these instances, starting doses are based upon the most sensitive species. Starting doses have traditionally been 1/10 LD10 (dose that was lethal to 10% of animals) in the mouse or 1/3 toxic.

*The US Government right to retain a non-exclusive, royalty free licence in and to any copyright is acknowledged.
dose low (TDL or first toxic dose) in the dog. The distributions of the ratio of 1/10 LD10 or 1/3 TDL to the human MTD have been shown to be comparable, but for individual drugs these ratios may be quite different. Multiple reviews that update these extrapolations have been published.

Dr. Verweij summarized his review of published phase I data for 100 anticancer drugs [1]. His objectives were to determine what starting doses were used, whether they were safe but not overly conservative, and if toxicology studies from a second or third species contributed to choice of a safe phase I starting dose. Of the 100 drugs, he identified 71 for which he could extract published toxicology information (including that obtained from mouse, rat and/or dog studies), and phase I results. For 57 of these drugs (80%), the selection of starting dose was based upon 1/10 the mouse LD10. For the remaining 14 drugs, a different starting dose was used, frequently because of results from a more sensitive species, often the dog.

In an effort to assess both safety and efficiency, the ratio between 1/10 mouse LD10 and the dose resulting in toxicity in patients was determined for all drugs. Clinical toxicity was evaluated using two definitions: (1) the maximally tolerated dose (MTD) defined as the dose producing dose limiting toxicity (DLT); and (2) the first toxic dose (FTD), defined as the dose at which the first relevant toxicity occurs, i.e., myelosuppression and/or nausea/vomiting \( \geq \) grade 2 or any organ toxicity. The median ratios of the MTD (dose that produced dose limiting toxicity) in the phase I trials to 1/10 of the LD10 for drugs using mouse and those using non-mouse based starting doses were 20 (range –2 to 248) and 11 (range –2 to 93), respectively. The median ratio of the first toxic dose (FTD) level in the clinic to 1/10 LD10 was also larger for the mouse group (8 vs. 5) with ranges of –2 to 127 and –2 to 12, respectively. These ratios indicate that 1/10 the mouse LD10 sometimes exceeded the MTD (the use of dose de-escalation is indicated by negative numbers), and they ranged over two logs, as determined in previous studies. Both the ratios and the ranges are smaller for the drugs where non-mouse data were used, consistent with the hypothesis that this group had a lower margin of safety when only mouse data were considered. The median number of dose escalation steps in the phase I trials was similar for the mouse and non-mouse groups, 9 (range –1 to 18) and 10 (range 4 to 21), respectively. These data indicate that use of non-mouse toxicology did not alter the number of dose escalation steps in the phase I studies, but did avoid dose de-escalation. These ratios were also considered for various subgroups defined by schedule (bolus vs. daily \( \times 5 \)) and type of toxicity (hematologic or non-hematologic). Results appear similar but are difficult to assess because of the small numbers in the non-mouse subgroups. Similarly, it is difficult to draw conclusions about differences in ratios of MTD/LD10 for different classes of agents.

The most crucial part of the review determined the number of drugs for which the ratio of clinical MTD: 1/10 mouse LD10 was small, indicating little difference between the mouse LD10, and the MTD in man. Since safety is a primary consideration in the execution of phase I trials, one might wish to observe a ratio of at least 2, indicating that the MTD was two-fold higher than the starting dose. For 57 agents where the starting dose was 1/10 mouse LD10, the ratio was \( \leq \) 2 in 2% of cases and \( \geq \) 5 in the remaining 98%. In contrast, for the remaining 14 drugs where the starting dose was based upon a more sensitive animal species, if 1/10 mouse LD10 had been used, 7% would have resulted in ratio \( \leq \) 1 and 14% a ratio of 2. The remaining 79% had a safer ratio of \( \geq \) 5. These data document the need for toxicology studies in a second species, since the second species contributed to choice of a safer starting dose in 21% of cases. However, Dr. Verweij also noted that for these 14 drugs (almost 20%) with marked interspecies differences in toxicity, rats were adequately predictive and dog toxicology data did not contribute to choosing a safe starting dose for phase I trials. Based upon his retrospective evaluation, Dr. Verweij suggested that a less conservative starting doses above 1/10 LD10 in mice could be considered, especially for drugs to which mice and rats are equally sensitive.

**Commentary**

Dr. Verweij’s review documents that a second species is required when 1/10 the mouse LD10 is considered for choice of starting dose to avoid initiating phase I trials with a starting dose \( \leq \) 2 times the MTD in humans. These results are consistent with NCI results with 20 drugs where preclinical toxicity studies using 2 or 3 species provided an acceptable clinical starting dose 97% of the time, but the use of mouse data alone provided an acceptable dose only 83% of the time [2]. Others have recommended the use of only a single rodent species [3]. However, if only mouse toxicology was used to determine starting dose, this review demonstrates that a fraction smaller than 1/10 the mouse LD10 would be required to ensure patient safety. It is likely that use of a smaller fraction of the mouse LD10 would be less efficient than the current approach for the vast majority of drugs studied. Differing views exist regarding the contribution of dog toxicology, particularly for determining starting doses for continuous intravenous infusion schedules.

Dr. Verweij also suggested using starting doses above 1/10 LD10 in mice, especially for drugs to which mice and rats are equally sensitive. However, it should be noted that if safety is to remain a primary concern, one would not save more than one or two dose escalation steps. Furthermore, one would not wish to use a higher starting dose in conjunction with more aggressive dose escalation schemes and smaller initial patient cohorts, as proposed later by Dr. Christian.
Dose-limiting toxicity: Can we make uniform definitions?

Conventionally, the MTD is the dose level at which at least one third of the patients experience DLT. A minimum of 6 patients are usually treated at this level. It is important to note however, that the MTD will depend upon the criteria set for DLT. These sometimes vary because investigators may be more conservative or more aggressive in their definition of acceptable and unacceptable toxicity.

Dr. Kaye’s first recommendation was that more than 6 patients be entered at the dose recommended for phase II evaluation (usually one dose level below the MTD) to help ensure that an appropriate phase II dose has been identified. Additional recommendations were made to help standardize the determination of the MTD. These included simultaneous consideration of hematologic and non-hematologic toxicities to define the MTD. He then considered criteria that should define dose-limiting toxicity.

Prior to the availability of G-CSF, dose-limiting myelosuppression was usually defined as grade 3 myelosuppression (WBC <2000/ul or ANC <1000/ul). More recently, with the availability of G-CSF, many phase I trials have defined dose-limiting myelosuppression as grade 4 (ANC <500/ul; platelets <25,000/ul). Some investigators have included additional requirements for a defined duration of grade 4 myelosuppression (often 7 or more days) or the occurrence of complications (for example, fever requiring antibiotics, or bleeding). While these additional requirements were intended to emphasize complications and increased risks associated with myelosuppression, they provide little margin for error and toxicity could be excessive. Nevertheless, Dr. Kaye believes that this approach should continue to be studied [4]. He did advocate, however, that attempts at escalating doses beyond the MTD with growth factors should not be incorporated into initial phase I trials. When myelosuppression is dose limiting, it is important to define an MTD according to the extent of prior exposure to cytotoxic therapy. At least two courses should be given to determine if hematologic toxicity is cumulative.

For non-hematologic toxicity, he suggested that more careful and specialized monitoring may be advisable. He noted that dose-limiting toxicity is usually defined as grade III toxicity which is sometimes associated with significant organ damage. All organ systems are appropriately included in this determination, but alopecia and nausea/vomiting should be excluded. Dr. Kaye suggested that phase I trials attempt to assess whether non-hematological toxicities are cumulative, reversible, schedule dependent and preventable, but noted that these issues frequently require further study in phase II trials.

He emphasized that the MTD should not automatically be chosen as the phase II dose without careful consideration of other factors, including, for example, how excretory organ function impairment impacts on the MTD.

Commentary

The term MTD has generated confusion because it is used in two ways by investigators. Some use it to describe the dose associated with dose-limiting toxicity in more than 1 of 6 patients. This dose might more accurately be described as the ‘highest administered dose’. The next lower dose level, which is frequently the recommended phase II dose, is also sometimes referred to as the maximally tolerated dose. To avoid confusion, however, this dose level could be called the ‘recommended phase II dose’.

The challenging task of developing a uniform definition of dose-limiting toxicity provides an opportunity to consider reassessment and revisions to the Common Toxicity Criteria to ensure as much as possible, equivalence of toxicity by grade, across organ systems. Such reassessment would also provide an opportunity to include additional toxicities that have been identified in trials in recent years, for example, edema and fluid accumulation.

Once toxicity is well defined, definitions of acceptable and unacceptable toxicity must be agreed upon. While the evaluation of a more aggressive definition of dose-limiting neutropenia (grade 4 neutropenia <7 days) was appropriately performed, results of these studies must be considered. The incidence of fever increases by 10% for each day of grade 4 neutropenia [5]. Seven days of grade 4 neutropenia would likely lead to an incidence of febrile neutropenia of 60%–70%, a rate that most would consider excessive for a phase II dose to screen for anticancer activity. The risks of this approach are further demonstrated by recent experience in CALGB 9430, a randomized phase II trial in patients with previously untreated extensive small cell lung cancer (T. Lynch et al., unpublished data). The phase II doses for two of three of the arms on this trial were determined using phase I designs that permitted up to 7 days of neutropenia in the absence of complications. The death rates 3/12 (25%) and 3/13 (23%) were unacceptably high in the phase II study, leading to premature closure of these arms, and reconsideration of appropriate phase II doses. With these results now available, the determination of a definition of dose-limiting neutropenia should be reconsidered, by first determining an acceptable incidence of febrile neutropenia. With this target agreed upon, data are available to determine the number of days of grade 4 neutropenia that should be permitted in the phase I trial [5]. With the limited number of patients at any phase I dose level, and with less frequent monitoring of blood counts, it may still be difficult to carefully define a safe phase II dose based upon duration of dose-limiting grade 4 neutropenia in phase I trials. This issue is important and requires careful resolution.
Alternate endpoints

Phase I clinical trials of new chemotherapeutic agents have classically utilized toxicity as the endpoint to determine the dose and schedule to be used for efficacy trials. Although a dose effect is more easily demonstrated for toxicity than for efficacy, the maximally tolerated dose is assumed to be associated with the greatest chance of efficacy. The MTD or a dose close to the MTD is used for early phase II trials that purport to detect ineffective drugs and are designed to minimize the probability of a false negative screening result.

Dr. Steward explained that toxicity may no longer be an appropriate endpoint in some phase I studies [6]. These include studies of drugs characterized by a dose-response curve that reaches a plateau at non-toxic doses, or that have the desired effect at doses without significant toxicity. Probable examples include angiogenesis inhibitors, metalloproteinase inhibitors, signal transduction inhibitors and colony stimulating factors. Toxicity would also not be an appropriate endpoint for phase I studies of agents with a bell-shaped dose-response curve and a dose-toxicity curve that rises after the maximal therapeutic dose. Probable examples include interferons, some interleukins and negative regulators of hemopoiesis. With development of novel biologic agents such as these, the use of toxicity as an endpoint would result in unnecessary toxicity and wasted resources. The need for new endpoints is also apparent for new modalities including monoclonal antibodies and gene therapy, which may produce nonspecific and sporadic toxicities that are not clearly dose-related.

Phase I clinical trials are becoming more complex as new drug development focuses less on classical cytotoxic agents and more on agents with greater specificity for tumor cells. While assessment of toxicity is an essential part of phase I studies, increasing emphasis on alternate endpoints will require measurement of the effect of different drug concentrations on the targets of interest using biochemical and biological assays. These studies will become increasingly important determinants of the dose chosen for future development. Pharmacokinetic assessment and correlation of drug levels with target effects, in addition to pharmacodynamic measures from tumor biopsies, will be necessary.

Commentary

The more sophisticated approaches described by Dr. Steward will be required for rational development of novel therapeutics with different targets. Development and validation of these approaches presents major challenges. It is critical to validate that the proposed target endpoints are correlated with activity. Many new agents have multiple effects (for example, suramin), and sometimes variable effects on the same target (for example, bryostatin). Thus, careful laboratory evaluation will be required. In addition, when normal cells are used to predict effects on tumor cells, it is important to validate that the surrogates are appropriate. Validation studies are more easily performed in animal models where studies can be repeated, and where tumor and surrogate tissues can be compared at different time points and with different drug doses. However, it must be remembered that species differences may decrease the relevance of such evaluations in the clinic; therefore, some clinical validation and confirmation will be required. Obtaining tissue for study, particularly in phase I trials with heterogeneous patient populations and different tumor types, is generally very difficult. These are complex issues which will demand continued attention.

Alternate phase I designs

The conventional phase I design uses three patients at each dose level until 1 of 3 have dose limiting toxicity, then up to three more patients are added. If none of the 3 additional patients have dose limiting toxicity, escalation continues. If one or more of the additional patients have DLT, and only 3 patients were evaluated at the next lower dose level, additional patients are added to the lower dose level. Dose steps are generally determined by the modified Fibonacci dose escalation scheme with escalations in decreasing increments. In theory, this approach should decrease the risk of overshooting the MTD as it is approached, but scientific justification for it is lacking. Inpatient dose escalation is frequently not permitted because of concerns about cumulative toxicity obscuring effects at a subsequent dose level. The impetus for developing alternate designs is the desire to treat as few patients as possible at biologically inactive dose levels.

Dr. Christian used an NCI database consisting of phase I studies of 9 drugs evaluated in 20 protocols with 756 patients treated between 1985 and 1995 to document that the standard approach, or minor variations of it, have proven safe with few drug related deaths (0.53%) [7]. However, it also demonstrated certain limitations. The majority of patients were treated with biologically inactive doses and had no opportunity for dose escalation, and relatively few patients were treated at the dose that was recommended for phase II evaluation. The overall response rate was only 0.93%. Dr. Christian also noted that the standard approach determines a dose for a population, rather than for individual patients, whose tolerance may differ substantially for some drugs. With few patients treated at the MTD, conventional phase I trial designs do not yield a good estimate of the dose at which some target probability of toxicity (for example, 25%, 33% or 50%) is achieved.

She summarized alternative designs that have been proposed (including pharmacology-guided, model-based and less complex approaches, such as dose
Many of the alternative dose escalation schemes begin with a more aggressive approach (for example, dose doubling). These approaches mandate switching to a more conservative, frequently standard approach, when some target is achieved (usually a defined incidence of some level of toxicity, sometimes a pharmacologic endpoint such as 40% of the AUC at the mouse LD10). Dr. Christian emphasized that this stop/switch point has not received adequate attention and used the NCI database to demonstrate how different stop/switch points would alter both trial safety and efficiency. She considered various possibilities including 'evidence of biologic activity' defined by first occurrence of grade 1 toxicity, first occurrence of grade 2 toxicity, second occurrence of grade 1, second occurrence of grade 2 toxicity, reproducible grade 1 or reproducible grade 2 toxicity and/or the first dose-limiting toxicity. If grade 1 toxicity is used, any alternate escalation scheme would have little impact, as switching would usually occur after the first or second dose level. An efficient approach which proved safe with computer simulations, used second occurrence of any grade 2 toxicity or first occurrence of any dose-limiting toxicity. For this determination, the following toxicities were not included: nausea, vomiting, fatigue, anorexia, anemia, alopecia, alkaline phosphatase elevation, fever and local reactions. In the NCI phase I Database, the second grade 2 toxicity occurred at median dose level 6, was 55% of the MTD and 9 times the starting dose. The first DLT occurred at median dose level 6.5, was 80% of the MTD and 10 times the starting dose.

In the NCI database, only 72 of 560 potentially eligible patients (13%) received escalated doses. Where intrapatient dose escalation was used, there was no evidence that it impacted adversely on the trial or on individual patient safety. Cumulative toxicity does not appear to be a valid reason to prohibit intrapatient dose escalation, as it occurs rarely.

Dr. Christian presented a preliminary proposal for phase I trial design developed through use of computer simulations in collaboration with her NCI colleagues Drs. Richard Simon, Boris Freidlin, Larry Rubinstein, and Susan Arbuck. The model permitted variation of several parameters, including interpatient variability in toxicity (a little or a lot), dose-toxicity curve (shallow or steep), and cumulative toxicity (none or a lot). Using these parameters, several phase I designs and dose escalation schemes were evaluated, each with computer simulations of 5000 phase I trials, to determine the number of patients enrolled, and the incidence of none or mild, moderate, and severe and unacceptable toxicity. The designs, include the standard design with 3 patients at each dose level, with addition of 3 more if 1 of the first 3 experience DLT, and designs with variations of one patient per cohort until the second grade 2 toxicity or the first DLT. Other variations included use of either 40% or 100% dose increments for the initial escalation portion of the trial. The impact of intrapatient dose escalation and use of data from all administered courses was also evaluated. Based upon simulations, and review of the NCI database, two novel phase I trial designs were proposed. Both use one patient per cohort until occurrence of second grade 2 or first dose limiting toxicity. Then standard cohorts (3 to 6 patients) are used. One design uses 40% dose increments throughout and considers only first course toxicity in decision rules, in contrast to another design which permits dose doubling in the initial escalation phase, and also incorporates information from all treatment courses, not just the first. For both designs, intrapatient dose escalation is permitted. The intrapatient dose escalation rule permits escalation for a patient who had grade 0 or 1 toxicity during the previous course. The dose is de-escalated if grade 3 or worse toxicity occurred during the previous course. Patients are not escalated to a dose at which 2 previous patients experienced grade 3+ toxicity. In addition, patients are de-escalated from a dose at which 2 previous patients experienced grade 3+ toxicity in that course (or earlier). The adoption of intrapatient dose escalation, and toxicity data from each course, provides the possibility for increasing the precision of the phase II dose estimate with model based schemes.

Although experience remains limited and implementation has proven problematic, alternative dose escalation schemes that attempt to minimize patient undertreatment and reduce the duration of phase I trials, are useful. Methods to accomplish these objectives include use of 1 patient per cohort until evidence of a biological effect, practical switching rules, larger dose increments, adoption of intrapatient dose escalation, and perhaps incorporation of methods to improve or standardize drug exposure. Another goal is to increase the precision of the phase II dose estimate which might be accomplished with use of model based schemes and/or pharmacologically based dosing.

**Commentary**

Dr. Christian's extensive review provides justification for evaluation of alternate dose escalation approaches, especially if we maintain the standard approach for determining the starting dose. (If less conservative starting doses are used, a more conservative dose escalation scheme would be required.) Several of the proposed alternate approaches include one patient cohorts for the first phase of the trial. Although modeling studies and retrospective evaluation using the NCI database suggest that this approach is safe, the heterogeneity of phase I patient populations is well known and patient selection, which is known to affect toxicity results,
would be expected to impact upon both toxicity and efficiency of a trial.

The standard escalation design, despite its limitations, has proven safe with few toxic deaths. However, to avoid treating patients at biologically inactive doses, and based upon reassessment of standard approaches and modeling studies of new approaches, the NCI designs and others should be evaluated promptly and systematically, in order to document that the stated objectives can be achieved with the additional goal of ensuring patient safety.

**Responses in phase I trials**

Traditionally many patients enrolled in phase I studies have had extensive prior therapy, extensive disease, and frequently, other poor prognostic factors. Some have not had measurable disease. For these reasons, although antitumor activity is always recorded in phase I studies, lack of clinical response in phase I trials does not alter development plans for the agent. If toxicity is acceptable, phase II trials are performed to screen for activity using a disease oriented approach.

However, Dr. Von Hoff demonstrated that response in a phase I trial is a useful predictor for the future of a new agent [8]. This determination was based upon a literature review of 12,160 patients treated in 428 phase I trials with 228 agents over the past 25 years. Each agent eventually approved for marketing produced at least one response during phase I evaluation. However, responses in phase I trials did not guarantee that the drug would subsequently prove efficacious. Using standard response criteria, the overall response rate was 6.8%. One percent of patients had complete responses with a median duration of 6 months and 5.8% had partial responses with median duration of 3 months. The median number of responses in phase I trials was 5 among drugs that were eventually approved and 1 among the others.

With regard to whether drug development should be is continued if no responses are observed in phase I trials, Dr. Von Hoff emphasized the need for sufficient numbers of patients treated at adequate doses among the phase I trials of a new agent. His suggested requirements include at least 2 phase I trials, with at least 68 patients, of whom 20 patients were treated at doses ≥80% of the MTD.

Responses in phase I trials are important for the patient who is responding. Observing a response in a phase I trial is also an important finding for the conduct of the trial. It makes it easier to talk to a new patient about the trial. It boosts accrual and improves the psyche of the phase I team. It is, therefore, important to maximize the probability of achieving a response. This objective might be accomplished by tougher selection criteria for proposing new drugs for clinical investigations, improved patient selection criteria that could be guided, for example, by results in the human tumor cloning assay, more efficient dose escalation schemes which would allow fewer patients to receive biologically ineffective doses of the new agent, and earlier incorporation into combination chemotherapy regimens even in the absence of single agent experience, (for example, a phase I trial for gemcitabine + a farnesyl transferase inhibitor in pancreatic cancer). In addition, it may be important to identify or develop indicators of antitumor activity other than a documented tumor shrinkage. A clinical benefit response could include pain relief, improvement in performance status or weight gain. Better tumor markers, PET scanning and other techniques could also be employed. Alternate indicators of antitumor activity will likely prove more important as novel agents undergo development, since effective differentiating agents, antiangiogenic agents or farnesyl transferase inhibitors, for example, are expected to be cytostatic and unlikely to cause dramatic tumor shrinkage.

**Commentary**

The question as to whether drug development should be discontinued if no responses are seen in phase I trials has important implications. As Dr. Von Hoff emphasized, a new agent should not be dropped after phase I evaluation because of lack of antitumor activity without examining in detail how the phase I trials were conducted. In addition to his requirements for sufficient patients among all phase I trials of a new agent who receive adequate doses, since phase I trials are not substitutes for phase II trials, one must also ensure that enough patients had measurable or evaluable disease, with a sufficiently broad distribution of tumor types, to ensure opportunity to document activity. Furthermore, if alternate dose escalation schemes are adopted, fewer patients may be required for phase I trials. If the criteria suggested by Dr. Von Hoff (20 patients treated at 80% or more of the MTD) are not met once phase I trials are completed (or if there is not adequate representation of different tumor types, for example), his suggestions mandate consideration of whether the phase I experience should be expanded prior to making a decision about proceeding to phase II evaluation. If the phase I experience was expanded for this purpose, it would be reminiscent of the broad phase II trial that included patients with multiple tumor types, a study design that is no longer employed.

**Conclusion**

Drs. Eisenhauer and Cavalli organized this workshop with the intent of beginning a dialogue on critical issues in cancer drug development, and with the hope that discussion would stimulate thoughtful reevaluation of current practice and lead to investigation of promising new approaches. They and the NCI-EORTC New
Drug Development Organizing Committee deserve congratulations for initiating an important activity that should lead to continuing reassessment, and to evaluation of alternative approaches that may prove advantageous.

References


Received 14 May 1996; accepted 22 May 1996.

Correspondence to:
Dr. Susan G. Arbuck
Investigational Drug Branch
CTEP
National Cancer Institute
6130 Executive Blvd., Room 715,
Rockville, MD 20852
USA