Novel endpoints and design of early clinical trials

W. R. Parulekar & E. A. Eisenhauer

National Cancer Institute of Canada (NCIC) Clinical Trials Group, Queen’s University, Kingston, Canada

Introduction

Early drug development trials in cancer medicine are designed to assess new therapies in a rapid and efficient manner. The traditional paradigm of dose selection based on maximum tolerable toxicity and screening for efficacy based on tumour shrinkage may not apply to new agents that target the extracellular environment or cell signalling molecules. This paper will review challenges involved in the evaluation of these agents and potential study designs to address these concerns.

Background: traditional design of early clinical trials of cytotoxic agents

Strategies for new drug evaluation until recently have focused almost exclusively on cytotoxic chemotherapy and were based on the following general observations:

- Preclinical studies have demonstrated that there is a direct relationship between chemotherapy dose and tumour cell kill.
- Preclinical and clinical studies have demonstrated that there is a direct relationship between chemotherapy dose and normal tissue toxicity.

The concept that toxicity is a useful means of drug dose or schedule selection represents an extrapolation of these observations and has formed the basis of the design of early drug development trials.

Phase I trials of cytotoxics

This type of study seeks to define the pharmacodynamic and pharmacokinetic profiles of new anticancer agents. The traditional design uses toxicological data from two animal species to determine a starting dose. Doses are escalated using one to three patients per dose level until a prespecified percentage of patients with dose-limiting toxicity is seen in an expanded cohort treated at a particular dose level. The major endpoint of the study is assessment of toxic effects, measured using standardised criteria such as the Common Toxicity Criteria (CTC) version 2.0. The primary study goal is determination of the maximum tolerated dose (MTD) [or more recently called maximal administered dose (MAD)] and the recommended dose for further single agent study which is generally deemed to be one dose level below the MTD.

Phase II trials of cytotoxics

The goal of the phase II trial is to further define toxicity and screen agents for future clinical development based on an estimate of antitumour effect. Tumour shrinkage, as measured by standardised objective response criteria such as those of RECIST [1] or World Health Organization (WHO) [2], is used as a surrogate marker of activity and is the only endpoint which has been validated to be predictive of drug efficacy (in some but not all cases) to date. These trials are usually conducted in two or more stages which allows for early termination of the study if little or no activity is demonstrated. At best, phase II trials are viewed as screening tools since the ultimate demonstration of a drug’s efficacy requires evaluation in a phase III trial.

Phase III trials

Phase III studies seek to define drug efficacy utilising definitive endpoints such as overall survival, cure rate or quality of life.

Issues with novel targeted ‘non-cytotoxic’ agents

The application of traditional phase I and II methodologies and endpoints to new agents that target signalling molecules or the extra-cellular milieu may be problematic since it is possible that neither toxicity nor tumour shrinkage are useful surrogates for dose selection or activity. The mechanisms of action and effects of these agents are thought to be distinct from cytotoxic chemotherapies which affect DNA or the process of cell division by interfering with DNA enzymes or tubulin. In contrast, in laboratory models, ‘cytostatic’ or ‘non-cytotoxic’ agents often appear to inhibit tumour progression rather than cause tumour regression although this effect may not be absolute. In addition, the concentration of drug needed for inhibition of a specific molecular target may not be one that induces significant toxicity.

Since the goals of early clinical studies remain the same regardless of the agent being evaluated, i.e. determination of the optimal drug dosage/schedule and screening for activity, modifications to trial design may be needed. These are discussed below.
Alternative endpoints/designs for novel agents

Phase I trials

Alternative measures for optimal dose selection for signalling molecule inhibitors include measurement of inhibition of subcellular targets or plasma/blood drug levels. The former may be problematic for a number of reasons. Measurement of effect on a specific target implies precise knowledge of the mechanism(s) of action of the new agent and prior validation of the endpoint in preclinical models, i.e. does target inhibition correlate with survival or lack of tumour progression? Other considerations include proper identification of the appropriate target. The complexity of cell signalling pathways suggests that this may be difficult since more that one molecule or signal may be important. In reality, this discussion becomes irrelevant if a valid and reliable target assay method is not available, as is often the case in the early stages of drug discovery. Another potential obstacle is the requirement for serial tumour biopsies for correlative study purposes; a process which is both invasive (especially in an experimental setting) and exclusive since eligible patients must have accessible tumours of an adequate size to participate in such a study. Finally, even if accessible tumour is present, sampling errors may occur due to the admixture of fibrotic and necrotic tissue invariably present in any tumour biopsy. A potential solution to this problem would be the use of normal tissue such as skin biopsies or peripheral blood cells to define the optimal dose. This method is reasonable providing that evidence exists demonstrating that target level changes in the substitute tissue parallel those in the tumour.

Alternatively, plasma or blood drug levels may serve as useful indicators of appropriate drug dose. Once again, this method is reasonable providing that preclinical data have demonstrated that the target blood or serum level correlates with tumour response.

Despite the relatively favourable side effect profile of these types of agents, toxicity remains an important consideration and may limit dose escalation of a cytostatic drug. Although dose escalation to maximal toxicity or tolerability may result in higher than necessary doses of this type of agent, this strategy minimises the chance that an ineffective low dose is chosen for further study.

Practically speaking, phase I trials incorporating multiple endpoints may be the most reasonable approach for evaluation of targeted, non-cytotoxic anti-cancer therapies and thus alternative endpoints to tumour shrinkage may be needed. One possible surrogate indicator of drug activity is change in tumour markers. Although this is appealing from a study design point of view, this technique remains largely unvalidated. For example, measurement of prostate-specific antigen (PSA) is widely regarded as an indirect measurement of prostate cancer tumour burden and changes in this marker have been incorporated in many study protocols as an endpoint of interest. Using in vivo and in vitro models, Thalmann et al. [3] demonstrated that changes in PSA expression may be discordant with tumour response. Androgen independent human prostate cancer cell line tumours were induced in mice and then treated with suramin. The anti-tumour and serum PSA/PSA mRNA effects of the drug were evaluated. Suramin failed to inhibit the growth of the androgen independent line in vitro and in vivo, despite significant decreases in PSA serum and PSA mRNA levels. This experience means that, should tumour markers be followed to assess activity of a new agent, researchers must first be assured that there is no effect of the drug on the marker level directly such as protein degradation or interference with assay methods, independent of tumour burden.

Another method of evaluating the activity of a cytostatic drug would be to perform a single stage phase II trial in which the endpoint of interest is progression-free survival. The potential usefulness of the agent could then be estimated by comparison with a historical cohort treated with the current standard of treatment or best supportive care. Issues of selection bias make this type of trial design highly subjective so it may be appropriate to choose an historical cohort from the same institution to maximise comparability of the two populations.

Alternatively, a study could be done in which patients act as their own controls. Examples of this type of trial include measurement of tumour growth before and after institution of the new therapy or randomising patients with stable disease on therapy to continuation or discontinuation of drug [4]. Progression-free survival would be the endpoint of interest. Issues of concern with these types of studies include the need to enter patients on study prior to failure of the standard regimen, physician bias in declaring progressive disease early on in the standard therapy to allow patients access to the experimental regimen, potentially large numbers of patients treated in the inception cohort to acquire sufficient numbers to randomise and the tendency for growth rates of tumours to vary spontaneously in the absence of treatment. Nonetheless, some researchers are incorporating randomised discontinuation designs into their evaluation of new therapeutics and the results of these trials are awaited with interest.

Others have proposed the use of a multinomial stopping rule as a means of efficiently evaluating new therapies in the

Phase II trials

As noted previously, inhibition of tumour growth may be the major mechanism of action of targeted, non-cytotoxic anti-cancer therapies and thus alternative endpoints to tumour shrinkage may be needed. One possible surrogate indicator of drug activity is change in tumour markers. Although this is appealing from a study design point of view, this technique remains largely unvalidated. For example, measurement of prostate-specific antigen (PSA) is widely regarded as an indirect measurement of prostate cancer tumour burden and changes in this marker have been incorporated in many study protocols as an endpoint of interest. Using in vivo and in vitro models, Thalmann et al. [3] demonstrated that changes in PSA expression may be discordant with tumour response. Androgen independent human prostate cancer cell line tumours were induced in mice and then treated with suramin. The anti-tumour and serum PSA/PSA mRNA effects of the drug were evaluated. Suramin failed to inhibit the growth of the androgen independent line in vitro and in vivo, despite significant decreases in PSA serum and PSA mRNA levels. This experience means that, should tumour markers be followed to assess activity of a new agent, researchers must first be assured that there is no effect of the drug on the marker level directly such as protein degradation or interference with assay methods, independent of tumour burden.

Another method of evaluating the activity of a cytostatic drug would be to perform a single stage phase II trial in which the endpoint of interest is progression-free survival. The potential usefulness of the agent could then be estimated by comparison with a historical cohort treated with the current standard of treatment or best supportive care. Issues of selection bias make this type of trial design highly subjective so it may be appropriate to choose an historical cohort from the same institution to maximise comparability of the two populations.

Alternatively, a study could be done in which patients act as their own controls. Examples of this type of trial include measurement of tumour growth before and after institution of the new therapy or randomising patients with stable disease on therapy to continuation or discontinuation of drug [4]. Progression-free survival would be the endpoint of interest. Issues of concern with these types of studies include the need to enter patients on study prior to failure of the standard regimen, physician bias in declaring progressive disease early on in the standard therapy to allow patients access to the experimental regimen, potentially large numbers of patients treated in the inception cohort to acquire sufficient numbers to randomise and the tendency for growth rates of tumours to vary spontaneously in the absence of treatment. Nonetheless, some researchers are incorporating randomised discontinuation designs into their evaluation of new therapeutics and the results of these trials are awaited with interest.
phase II setting [5]. This may have particular relevance for drugs whose main mechanism of action is induction of tumour stasis rather than shrinkage. The stopping rule in this type of trial incorporates objective response and early progression rates in a two-stage design. Using null and alternate hypotheses for both response and progression, decisions regarding early termination or continuation of the study to the next stage can be made based on various combinations of response and progression seen in the first cohort of patients. This study design is economical and evidence suggests that it may be more efficient than the Fleming or Gehan stopping rules in terminating a trial early due to drug inactivity [5]. Potential benefits to using the multinomial stopping rule include minimising patient exposure to inactive drugs and decreasing the probability of falsely rejecting an active drug since lack of early progression is considered a sign of drug activity. This approach has been retrospectively evaluated on a series of cytotoxic agents that are expected to produce tumour shrinkage. For non-cytotoxic agents, it may be speculated that the use of early progression alone could allow early decisions regarding activity: those agents that are not expected to cause response but rather growth delay, would be expected to show a decrease in the rate of early progression. If the trial shows the opposite, then the likelihood that the new drug will make a meaningful contribution to cancer treatment is low. Under these circumstances, early termination seems warranted. However, very few investigators have as yet applied this approach so there is no body of data to support its validity.

Randomized screening trials have been suggested by some as a means of overcoming study design limitations since there is no reliance on historical data and issues of bias are minimised [6]. When multiple new agents or different schedules of the same agent are involved in the trial, the objective is not to choose the therapy which is statistically superior in activity compared with the rest but one which is not likely to be inferior, i.e. ‘not pick the loser’ as opposed to ‘pick the winner’ [5]. In certain circumstances, randomised phase II trials can lead directly into phase III trials provided that this intention is clearly stated at the start of the trial, criteria for continuation are clearly defined in the protocol and there is sufficient patient availability to conduct a phase III study.

Phase III trials
The need to follow the traditional sequence of drug studies, i.e. phase I, followed by phase II and then phase III, may not be necessary with cytostatic agents. The strategy of moving from phase I directly to phase III may be justified if there is compelling preclinical evidence of drug activity, early stopping rules based on lack of drug effect are planned for the phase III trial (futility analysis) and the decision has been made to proceed to the phase III trial regardless of phase II data.

The conduct of phase III trials involving targeted non-cytotoxic therapy is no different than that involving cytotoxic agents since efficacy and safety remain the primary endpoints. An exception to this is the addition of correlative studies to the traditional phase III design as a means of increasing the understanding of drug effect(s) and tumour biology. Further, should the agent move from phase I directly into a phase III study, there is a need for interim stopping rules that could halt the study should the experimental arm show an adverse trend on outcome. In reality, this is little different from the concept of randomised phase II extending to phase III: in both cases there is a planned early review of patient data and an active decision must be taken on the basis of pre-determined rules to carry on accrual. The differences are perhaps more ones of semantics than in the study plan.

Examples
An example of the evaluation of non-cytotoxic agents in the National Cancer Institute of Canada Clinical Trials Group (NCIC CTG) is that of the matrix metalloproteinase inhibitor (MMPI) BAY 12 9566. Matrix metalloproteinases are secreted by both benign and malignant cells and are essential in normal tissue remodelling [7, 8]. Their importance in tumour growth relates to their ability to degrade the extracellular matrix, allowing for tumour cell invasion and endothelial cell migration. Matrix metalloproteinases are thought to be active in tumours of the breast, ovary, colon and lung and evidence suggests that there is a positive correlation between levels and aggressive tumour biology and the development of metastases [9, 10]. Preclinical evidence has shown that inhibition of these proteinases can result in inhibition of tumour growth and metastases thus providing the rationale behind development and use of inhibitors such as BAY 12 9566 [11, 12]. Studies in healthy volunteers have demonstrated a favourable side effect profile including headaches, drowsiness and mild gastrointestinal symptoms [13].

A phase I study was performed by the NCIC CTG to determine the MTD of BAY 12 9566 as well as characterise the pharmacokinetic and toxicity profile when administered orally in a dose escalation design [14]. Due to the relatively favourable side effect profile of the drug, dose selection based on achievable drug levels was included in the dose finding strategy. The target drug dose was defined as one that would result in a drug level 10 times that needed for efficacy in preclinical models. The drug was administered daily in a chronic fashion and discontinued for reasons of toxicity and/or progression. Pharmacokinetic sampling was done on days 1–5, 29 and 30. Response was assessed based on the WHO criteria [2]. At the conclusion of the study, MTD was not determined due to the lack of dose-limiting side effects. In addition, the target plasma levels as defined in the protocol could not be achieved despite increasing doses of drug. The choice of 800 mg p.o. b.i.d. as the recommended dose was based on patient compliance factors as well as the acceptable plasma levels obtained with this dose. No responses were seen but stability of disease was noted in a number of patients.
The decision to proceed with a randomised phase III trial in pancreatic cancer patients was based on the following rationale:

- There was strong preclinical data demonstrating tumour growth inhibition and reduction of metastases with BAY 12-9566.
- Phase I data in healthy volunteers as well as cancer patients confirmed good tolerability and achievement of therapeutic drug levels of this agent given in a daily oral fashion.
- Since little tumour shrinkage would be expected with this agent, results of a phase II study would not have an impact on decisions regarding further development of the drug and thus a phase II trial was judged to be unnecessary.
- New agents were needed in the treatment of pancreatic cancer since current therapies had little if any impact on the natural history of the disease [14].

The phase III trial accrued 277 patients from December 1997 to July 1999. Patients with locally advanced or metastatic pancreatic cancer were randomised to single-agent gemcitabine (1000 mg/m² i.v. weekly) versus BAY 12 9566, 800 mg p.o. b.i.d. The primary endpoint was overall survival and secondary endpoints included progression-free survival, quality of life, clinical benefit [15], response rates and pharmacokinetic analysis. Two interim analyses were planned. The first, performed after 30 patients had been randomised to each arm, verified that the activity of gemcitabine was as expected and that of the study arm was acceptable. The second analysis was scheduled after 140 deaths and involved a formal comparison of overall survival between arms. The study was closed after the second analysis revealed that the median overall survival was 6.4 months versus 3.2 months favouring the gemcitabine arm (P = 0.0001) [16].

This example illustrates many of the issues surrounding the evaluation of new agents that target the extra-cellular environment or cell signalling molecules. Dose selection in the phase I trial was on the basis of serum drug levels rather than toxicity since the favourable side effect profile precluded determination of the MTD. This was justified since the alternate endpoint (serum levels of drug) had been validated in preclinical models. Although the decision to proceed directly from a phase I to a phase III trial was made, early stopping rules were incorporated in the trial design to allow for timely termination of the study if significant differences in efficacy between the treatment arms were detected. Finally, despite promising preclinical and early clinical data, the determination of drug efficacy required evaluation using a standard randomised phase III trial design.

In addition to the example cited above, recently many other ‘non-cytotoxic’ agents have been explored in early clinical studies. Information from such work is now providing the beginnings of a clinical database from which we are learning several important lessons vis-à-vis early clinical trial design.

The first of these is that, despite the rational appeal of utilising measures of change in target effect in tumour samples, this remains problematic: while investigators may sample tissue during dose-seeking trials, this has not yet been the means by which doses for further study are determined. Toxic effects supplemented by information regarding blood levels of the drug remain the mainstay of dose selection (Table 1).

Furthermore, despite predictions to the contrary, some of the agents being evaluated are inducing objective tumour regression in single-agent studies. Agents targeting several members of the EGFR/ErbB family, such as OSI-774, ZD1839 and trastuzumab, have all been reported to show tumour regression in a small proportion of patients in single-agent phase II trials [16–21]. At least one of these, trastuzumab, has subsequently been shown to improve survival in combination with chemotherapy in a breast cancer study [22].

Thus it may be the case that as the database for non-cytotoxic anticancer drugs evolves and results of phase II and III studies are amassed, a pattern may appear that allows us to draw inferences about the type of phase II result which will predict for efficacy in the phase III setting.

**Summary**

In summary, the clinical development of non-cytotoxic targeted therapy requires a critical examination of existing models of phase I, II and III study design. The investigator must

---

**Table 1. Examples of endpoints/outcomes of non-cytotoxic agents in clinical trials**

<table>
<thead>
<tr>
<th>Agent</th>
<th>Target</th>
<th>Basis of dose selection in phase I</th>
<th>Phase II results/endpoints</th>
<th>Phase III results</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAY 129566</td>
<td>MMP</td>
<td>Blood levels (toxicity minimal)</td>
<td>Not done</td>
<td>No evidence of activity in several studies</td>
</tr>
<tr>
<td>BB2516</td>
<td>MMP</td>
<td>Toxicity and blood levels</td>
<td>Rate of rise in tumour marker showed dose effect</td>
<td>No clear benefit to date in multiple studies</td>
</tr>
<tr>
<td>ZD1839</td>
<td>EGFR</td>
<td>Toxic effects</td>
<td>Objective responses seen</td>
<td>Too early</td>
</tr>
<tr>
<td>OSI-774</td>
<td>EGFR</td>
<td>Toxic effects</td>
<td>Objective responses seen</td>
<td>Too early</td>
</tr>
<tr>
<td>RhuMAB HER2</td>
<td>HER2</td>
<td>Blood levels</td>
<td>Objective responses seen</td>
<td>Improved survival in combination with chemotherapy in randomised breast cancer trial</td>
</tr>
</tbody>
</table>

EGFR, epidermal growth factor receptor; MMP, matrix metalloproteinases.
carefully consider all aspects of study conduct including dose selection criteria, methods of determination/confirmation of drug activity and validation of new technologies and laboratory methods. Since the goals of phase I, II and III studies remain the same regardless of the agent being evaluated, modifications but not abandonment of traditional trial designs may be the most rational approach to this new challenge. Data emerging from early clinical trials of non-cytotoxic agents currently undergoing evaluation are expected to shed some light on optimal design and endpoint selection for these types of agents in the next decade.

References