SPECIAL ARTICLE

Phase I Clinical Trial Design in Cancer Drug Development

By E.A. Eisenhauer, P.J. O'Dwyer, M. Christian, and J.S. Humphrey

Abstract: The past decade has seen the publication of a number of new proposals for the design of phase I trials of anticancer agents. The purpose of these proposals has been to address ethical concerns about treating excessive numbers of patients at subtherapeutic doses of a new agent and to increase the overall efficiency of the process while enhancing the precision of the recommended phase II dose. In early 1998, a workshop of phase I investigators was held under the sponsorship of Bristol-Myers Squibb Pharmaceutical Research Institute (Wallingford, CT) to review the experience to date with novel phase I methodologies, with a particular focus on their efficiency and safety. This report summarizes the material presented. It was concluded that for phase I trials of antineoplastics (cytotoxics), which begin at 0.1 mouse-equivalent LD10 doses, evidence to date suggests that the historic approach of using a modified Fibonacci escalation and three patients per dose level is not necessary and is seldom used. One patient per dose level and more rapid escalation schemes, both empirically based and statistically based, are commonly used with apparent safety. There remain questions, however: Which of the dose escalation schemes is optimal? Are there alternatives to toxicity as a phase I end point, and will these end points be reliable in defining active doses? Answering these questions in a reasonable time frame will be important if new anticancer agents are not to suffer undue delays in phase I evaluation.

J Clin Oncol 18:684-692. © 2000 by American Society of Clinical Oncology.

P HASE I TRIAL design in cancer therapeutics has changed little in 20 years. Unlike most therapeutic areas, there are two goals in cancer trials: precise definition of an optimal (recommended phase II) dose and safe treatment of the individual patient at doses that are close to therapeutic. The latter concern has led to proposals for more rapid and efficient dose escalation schemes in recent years: dose escalation on the basis of pharmacokinetic observations and statistically based models. The former uses toxicologic projections that are based on pharmacologic information from preclinical models. The latter are driven by accumulating patient observations that refine a model pre-

dictive of the optimal dose. As indicated in a recent publication,[1] neither approach has gained wide usage in the field.

In 1996 at the 9th National Cancer Institute (NCI)/ European Organization for the Research and Treatment of Cancer Symposium on New Drugs in Cancer Therapy, a workshop was held to examine both standard and novel approaches to phase I trial design. As was described in a comprehensive summary of the workshop,[2] at least for cytotoxic agents, there was a clear interest in modifications to standard phase I design to make them more efficient, minimizing the numbers of patients treated at nontoxic dose levels and maximizing the precision of phase II dose recommendations. Thus higher starting doses, recruitment of only one patient per dose level, and accelerated dose escalation schemes were discussed in detail. At approximately the same time, a review of the phase I trial literature from 1993 to 1995[1] concluded that, despite the publication of several novel dose escalation approaches over the preceding decade, few were being used in practice. With this background, a colloquium was organized in early 1998 to bring together investigators with experience in the use of novel phase I trial methodologies to review the relative efficiencies, inefficiencies, and safety of such methodologies. In particular, the following questions were to be addressed with respect to patient safety and trial efficiency:

1. Should higher starting doses be used? If so, when and with what restrictions?
2. Is the entry of one patient per dose level appropriate? If so, when and with what restrictions?
3. Are novel dose escalation schemes being used? Are

they more efficient than modified Fibonacci? Is there a dose escalation method that is preferred? If so, on what basis?

GOALS OF PHASE I TRIALS OF CYTOTOXIC AGENTS

Dr Elizabeth Eisenhauer introduced the meeting by reviewing the primary goal of phase I studies: to determine the appropriate dose for phase II evaluation. In the case of cytotoxic agents, an assumption is made that the higher the dose, the greater the likelihood of efficacy. Because most of these agents exhibit a dose-toxicity relationship, dose-related toxicity is regarded, in general, as a surrogate for efficacy: the highest safe dose is assumed to be the one most likely to be efficacious. This view creates a situation where the achievement of significant, but reversible, toxicity is desirable. Those toxic effects that by nature of their severity limit further dose escalation (dose-limiting toxicity; DLT) are defined in advance in phase I trials, and the maximum-tolerated dose (MTD) is defined as that dose producing a certain frequency of DLT within the treated patient population. At the same time, the investigators conducting these trials have a responsibility to limit the risk of individual patients to unacceptable levels of toxicity. Historically this has been accomplished using a "conventional" phase I trial design conducted by selecting a safe starting dose of 0.1 MELD10 (one tenth of the mouse equivalent LD10) or lower, accruing patients in cohorts of three, and escalating the dose according to a modified Fibonacci sequence in which ever higher escalation steps have ever decreasing relative increments (eg, dose increases of 100%, 65%, 50%, 40%, and 30% to 35% thereafter). The dose escalation is continued in cohorts of three patients until the MTD is reached. The next lower dose level is the recommended phase II dose (RPTD).

LIMITATIONS OF STANDARD PHASE I DESIGN AND POTENTIAL SOLUTIONS

The major problems raised with respect to the "standard" phase I approach described above have been the following:

- *Ethical*: With three patients entered per dose level, substantial numbers of patients are treated at doses that are retrospectively predicted to be nontherapeutic. Although the overall response rates in phase I trials are low, the majority occur within 80% to 120% of the recommended phase II dose.[3] These considerations raise ethical pressures to treat fewer patients at the initial dose levels in the absence of toxicity.
- *Efficiency*: The Fibonacci escalation scheme may result in quite lengthy trials in which dozens of patients and many months are required to determine the phase II dose. With a plethora of molecularly defined antitumor

targets and an increasingly clear description of tumor biology, there are now more antitumor candidate therapies requiring phase I study than ever. Unless more efficient approaches are undertaken, phase I trials may be a rate-limiting step in the process of evaluation of novel anticancer agents.

The variables in phase I design that, if modified, may lead to solutions to one or both of these problems are three: (1) the starting dose, (2) the number of patients per dose level, and (3) the method/rapidity of dose escalation. Increasing the starting dose could potentially reduce the trial length and limit the number of patients receiving nontoxic drug doses. Fewer patients per dose level would also limit numbers exposed to low doses and might shorten trial length if recruitment of three patients per level is rate limiting. Finally, more aggressive escalation in the initial portion of the trial or escalation targeted to the estimated MTD could also shorten the length of a trial as compared with Fibonacci escalation.

Within the colloquium, all of these approaches were discussed. Participants also noted that an important principle in phase I design was the protection of patients from exposure to unacceptable levels of risk (toxicity), so evaluation of novel methodology must include not only a measure of its relative efficiency but also a determination of its relative safety. Finally, all agreed that any new phase I design must permit precise determination of the phase II dose. To complete a trial quickly and with few patients receiving nontoxic doses is not helpful if the recommended dose is subsequently shown to be inaccurate.

STARTING DOSE LEVELS FOR PHASE I STUDIES

As noted above, preclinical studies in mice define a dose at which approximately 10% of the mice die (the murine LD10). One tenth of the murine equivalent LD10 (0.1 MELD10), expressed in milligrams per meters squared, has historically been a safe starting dose in humans when toxicologic studies in a second species (eg, rat, dog) do not show substantial differences in the dose-toxicity relationship. Under conditions in which murine toxicity and data from a second species show no marked interspecies differences (or where mouse was the most sensitive of the two species), Eisenhauer asked the question of whether higher starting doses can be safely used. To address this, a review of compounds evaluated in phase I trials over the past few years was undertaken. Agents selected for review were cytotoxic drugs studied as single agents in an initial phase I trial performed to determine the MTD. All published trials of such agents were included, provided their starting dose was based on murine LD10 information. With the knowledge of the "true" MTD determined in each trial, the

Table 1. Agents/Trials Selected for Starting Dose Review: 14 Agents Studied in 21 Trials

Agent	No. of Trials
DUP 937	2
Ormaplatin	2
Adozelesin	3
CPT-11	1
CB10-277	1
EO9	1
Bryostatin	1
PZDH	3
JM216	1
PZA	1
FCE 23762	1
KW 2149	1
Penclomedine	1
RP49532A	2

Table 3. Ratios of Human/Murine Toxicity in Phase I Trials of 71 Agents

Basis for Starting Dose	No. of Agents	MTD/0.1LD10* (mg/m²)		FTD/0.1LD10† (mg/m²)	
		Median	Range	Median	Range
Murine toxicology	57	20	0.5 to 248	8	0.25 to 127
Nonmurine toxicology	14	11	0.25 to 93	5	0.25 to 12

*Ratio of human MTD to one tenth of the murine LD10.
†Ratio of first toxic dose in clinical trials to one tenth of the murine LD10.

number of dose-escalation steps to achieve MTD was calculated based on the actual starting dose of 0.1 MELD10 and theoretical starting doses of 0.2 and 0.3 MELD10. To assure comparability, dose escalation was performed in all cases according to the modified Fibonacci scheme. Major end points of the exercise were to determine if increasing the starting dose shortened dose escalation and trial length and to assess the safety of the use of higher starting doses. With respect to the latter, a trial was arbitrarily considered unsafe if three or fewer dose levels (including the starting dose) were required to reach MTD. This was based on the notion that escalation schemes that reached MTD (a dose considered to be too toxic) in three or fewer steps would occasionally be expected to result in serious toxicity at the starting dose level. Fourteen agents studied in 21 trials met the criteria for inclusion (Table 1). The major results are listed in Table 2. For this group of agents and trials, a starting dose of 0.1 MELD10 led to a median of seven dose levels to attain the MTD (range, four to 14 dose levels). A starting dose of 0.2 MELD10 yielded a median of five dose levels (range, three to 11 dose levels) to attain the MTD, and when the starting dose level was increased to 0.3 MELD10,

Table 2. Safety and Number of Dose Levels in 21 Trials (14 Agents) With Varying Starting Doses

Starting Dose*	No. of Dose Levels to Reach MTD		No. of Unsafe† Trials (n = 21)	No. of Unsafe† Agents (n = 14)
	Median	Range		
0.1	7	4-14	0	0
0.2	5	3-11	5	2
0.3	3	2-9	11	6

*Expressed as a fraction of MELD10.
†Unsafe defined as three or fewer dose levels (including starting dose) needed to reach MTD.

a median of three dose levels were required to reach the MTD (range, two to nine dose levels). If an unsafe trial was defined as three or fewer levels to attain the MTD, then 0 of 21 trials (0 of 14 agents) were considered unsafe with 0.1 MELD10 starting dose, five of 21 (two of 14 agents) were considered unsafe at the 0.2 MELD10, and 11 of 21 (six of 14 agents) were considered unsafe at the 0.3 MELD10 starting dose. No agent at any of the three starting doses would have entered a phase I trial at a dose level above the MTD.

In summary, Eisenhauer concluded that a starting dose of 0.2 MELD10 may be a reasonable approach to shorten duration of phase I trials and limit the number of patients who are treated at very low doses when agents under study show no significant interspecies variation in toxicology. To increase the safety of this approach, she suggested this should only be undertaken if three patients per dose level are enrolled and escalation of the dose in 100% increments is limited to only one or two dose steps.

Dr Jaap Verweij subsequently addressed the question of which preclinical toxicologic studies predict a safe starting dose. He reviewed 100 drugs from the literature, for which 71 drugs had full data available, and compared the following ratios: (1) the ratio between the human MTD and 0.1 mouse LD10 (both expressed in milligrams per meters squared), and (2) the ratio between the first toxic dose (FTD) in humans and 0.1 mouse LD10. For 57 agents, the starting dose was based on mouse toxicology data. For these agents, the median ratio of the MTD/0.1LD10 was 20 (range, 0.5 to 248), with fludarabine being the only drug for which the ratio was less than 1 (Table 3). For these 57 agents, the median ratio of the first toxic dose to 0.1 mouse LD10 was 8 (range, 0.25 to 127). For 14 agents for which the starting dose was based on nonmouse toxicology, the median ratio of MTD/0.1 LD10 was 11 (range, 0.25 to 93). The median FTD/0.1 LD10 ratio was 5 (range, 0.25 to 12), with docetaxel being the only drug for which the ratio was less than 1. In general, the ratios were similar for antibiotics, antimetabolites, alkylating agents, metals, and antimitotics. Topoisomerase inhibitors had slightly lower ratios. Compounds that are considered to be outside of these classes had higher median MTD/0.1 LD10 ratios.

When Verweij examined the 14 agents with marked interspecies differences in toxicology in greater detail, he noted that in all cases, rat toxicology was adequately predictive of a safe starting dose. Dog toxicology data did not give additional information to that provided by the rat toxicology data.

He concluded that a limited toxicology model is feasible for developing anticancer agents, such that starting doses for phase I studies can appropriately be defined using mice and rat toxicology only. Furthermore, if toxicology in mice and rats is similar, then a starting dose of 20% to 40% of the mouse LD10 would be safe in most cases. This was very much in agreement with Eisenhauer's data. In fact, the European Organization for Research and Treatment of Cancer has adopted the rat as the second species for toxicity studies.

PHARMACOKINETICS IN PHASE I TRIALS AND PHARMACOLOGICALLY GUIDED DOSE ESCALATION

Dr Jerry Collins presented examples of how pharmacokinetics have materially aided in the development of drugs. He began by reviewing the experience with pharmacologically guided dose escalation,[4,5] which sought to escalate doses in a rapid fashion by doubling dose to a target area under the curve value derived from murine pharmacokinetic information. This approach was useful in saving dose levels (compared with the modified Fibonacci) in phase I trials of several agents over the past decade, including flavone acetic acid, hexamethylene bisacetamide, piroxantrone, and iododeoxydoxorubicin (I-Dox). However, the methodology has not been widely adopted for a number of primarily practical reasons. These include (1) logistical difficulties in obtaining "real-time" pharmacokinetic results, which are required to determine the safety of the subsequent escalation, (2) problems in extrapolating preclinical pharmacokinetic data to phase I studies of differing schedules, and (3) interpatient variability in results.

However, in addition to aiding dose escalation, pharmacokinetic studies contribute to other important end points. The magnitude of interpatient variability in drug metabolism is often identified in the course of phase I trials. Furthermore, identification and characterization of drug metabolites in humans is an important by-product of such studies. For example, I-Dox is not metabolized in the mouse but is metabolized in humans. The metabolite is active, and, therefore, dosing of I-Dox in humans is effectively treatment with the I-Dox metabolite, iododoxorubicinol. Similarly, paclitaxel has different metabolites in humans than in rats. In the example of penclomedine, pharmacokinetics demonstrated that neurotoxicity was related to the accumulation of a parent compound, whereas the demethylated

metabolite is just as cytotoxic but without neurotoxic side effects. Thus, under conditions of increasing pressures to improve the efficiency of drug development, pharmacokinetics from phase I trials can add information.

NOVEL DOSE-ESCALATION METHODS: STATISTICALLY BASED METHODS

Recently, several statistical methods have been developed for phase I dose escalation, with their primary goal being to shorten the duration of phase I trials and to enhance the precision of the phase II dose recommendation. Like the traditional modified Fibonacci, these methods use toxicity, and specifically DLT, as the end point of the trial. A mathematical function is created that describes the hypothesized relationship (curve) between the incidence of DLT and dose. This curve is reasonably predicted to assume a sigmoid shape that can be generically described by a "logit function" and for which the MTD must be estimated first. As information regarding the occurrence or absence of toxicity accumulates from the trial, the original estimate of the MTD is updated to more accurately fit the hypothesized curve to the actual data. Under these types of trial designs, the occurrence of toxicity results in an adjustment of the curve to match the probability that one is now approaching the MTD. Conversely, the absence of toxicity results in adjustments of the curve to match the probability that one is not yet at the MTD. Therefore, the occurrence of no DLT in several sequential patients results in a statistical prediction that the dose can be more rapidly escalated in a safe manner.

Two such statistical approaches were discussed. In the first of these, Dr André Rogatko presented a dose-escalation scheme that controls the probability that a patient will receive an overdose (Escalation With Overdose Control, or EWOC).[6,7] In this method, each dose level is selected such that the probability that the dose exceeds the MTD is ≤ a prespecified value. At the time of dose assignment, the dose level is selected by computing the most likely curve describing probability of DLT versus dose based on the experience in previous patients. At one extreme, the failure to observe toxic effects at any preceding dose level results in a prediction that one can escalate dose more rapidly. At the other extreme, the appearance of DLT in every other patient results in a statistical prediction that one is at the MTD. Rogatko presented data from two phase I trials of combination cytotoxic therapy to demonstrate how this method resulted in an ability to calculate the most likely value for the MTD together with the confidence intervals around that dose, corresponding to a 95% confidence interval that the true MTD lies within a certain dose range. The major advantage of this design as that the method is statistically designed to converge towards the MTD from

doses below the MTD, and it provides a confidence interval for the MTD by the end of the phase I trial. In addition, the dose escalations are chosen in such a way as to control the probability of delivering a dose that causes DLT. Although no examples were presented for which this method was used in the initial clinical trials of a new cytotoxic agent, it might be anticipated that this method would save dose-escalation steps as compared with the modified Fibonacci.

A second statistical approach was presented by Dr Eric Rowinsky, who described the application of the San Antonio version of the modified Continual Reassessment Method (mCRM)[8-10] in several phase I trials carried out at the University of Texas in San Antonio, TX. This method, like the EWOC method, constantly modifies the predicted function describing the dose-toxicity curve based on toxicity experience of all patients entered onto the trial at the completion of each dose level, resulting in updated predictions of the MTD. Also, like the EWOC method, the objectives of the mCRM are to reduce the number of patients treated at dose levels that are not likely to be efficacious, optimize evaluation of dose levels that are likely to be clinically relevant, and improve estimates of the MTD. As for other methods, the MTD is defined as the dose at which a certain percentage of patients would have DLT. To use the mCRM, the investigator must estimate a dose-response curve before the trial and also predict, on the basis of preclinical toxicity data and experience with agents in that class, an estimate of the MTD. In the mCRM, few dose levels are studied thoroughly. Most dose levels have a single patient. Therefore, the patients enrolled onto the study must be appropriate and representative. The small number of patients per dose level limits the dose-related pharmacokinetic data that can be obtained. In addition, the mCRM relies primarily on acute toxicities to predict the MTD, and chronic toxicities are less easily factored into the assessment. In fact, this is not different from other phase I designs in which acute rather than delayed effects are study end points.

Under the mCRM, a conservative starting dose is selected, such as 0.1 or 0.2 of the MELD10. A pre-estimation of the sample size is required. A dose-toxicity model must be selected, and an estimate of the MTD must be selected. As used in San Antonio, moderate toxicity grades, chronic toxicity, and patient characteristics are factored into the statistical modeling (detailed method in Table 4).

Rowinsky presented six trials conducted with the mCRM. The first was a trial conducted with AN-9, a butyrate-differentiating agent. After 16 accrued patients, mCRM dose escalation had permitted increasing the dose from 0.047 g/m^2/d up to 1.875 g/m^2/d. A modified Fibonacci approach, if conducted with three patients per dose level,

Table 4. San Antonio Application of the mCRM

Is drug or drug class appropriate for mCRM?
- Dose-toxicity relationship in preclinical studies
- Qualitative toxicity profile
- Interspecies difference (preclinical)

Starting dose based on traditional criteria
- One to three patients at starting dose
- Depends of knowledge of drug or class

Statistician and clinician interaction pretrial to define:
- Estimation of sample size
- MTD defined (20% to 30% incidence of DLT)
- Estimation of MTD and posterior distribution
- Estimation of dose-toxicity curves

Patient treatment:
- Patient should typify targeted population (phase II)
- Full observation period between patients

Preselection of doses (escalation scheme)
- Example: 100% to 33% with one patient per dose level until DLT or moderate toxicity

Possible scenarios (dose assignment)
- If mild toxicity, may treat additional patients. Next dose escalation is 50% to 100% (if initially 100%).
- If moderate toxicity, may treat additional patients. Next dose escalation is 33% to 100% (if initially 100%).
- Up to 10 patients treated at projected phase II dose.
- After each patient is declared as having DLT or no DLT, an updated dose-toxicity curve or an updated estimate of the MTD will be calculated.
- Therefore, a current updated estimate of the MTD incorporating information from patients earlier in the trial is available to select the dose level for subsequent patients entering the trial.
- Calculations for all possible updated estimates of the MTD are performed ahead of time.
- After the first patient is enrolled but before the patient experiences DLT or no DLT, the updated MTD is calculated assuming both acceptable and unacceptable toxicity.
- If the calculations show that more than one patient is required before dose escalation, then two patients can be enrolled at the same time.

would have advanced the dose to only 0.4 g/m^2/d after 16 accrued patients. Similarly, an mCRM trial of CGP 48664 resulted in an escalation of the dose from 3.6 to 202.8 mg/m^2/d after eight dose levels and 13 accrued patients. Trial results were also presented for MGI-114, a DNA-interactive cytotoxic agent, and MDL 101, a ribonucleotide reductase inhibitor. In each case, the mCRM resulted in more rapid dose escalation than conventional modified Fibonacci trial design. Subsequent to this workshop, a comprehensive review of the San Antonio mCRM experience has been reported.[11] In this it was concluded that although fewer patients were treated using the mCRM as compared with the number of patients who would have been treated with traditional Fibonacci escalation, the time taken to complete studies was not substantially altered, likely because of the necessary observation time between dose

Table 5. Accelerated Titration Designs*

Design Type	No. of Patients per Dose Level	Increments Between Dose Levels (%)	Intrapatient Escalation	Stop/Switch Rule§ Invoked for First Course or Any Course Toxicity
1A	3	40	No†	NA
2B	1	40	Yes‡	First
3B	1	100	Yes	First
4B	1	100	Yes	Any

*Modified from Simon et al.[12]

†No within patient escalation. De-escalate if grade 3 or worse toxicity in previous course.

‡Escalate if grade 0-1 toxicity at previous course. De-escalate if grade 3 or worse toxicity at previous course.

§Stop/switch rule: After one occurrence of DLT or two occurrences of grade 2 toxicity, the design reverts to 40% increments between dose levels and three to six patients per dose level. Design 2B and 3B invoke the switch if these are first-course events. Design 4B invokes the switch with any course events.

levels. It should be noted that several variants of the mCRM design have been used in practical applications, making it somewhat difficult to discuss the merits or limitations of this approach.

Although there has been no direct comparison of the mCRM and the EWOC methods in a clinical trial, simulation studies conducted by Babb et al[6] suggested that use of the EWOC method resulted in overdose of a smaller proportion of patients, exhibition of fewer DLTs, and estimation of the MTD with a slightly lower average bias and marginally higher mean error as compared with the mCRM. Whether these differences would produce meaningful differences in practice is unknown and awaits comparative clinical applications of both methods.

ACCELERATED TITRATION DESIGNS

On behalf of colleagues at the NCI, Dr Michaele Christian presented three accelerated designs that are intended for use in phase I trials of drugs which have not been used previously in humans.[12] The goal of their work was to develop dose-escalation rules that would limit the numbers of patients receiving low and probably subtherapeutic doses of new agents. They developed the dose-escalation rules by fitting a stochastic model to data from 20 phase I trials involving the study of nine different drugs. They then simulated new data from the model with the parameters estimated from the actual trials and evaluated the performance of alternative phase I trial designs on this simulated data. In comparison with standard practice (listed in Table 5 as design 1A), each of these designs includes only one patient per cohort until one patient experiences DLT or two patients experience grade 2 toxic effects. In design 2B, a similar escalation sequence to design 1A is used for the incremental increase of doses. (Note that in their simula-

tions, 40% escalation steps were used in both designs 1 and 2 rather than the more typical modified Fibonacci sequence.) In design 2B, intrapatient dose escalation is permitted and the DLTs and/or grade 2 toxicities that invoke the stop/switch rule are only those experienced during the first course. In contrast, designs 3B and 4B have escalations of 100% between dose levels until a stop/switch point is reached, after which dose escalations are 40% between dose levels. As in design 2B, the 3B and 4B designs permit intrapatient dose escalation, and the stop/switch rule is invoked for a first-course (3B) DLT or any course (4B) DLT or the second incidence of grade 2 toxicity. In all cases, the stop/switch rule dictates that the trial design revert to 40% dose escalation steps, with three to six patients per dose level. The use of toxicity from all courses, as well as grade 2 toxicity data, would be expected to provide more toxicity information on which to base a recommended dose at the conclusion of the trial. Christian presented a summary of the simulated data for each of these trial designs that showed that the average number of patients required for a phase I trial is reduced from 39.9 for a conventional design (1A) to 24.4 patients per trial with design 2B and 21.2 patients per trial with design 4B. Although design 2B results in a reduction of the number of patients required compared with design 1, design 2 required slightly more cohorts as a result of occasional overshooting of the targeted MTD. Because design 2 had slightly more cohorts per trial, design 2 may not offer any savings in the time required to complete the study over design 1A in situations where eligible patients are readily available. In such a situation, it would take little more time to place three patients on a dose level than to place a single patient on a dose level as in design 2B. This design would, however, offer an advantage in terms of the total number of patients required to complete the trial. Conversely, design 4 resulted in a reduction in the number of cohorts primarily as a result of the 100% increments in dose from dose level to dose level. The average number of patients with grade 4 toxicity as their worst toxicity increased from 1.9 for design 1 to 3.0 for design 2B and 3.2 for design 4B.

Dr Pat Lorusso compared the 1A and 2B designs in two otherwise similar trials of the same agent, KRN 5500, a spicamycin derivative. The conventional (1A) trial design is being carried out at the Dana-Farber Cancer Institute and the accelerated titration design (2B) is being carried out by Lorusso at Wayne State University. Both trials used a modified Fibonacci escalation (unlike the simulations referred to in the published article, in which 40% increments only were used), with the only difference being the number of patients per dose level. Over a similar period of time, the 2B design allowed the study of eight dose levels (15

patients) as compared with three dose levels (12 patients) in the 1A design. This example suggested that the 2B design, even without more aggressive dose-escalation steps than standard, has resulted in much time being saved by simply reducing the number of patients per dose level.

More information on the savings and safety afforded by these designs are desirable, and Christian stated an interest on the part of the NCI for investigators to use the 2B and 4B accelerated titration designs where appropriate. It is known that many research groups use a variation of these approaches (ie, double the dose until toxic events dictate a more conservative approach), but many such trials continue to enroll three patients per dose level.

COMMENTARY AND DISCUSSION

Dr Hilary Calvert provided some comments on the subject of phase I design, drawing on his experience with a number of methods of dose escalation. Although advantages to enrolling one patient per dose level have been well articulated in terms of speed of study completion and reduction in the numbers of patients receiving nontoxic doses, he noted that such an approach limits the quantity of information available from a phase I trial on interpatient variability in toxicity and pharmacokinetics. In fact, the underlying assumption that permits the entry of one patient per dose level is that dose, and not the patient, is the primary determinant of toxicity. At least for some classes of agents (eg, antifols) this is a risky assumption. The expansion of dose levels near or at the MTD, as is the case with some of the new designs, helps to address the issue of interpatient variability. However, it is only in phase I trials that dose-related pharmacokinetic effects can be studied over a broad range of doses, and if this is an important element of evaluation of the drug, then sufficient patients must be recruited at each level to make it feasible. In addition, Calvert pointed out that end points other than toxicity (eg, measures of target effect) are of interest and that more than one patient per level may be required to adequately assess such end points.

Dr Mark Ratain suggested that beyond specific accelerated titration and statistical designs, general principles and common sense are of importance: not all designs may fit all situations. The general principles he noted are to begin at a safe starting dose, to minimize the number of patients treated at the subtoxic dose levels, to escalate dose rapidly in the absence of toxicity, and to escalate dose slowly in the presence of toxicity. In addition, he advocated expanding the recommended phase II dose level to include 20 to 30 patients to permit pharmacokinetic and pharmacodynamic characterization before embarking on phase II studies. He presented results of several trials designed with conserva-

tive starting doses, which permitted single patients per dose level with 100% dose escalation in the absence of any toxicity. After the occurrence of grade 1 toxicity, the dose levels are expanded to three patients per dose level, with only a 50% incremental increase in dosage in advancing to the next dose level. For the occurrence of grade 2 or 3 toxicity, there would likewise be three patients in the study per dose level, with a 25% incremental increase in dose to the next level. In the presence of grade 4 toxicity, the dose level would be expanded to six patients or more with a less than 25% dose escalation for the next dose level. Data from two ongoing trials revealed rapid escalation in the absence of toxicity for one novel agent. In essence, these trials are similar in design to those termed "accelerated titration" by the NCI group. In these examples, conventional modified Fibonacci with three patients at all dose levels would have led to much longer and larger studies.

In conclusion, a tension that is inherent in phase I design comes from the competing goals of limiting the number of patients who are exposed to nontherapeutic doses of the new drug and ensuring the safety of those enrolled in the trial. As reviewed in this meeting, options for limiting the number of patients who are exposed to nontherapeutic doses include increasing the starting dose, limiting the number of patients accrued on each dose level, and accelerating the dose escalation process. Increasing the starting dose from 0.1 to 0.2 MELD10 would seem to be safe for trials of those agents that show no interspecies differences in toxicology. Similarly, entering one patient per dose level at low doses is a growing practice that seems to be safe, although it may limit ancillary information that may be obtained from the study, such as pharmacokinetic data. Several approaches to altered dose escalation were reviewed. Both statistical methods and accelerated escalation methods seemed to result in more rapid achievement of toxic doses as compared with the modified Fibonacci method, thus limiting patient numbers treated at nontoxic doses. Higher starting doses and newer dose escalation methods also both lead to more rapid completion of phase I trials with fewer dose levels. Although recruitment of one patient per dose level may also shorten trial duration, this modification alone would not offer an advantage to the more standard three patients per level in the presence of readily available patients.

Participants at this colloquium were able to reach certain specific conclusions related to the three major questions posed in the introductory section of this article:

Should higher starting doses be used? If so, when and with what restrictions?   Although, as noted previously, it seems that starting doses of 0.2 MELD10 are safe under some conditions, this does not offer a large advantage over the usual starting dose of 0.1 MELD10 when it is combined

with a more aggressive dose-escalation scheme. Thus the standard starting dose need not be modified, because all agreed that more aggressive dose escalation was now the norm (vide infra).

Is the entry of one patient per dose level appropriate? If so, when and with what restrictions? The use of one patient per dose level, at the lowest doses of a phase I trial, has become a frequent and apparently safe approach. An exception to this might be foreseen in the circumstance in which it is suspected that interpatient variability in toxicity will be likely. In this setting, the observation of no toxicity in one patient may be misleading. Certain agents such as antifolate compounds may be expected to produce greater interpatient variability.

Are novel dose-escalation schemes being used? Are they more efficient than modified Fibonacci? Is there a dose-escalation method that is preferred? If so, on what basis? All participants in this colloquium stated that they no longer would routinely use modified Fibonacci dose escalation when designing a phase I trial. All of the methods presented seem to offer an advantage in shortening trial length in comparison with the modified Fibonacci, with no suggestion that they decrease safety. What is not clear is whether any one of these methods provides an advantage over the others. Statistically based methods in which the next dose level is assigned by renewed estimates of the MTD may ultimately prove most useful in determining a more precise estimate of the MTD (and thus recommended phase II dose). However, evaluation of these (in particular the mCRM) against conventional designs has been complicated by the fact that many variants of this design have been used in practice. Strategies for dose selection using mCRM are missing from most of the literature descriptions of its use. In fact, many have followed the same approach as that described by Rowinsky: to define planned dose levels arbitrarily in advance of the trial using the statistical estimates of MTD to assure that sufficient patients have been enrolled at each level and that the next planned level will not exceed the MTD estimate. It is difficult to ascertain if any have used the mCRM to derive the next dose level. Phase I designs in which continued rapid escalation by 100% steps with a switch to more conservative escalation when certain toxic (or pharmacologic) criteria are met can be written into the protocol and implemented without relying on repeated recalculation of the dose-toxicity curves. Any design used may, at the conclusion of the trial, have incorporated a statistical analysis of all toxic events to refine further the estimate of the recommended phase II dose. Comparisons of the performance and ease of application between these novel methods are desirable.

It should be noted that both statistically based and accelerated dose-escalation approaches might lend themselves to study of novel agents where nontoxicity end points are evaluated. Under these circumstances, measures other than toxicity (eg, target inhibition) could be substituted in the dose–end point relationship. Whether small numbers of patients per dose level would be appropriate to include in such trials would depend two factors: whether preclinical data suggested that the dose-effect curve for these end points is as steep as is usually the case for dose-toxicity effects and whether limited interpatient variability is anticipated. If both of these conditions were met, then the designs described in this article could be reasonably applied (although it remains debatable whether as few as one patient per dose level would allow adequate assessment of the end point). If the conditions were not met, then larger sample sizes might be required and different designs used, as is the case with the evaluation of noncancer compounds for which the first phase I studies are conducted in healthy volunteers. These issues are of importance to consider given the growing number of novel agents that may not be expected to exhibit the type of toxicity/efficacy relationship seen with traditional cytotoxic anticancer drugs. It might be possible to conduct early investigation of such agents in volunteers, but if cancer patients are to comprise the phase I population, then concerns about excessive numbers of individuals enrolled at subtherapeutic dose levels once again become relevant.

The consensus of meeting participants was that for phase I trials that begin at 0.1 MELD10, the approach of accruing three patients per dose level with dose escalation based on a modified Fibonacci sequence should no longer be considered the standard design. A number of questions and challenges remain for those involved in designing and conducting phase I trials of anticancer agents:

1. Which of the various types of dose escalation schemes described meet the criteria for safety, efficiency, and precision of phase II dose estimation when assessed in a large number of phase I trials?

2. Given the volume of novel targeted anticancer agents that enter early clinical study, are there alternatives to toxicity that can be used to serve as end points in phase I dose escalation?

3. Can nontoxicity end points be efficiently incorporated into the novel escalation designs that have been described in this commentary? Are new end points reliable in identifying active doses?

Answering these questions in a reasonable time frame will be important if new anticancer agents are not to suffer undue delays in evaluation in phase I trials.

REFERENCES

1. Dent SF, Eisenhauer EA: Phase I trial design: Are new methodologies being put into practice? Ann Oncol 7:561-566, 1996

2. Arbuck SG: Workshop on Phase I design: Ninth NCI EORTC New Drug Development Symposium, Amsterdam, March 12, 1996. Ann Oncol 7:567-573, 1996

3. Von Hoff DD, Turner J: Response rates, duration of response and dose response effects in phase I trials of antineoplastics. Invest New Drugs 9:115-122, 1991

4. Collins JM, Zaharko DS, Dedrick RL, et al: Potential roles for preclinical pharmacology in phase I trials. Cancer Treat Rep 70:73-80, 1986

5. Collins JM, Grieshaber CK, Chabner BA: Pharmacologically guided phase I trials based upon preclinical development. J Natl Cancer Inst 82:1321-1326, 1990

6. Babb J, Rogatko A, Zacks S: Cancer phase I clinical trials: Efficient dose escalations with overdose control. Stat Med 17:1103-1120, 1998

7. Rogatko A, Babb J: Escalation With Overdose Control (EWOC). Http://www.fccc.edu/users/rogatko

8. O'Quigly J, Pepe M, Fisher L: Continual reassessment method: A practical design for phase I clinical trials in cancer. Biometrics 46:33-48, 1990

9. O'Quigley J: Estimating the probability of toxicity at the recommended dose following a phase I clinical trial in cancer. Biometrics 48:853-862, 1992

10. Goodman SN, Zahurak ML, Piantadosi S: Some practical improvements in the continual reassessment method for phase I studies. Stat Med 14:1149-1161, 1995

11. Eckhardt SG, Siu LL, 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. Proc Am Soc Clin Oncol 18:163a, 1999 (abstr 627)

12. Simon R, Freidlin B, Rubinstein L, et al: Accelerated titration designs for phase I clinical trials in oncology. J Natl Cancer Inst 89:1138-1147, 1997