COMMENTARY

Statistical and Ethical Issues in the Design and Conduct of Phase I and II Clinical Trials of New Anticancer Agents

*Mark J. Ratain, Rosemarie Mick, Richard L. Schilsky, Mark Siegler**

The development of new anticancer agents is a complex, stepwise process proceeding from discovery to demonstration of antitumor activity in preclinical models and evaluation of normal tissue toxicity prior to initiation of clinical trials. The purpose of the initial clinical trials (phase I) is to define the toxic effects of the agent and the recommended dosage for subsequent testing (*1*). The vast majority of drugs that reach phase I studies go on to phase II testing, which is aimed at finding evidence of efficacy in human cancer (*1*). Rarely, drugs are withdrawn at the completion of phase I testing, usually due to unpredictable and/or severe nonhematologic toxic effects that may be irreversible (*2-4*).

Although there is general agreement regarding the overall research goals of phase I and phase II studies of new anticancer agents, two problems emerge: 1) The research goals may differ from the patient's goals, and 2) there is no consensus on how to achieve the researcher's goals in the most efficient and ethically appropriate way. In this commentary, we discuss both statistical and ethical issues of early anticancer drug development and suggest an alternative framework that may improve the protocol design and address the ethical issues of phase I and phase II clinical trials.

Phase I Trials: Statistical Issues

Background

The major scientific objective of the phase I trial is to determine the recommended phase II dose of the drug being studied. Thus, phase I trials address an estimation problem rather than the testing of a hypothesis. The theoretically optimal dose for any individual patient is the dose resulting in the maximally achievable antitumor response with an acceptable level of toxicity. Since one cannot predict efficacy prior to treatment, the theoretically optimal individual dose is the maximal dose that does not exceed an acceptable level of toxicity. However, phase I trials are designed to determine the recommended phase II dose for a population, not an individual. Thus, since there is usually substantial interpatient variability in toxic effects, the recommended phase II dose will always be an imprecise estimate of the optimal dose for the individual patient (*5*). As a result, many phase II trials utilize intrapatient dose modifications based on observed toxicity in an attempt to treat each patient at the optimal dose.

A second problem with the design of phase I trials is that the optimal dose is usually administered to only a minority of the patients treated in the phase I trial. The reason for this occurrence is complicated. For determination of the recommended phase II dose, it is standard research practice to treat cohorts of patients beginning at a dose that is believed to be nontoxic and then escalating the dose for patients in successive cohorts until defined grades of toxic effects are observed. Table 1 lists the variables defined by the investigator who uses the traditional phase I design and some of the commonly considered alternatives.

Statistical issues in phase I trials were initially outlined by Schneiderman (*6*). Other investigators have subsequently considered such issues with regard to the traditional design, such as the optimal starting dose (*7-10*) and whether real-time pharmacokinetics can more quickly lead to dose escalations (*11-13*). However, investigators have not carefully focused on criteria for termination of phase I studies, which should only occur when the recommended phase II dose has been adequately defined.

If all three patients of a hypothetical cohort experience dose-limiting toxicity, most investigators would agree that the recommended phase II dose has been exceeded. However, even in that setting, the investigator must acknowledge that the 95% exact confidence interval for the incidence of

*See "Notes" section following "References."

Table 1. Variables to be defined for traditionally designed phase I trials*

Variable	Example	Suggested alternatives
Starting dose	$1/10$ LD_{10}	Higher
Patients per nontoxic dose level	3	Fewer
Definition of DLT	Any grade 3 toxicity	Grade 3-4†
Subsequent dose levels	"Modified Fibonacci"	Pharmacokinetically guided dosing
Patients per toxic dose level	6	Greater
Definition of RPTD	<2 of 6 patients with DLT	<3 of 6 patients with DLT
Patients at RPTD	6	Much greater

*DLT = dose-limiting toxicity; RPTD = recommended phase II dose.
†Grade 3 organ toxicity is dose-limiting, but grade 3 myelosuppression or nausea/vomiting is not.

dose-limiting toxicity is 37%-100%. If the target rate of dose-limiting toxicity is 50%, then the recommended phase II dose may be underestimated. Many investigators will not accept a 33% rate of dose-limiting toxicity (two of six patients) at the recommended phase II dose. Yet, the 95% exact confidence interval for this dose-limiting toxicity rate is quite broad (6%-73%).

Alternatives to "Standard" Phase I Trial Design

Because of the relatively small number of patients who actually receive the recommended phase II dose, several authors (*5,14-22*) have suggested substantial modifications to the traditional phase I trial design. These modifications have been aimed at efficiently determining the recommended phase II dose, while minimizing the risk of both undertreatment and excessive toxicity. Since most responses in phase I studies occur at dose levels that are 80%-120% of the recommended phase II dose (*23*), such modified trial designs would result in a greater response rate during phase I. For example, several investigators (*14,16*) have suggested "up and down" designs, in which the dose for each patient is assigned on the basis of the experience in the patient(s) most recently treated and may be adjusted either higher or lower. These designs are relatively simple modifications to the traditional design, but they base major protocol decisions on relatively small numbers of patients.

In contrast to the traditional design and these modifications of it, O'Quigley et al. (*17,18*) proposed a Bayesian approach, the continual reassessment method. A Bayesian method begins with assumptions about the main end points of the study, known as "priors," which are derived from the investigators' prior observations and/or beliefs based on their own experience and that of others. In a phase I trial that uses a Bayesian approach, information from preclinical studies and/or clinical studies of similar drugs is used to make an educated guess regarding the dose–toxicity curve and the recommended phase II dose. Patients are then treated at the current estimated recommended phase II dose, and these estimates are continually updated to reflect the accumulating dose–toxicity data. When the sample size (determined in advance) has been reached, the final estimate of the recommended phase II dose is made from all available data. Although it is desirable to meet the ethical goal of treating each patient at the hypothesized recommended phase II dose, there is a danger that patients would be exposed to highly toxic doses if major errors were made in the assumptions about the end points of the study.

We recently proposed a design (*22*) that is a bridge between the traditional design and the Bayesian design suggested by O'Quigley et al. (*17,18*). The scheme utilizes a cohort-based escalation approach similar to that used in traditional phase I studies. As dose–toxicity data are accumulated, a pharmacodynamic model is fit to the data. Model-guided dosing commences only after a prespecified number of patient cohorts have been treated and evaluated for toxic effects and after a dose–toxicity relationship is statistically defined. Thus, model-guided dosing is less dependent on assumptions and more dependent on observed toxic effects. Unlike the design proposed by O'Quigley et al. (*17,18*), each patient is not treated at the estimated recommended phase II dose, but instead at the dose calculated to yield a target nadir white blood cell count for the individual patient. This design is only useful when myelosuppression is dose limiting, although modifications to include graded nonhematologic toxicity would be feasible. Based on computer simulation studies, this method performs better than the traditional design because it requires entry of fewer patients in the phase I study and yields more reproducible estimates of the recommended phase II dose for an average patient. As pointed out by Mathe and Brienza (*24*), there may be substantial interstudy variability in the recommended phase II dose.

Phase I Trials: Ethical Issues

Background

Investigators conducting phase I trials must adhere to the ethical norms of clinical research (*25-30*) and therefore encounter a number of potential ethical issues:

1) Minimizing patients treated at ineffective doses;
2) Minimizing patients treated at toxic doses;
3) Historically low probability of response in phase I trials;
4) Unknown toxicity and benefit of new agent; and
5) Difficulty in obtaining true informed consent in vulnerable patient populations.

Some have argued that since phase I trials are designed to define the qualitative and quantitative aspects of toxicity,

patients should only enter such trials "for the benefit of future cancer victims" (*26*). While this situation may apply to normal volunteers without disease who enroll in phase I trials of some drugs, it surely does not apply to phase I trials of anticancer agents, since both the control and treatment arms of the trial only involve patients with cancer (*31*). Thus, cancer treatments in phase I trials are always administered with therapeutic intent, even though the major scientific objective of a phase I study is to define toxicity. Most importantly, the ethical concerns related to therapy for patients with advanced cancer do not depend on whether or not the treatment is administered in the context of a clinical trial or whether such a trial is phase I, phase II, or phase III.

Risk–Benefit Ratio

Patients offered treatment in the context of a phase I trial are informed that the purpose of the study is to find the optimal drug dose and that some patients will be treated at a dose that is too low or too high (i.e., too toxic). They are also informed that there is a small possibility that the treatment will be beneficial. The probability of benefit, however, does not depend on the type of trial, but rather on the disease being treated, the drug being tested, and the dose of the drug that is administered. For example, there is probably a greater response rate for lymphomas in phase I trials than for pancreatic cancer in phase III trials, since the latter is rarely responsive to any therapy.

Phase I studies often raise ethical concerns because of the perception that the patient is exposed to drug toxicity without being offered a reasonable expectation of benefit (*30*). Another common concern is the potential for undertreatment of those patients enrolled at dose levels that are eventually determined to be subtherapeutic. These concerns have led some to suggest the use of intrapatient dose escalation as a solution to this problem (*24*). With this approach, patients who experience little or no toxicity at their entry dose receive increased doses on subsequent cycles. However, intrapatient dose escalation only allows for the possibility that patients who do not rapidly succumb to their disease might eventually receive a "therapeutic" dose of the new agent. This approach does not recover the time lost while the patient received ineffective therapy, and it does not clearly increase the chance of tumor response, since the tumor is likely to be resistant even to doses that result in toxic effects unless such dose escalation is marked (>100%). In addition, it increases the period of time that a patient remains on a particular study, possibly limiting opportunities for treatment in subsequent trials. Traditionally designed phase I studies do have a unique property: At least one patient must experience dose-limiting toxicity for the trial to be completed. A statement to this effect is now included in the consent form of phase I cancer trials at our institution.

Another major concern in phase I trials is the low probability of response. There is no question that, on the basis of bidimensional tumor measurements, most patients entered in phase I trials do not achieve a partial or complete response (*32*). However, there is the small possibility of a substantial and rewarding response, as was observed during the phase I trials of cisplatin (*33*) and paclitaxel (Taxol) (*34*). In fact, all 17 drugs that were eventually marketed after beginning National Cancer Institute-sponsored phase I testing from 1970 through 1983 manifested activity in at least one phase I trial (*22*). Thus, the potential range of outcomes, both beneficial and harmful, is extremely broad.

At study initiation, the type and degree of toxicity likely to occur are unknown, although toxicology studies are often predictive for drug toxicity to organs. Patients must take risks that cannot be easily measured. Since both the probability of benefit and toxicity are unknown, estimating the therapeutic index is impossible. At low dose levels, there is a low probability of both toxicity and benefit. At dose levels near the recommended phase II dose, there is a high probability of severe toxicity and a maximal probability of benefit (*23*). The magnitude of the probability of benefit cannot be ascertained until phase II trials have been completed.

Informed Consent

An issue of great concern to both investigators and institutional review boards is the consent process (*27,29,30,35,36*), with the suggestion by some of procedural safeguards such as third-party consultation with noninvestigators (i.e., primary care physicians) as patient advocates (*30*). The major obstacle to true "informed consent" is that the investigator often has little information. Consent forms usually contain a litany of possible side effects, when in fact the clinical experience has been almost exclusively without toxicity. Furthermore, centers with a major interest in phase I trials may have 10 or more trials available to an individual patient. Should patients be offered all such trials as alternatives? Since there may be substantial deficiencies in participants' perceptions of the consent process in even less complex trials (*36*), subjects may be confused regarding what is involved in any specific trial.

Freedman (*35*) has suggested a cohort-specific approach to phase I cancer studies. In this approach, the informed consent process is different for the first patient entered in the study compared with that used to enroll patients after significant toxicity has been noted. This strategy requires a dynamic consent form, which varies according to the patient cohort. One expects that investigators would ordinarily communicate this information to the patient verbally, but inclusion of this information as an appendix to the standard consent form would ensure that the patient is informed of the actual experience to date.

Phase II Trials: Statistical Issues

Background

Phase II trials are generally studies with no control group that are aimed at estimating the antitumor efficacy of a new agent in a particular disease. The design of such trials has been discussed extensively in the literature; the discussion has recently focused on issues of sample size and hypothesis

testing as a basis for guidelines relating to early discontinuation of therapy if there is adequate evidence of inefficacy (*15,37-51*).

In contrast to phase I trials, phase II trials test a hypothesis: "New drug X is active against disease Y." The investigator must still define "active" as well as select which subset of patients with disease Y will be studied. For example, the investigator may wish to determine whether drug X has a 20% response rate in women with metastatic breast cancer refractory to doxorubicin. In the simplest design, based on binomial probability, 14 patients would be evaluated initially. If no patient responds, the investigator may conclude with 95% certainty that drug X has a response rate less than 20% in this patient population. If at least one patient responds, a second cohort of patients would be treated to better estimate the response rate. The greater the number of patients in this second cohort, the more precisely one can estimate the response rate.

Alternative Designs

A number of variations to this design have been suggested. Fleming (*43*) proposed testing two hypotheses simultaneously, one for a minimally acceptable response rate and one for a maximally unacceptable response rate (i.e., the highest response rate an investigator would be willing to miss). Simon et al. (*45*) suggested randomized phase II trials. This design is similar to the Fleming design in that it tests two (or more) hypotheses simultaneously: a) drug A is active against disease Y, and b) drug B is active against disease Y.

End Points

Virtually all phase II trials use a common end point, the objective response rate, which is defined as the percentage of assessable subjects who demonstrate a partial or complete response. Is screening new agents for a 15%-20% objective response rate the best approach? One could argue that we are using needlessly strict criteria and that we should simply be looking for improvement in quality of life or decrease in the rate of tumor growth, possibly by utilizing sequential bidimensional measurements (*52*). Clearly, the end point should be based on the patient population to be studied.

Phase II Trials: Ethical Issues

Probability of Response

Ethical issues in phase II trials have received little attention in the past. Whereas it is commonly known that very few patients achieve partial or complete responses during phase I studies (*32*), little concern has been expressed among investigators regarding the generally poor results of phase II trials. Marsoni et al. (*53*) tabulated the results of all National Cancer Institute-sponsored phase II trials from 1970 to 1985 and found a single active drug from among 42 phase II trials for colon cancer and from among 33 phase II trials for non–small-cell lung cancer. In a separate analysis (*54*), the overall response rate for new agents in phase II trials in non–small-cell lung cancer was 4%, which is less than the overall objective response rate in phase I trials of 6% reported by Von Hoff and Turner (*23*). Thus, one can conclude that the vast majority of patients with advanced solid tumors that are either refractory to standard therapy or for which no standard therapy exists will not respond to any treatment, whether it be noninvestigational, phase II, or phase I.

Informed Consent

One ethical issue associated with phase II trials is the construct used for hypothesis testing, i.e., attempting to disprove inactivity. The null hypothesis being tested in phase II trials is that "Drug X has less than a Z% objective response rate in disease Y." Thus, responses disprove the investigator's null hypothesis, and, usually, an expansion of the trial is then required to confirm the response rate. But how do investigators view their evolving data from phase II trials? What do we tell the 14th patient if all of the prior 13 have failed to respond? Do we say, "If you don't respond, we will be 95% certain that drug X has less than a 20% objective response rate in disease Y"?

Sordillo and Schaffner (*55*) have previously addressed this issue and have recommended that a statement be added to the consent form regarding the lack of activity to date. But do we really need the 14th patient? The investigator has already concluded, albeit with only 94% certainty, that drug X is inactive (<20% objective response rate) in disease Y. But if we accept 13 patients as the final sample size, then what do we tell the 13th patient if none of the first 12 responded? This recursive reasoning can be iterated back to the beginning of the study!

Vulnerability of Subjects

Patients entering phase II trials are as vulnerable as those entering phase I trials. Whereas patients may often select a phase I trial from multiple studies available at a major center, investigators have been cautioned against conducting more than one phase II trial for the same patient population, because of the risk of bias in patient selection in individual studies. Thus, patients offered a phase II trial at a particular center may feel that they have no good alternatives, regardless of the data accrued to date.

Patient Selection

As investigators recognize the problems with current phase II designs, new ethical issues must be addressed. Currently, an important issue under debate is whether to include in phase II trials previously untreated patients with tumors that are sensitive to chemotherapy but incurable, such as metastatic breast cancer or extensive small-cell lung cancer (*56-61*). This issue was recently considered in detail by Moore and Korn (*62*).

Comparison of Ethical Issues for Phase I and Phase II Trials

Both phase I and phase II trials place the investigator/physician under "ethical stress" (*26*); in both types of trials, the statistical and clinical objectives are in conflict. The conflict in phase I between the scientific goals of defining toxicity, recommended dosage, and clinical pharmacology and the therapeutic goal of antitumor effect is obvious to many and is generally recognized by both investigators and Institutional Review Board chairs (*30*). The conflict in phase II is less obvious, depending on the extent of accrual to the study and the results in the assessable patients to date (*55*).

Patients may enroll themselves in clinical trials for altruistic reasons, but most do so primarily in the hope of therapeutic benefit (*30,36*). Patients seek the benefit but may in turn benefit from the hope itself. Those patients who enroll in phase I trials are usually aware of the lack of knowledge of efficacy regarding drug X at dose Z in disease Y. However, one must also be concerned about the consent process in phase II trials. It is not clear that patients are told that "drug X has been shown to be safe, but there are no data yet to suggest that it is effective in disease Y." Even if there are no clinical data, the patient might feel that drug X is likely to be effective. It is even less clear that the patient will be informed of negative data as the trial evolves.

The decision to enroll in a clinical trial should include an assessment of the probability of both incremental benefit and harm. A major difference between phase I and phase II trials is in the assessment of toxicity. It may be paradoxical to some that most patients in phase I trials experience less drug toxicity than those in phase II trials. This paradox occurs because most patients in phase I studies are undertreated (*23,30*), whereas patients in phase II trials are generally treated at a dose that results in moderate to severe toxicity. Thus, on the average, toxicity in phase I studies is less than in phase II, but there is a much greater degree of variation.

Most patients in phase I studies have a low probability of both benefit and toxicity. Their major costs are the inconvenience of being in the study and "opportunity risk" if there are other possible treatments. In contrast, most patients in phase II trials have a significant risk of experiencing toxicity and an unknown likelihood of benefit, at least at study initiation. From the standpoint of the patient and treating physician, the late stages of a phase I trial, in which the dose being used is near the recommended phase II dose, are virtually equivalent to the early stages of a phase II trial, because the dose and toxicity have been identified and the benefit is unknown.

Considerations for the Future

Phase I Trials

The traditional cohort design is clearly imperfect, and it is important to consider implementation of new study designs, such as the design proposed by O'Quigley et al. (*17,18*) or the model-based dosing design that we have suggested (*22*). Phase I studies should attempt to define more than recommended phase II dose, emphasizing instead the development of a preliminary pharmacodynamic model to be tested prospectively in phase II trials (*63*), particularly in conjunction with limited sampling strategies based on prior pharmacokinetic studies (*64-66*).

How much patient autonomy should there be in phase I studies? The opportunity for a patient to choose among multiple phase I trials is one example. A more difficult autonomy issue is whether patients should have a say in the dose of the drug they wish to receive. If the investigator does not know the optimal dose, might it be reasonable to treat patients willing to take greater risks with higher doses? Suppose 10 dose levels have been defined for drug X, beginning at the standard starting dose (one-tenth of the dose lethal to 10% of treated animals). Patients might be given the choice of receiving the current dose level N, which has been incompletely evaluated, or dose level N + 1, which has not been evaluated. If a patient selected level N + 1 and experienced no toxicity, level N could be dropped. Then, patients would choose between level N + 1 and level N + 2. This approach could be further expanded to allow patients even greater autonomy, such as being allowed to choose level N + 2 or even N + 8. Thus, more aggressive patients could expedite completion of a study, assuming that the consent process can be completed in a fully informed, ethical setting. An important concern, however, is the patient's ability to make such difficult and complex decisions. Clearly, this would require a dynamic consent form (*35*) and probably third-party consultation with a noninvestigator as patient advocate.

Phase II Trials

Investigators must become more aware of the ethical issues in phase II trials. A dynamic consent process could be implemented to inform patients of results available to date, particularly when there is a high probability that the drug is ineffective (*55*). This may make single-institution phase II trials more difficult to complete unless patients are accrued rapidly. It would also require better communication among those institutions participating in multi-institution studies.

Phase II trials involving multiple institutions and multiple drugs (tested as single agents) would be a possible solution to this issue. In order to maximize patient autonomy, a randomized consent design could be used (*67*), or, alternatively, the patient and treating physician could jointly select an agent on the basis of data accrued to date. Since the purpose of phase II is to decide whether further testing is warranted, this self-selection by patients would not necessarily alter conclusions (*55*). Patients would be likely to "follow the winner" once activity is identified. If this activity is not confirmed, subsequent patients would be less likely to select the agent. In addition, patients would be guided by the type and degree of toxicity to be expected. The major drawback to such innovative designs, which require a high level of patient participation, is the concern that the complexity of the consent process could lead to the exclusion of those patients intellectually or emotionally unable to participate.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.