VOLUME 26 · NUMBER 8 · MARCH 10 2008

JOURNAL OF CLINICAL ONCOLOGY

REVIEW ARTICLE

Review of Phase II Trial Designs Used in Studies of Molecular Targeted Agents: Outcomes and Predictors of Success in Phase III

Robert H. El-Maraghi and Elizabeth A. Eisenhauer

From the Royal Victoria Hospital, Barrie; and National Cancer Institute of Canada Clinical Trials Group, Queen's University, Kingston, Ontario, Canada.

Submitted July 18, 2007; accepted November 1, 2007; published online ahead of print at www.jco.org on February 19, 2008.

The National Cancer Institute of Canada (NCIC) Clinical Trials Group (CTG) is supported by a core grant from the National Cancer Institute of Canada with funds received from the Canadian Cancer Society. R.H.E.-M. was a Drug Development Fellow at NCIC CTG from 2005 to 2006 whose salary was supported by a Transdisciplinary Training Program Grant from the Canadian Institutes of Health Research as well as a Drug Development Fellowship grant from AstraZeneca.

Presented in part at the 42nd Annual Meeting of the American Society of Clinical Oncology, June 2-6, 2006, Atlanta, GA; and the National Cancer Research Institute Cancer Conference, October 8-11, 2006, Birmingham, United Kingdom.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Elizabeth Eisenhauer, MD, FRCPC, National Cancer Institute of Canada Clinical Trials Group, Queen's University, 10 Stuart St, Kingston, Ontario, Canada K7L 3N6; e-mail: eeisenhauer@ctg.queensu.ca.

0732-183X/08/2608-1346/$20.00

DOI: 10.1200/JCO.2007.13.5913

ABSTRACT

Purpose
Because the appropriate design and end points for phase II evaluation of targeted anticancer agents are unclear, we undertook a review of recent reports of phase II trials of targeted agents to determine the types of designs used, the planned end points, the outcomes, and the relationship between trial outcomes and regulatory approval.

Methods
We retrieved reports of single-agent phase II trials in six solid tumors for 19 targeted drugs. For each, we abstracted data regarding planned design and actual results. Response rates were examined for any relationship to eventual success of the agents, as determined by US Food and Drug Administration approval for at least one indication.

Results
Eighty-nine trials were identified. Objective response was the primary or coprimary end point in the majority of trials (61 of 89 trials). Fourteen reports were of randomized studies generally evaluating different doses of agents, not as controlled experiments. Enrichment for target expression was uncommon. Objective responses were seen in 38 trials; in 19 trials, response rates were more than 10%, and in eight, they were more than 20%. Agents with high response rates tended to have high nonprogression rates; renal cell carcinoma was the exception to this. Higher overall response rates were predictive of regulatory approval in the tumor types reviewed ($P = .005$).

Conclusion
In practice, phase II design for targeted agents is similar to that for cytotoxics. Objective response seems to be a useful end point for screening new targeted agents because, in our review, its observation predicted for eventual success. Improvements in design are recommended, as is more frequent inclusion of biological questions as part of phase II trials.

J Clin Oncol 26:1346-1354. © 2008 by American Society of Clinical Oncology

INTRODUCTION

Recently, there has been a veritable explosion of knowledge with respect to the molecular biology of malignancy. This has led to the identification of potential new targets for cancer therapy and, subsequently, to the rational design of agents created to affect those targets in a clinically meaningful way. However, as a growing number of agents targeting molecular pathways are tested in the clinic, there has been increasing pressure to rethink the standard drug development paradigm, specifically early trial design, to ensure that these promising new drugs are appropriately evaluated.

The ultimate goal of drug development in oncology is to identify new agents that provide a meaningful clinical benefit for patients, with the gold standard being the prolongation of patient survival. This is both a time- and resource-intensive process, with an estimated monetary cost to bring new drugs to market of $800 million (or more)[1,2] with equally important, although perhaps less definable, costs to participating clinical subjects. With this in mind, it is understandable that there is pressure to optimize early clinical trial design so as to minimize the resources expended on drugs that are likely to fail in later development.

The primary objective of phase II trials, regardless of the nature of the compound, is to screen for preliminary evidence of efficacy in a given tumor type. For cytotoxic agents, the standard approach has been to enroll small numbers of patients in a nonrandomized design, often of two or more accrual stages, and use objective tumor regression assessed by standard criteria[3,4] as the end point to identify drugs with potential efficacy. Retrospective

data support the use of the response end point for studies of cytotoxic agents[5-7] and the efficiency of the nonrandomized design. The inherent differences in the mechanism of action between traditional cytotoxic drugs and molecularly targeted agents, coupled with interest in increasing the reliability of phase II results in identifying truly active agents, have led to considerable discussion about the so-called traditional approach to screening trials, particularly in regards to patient selection criteria, end points, and study design.

With respect to patient selection, it is reasonable to expect that not all patients with a given tumor type will have similar levels of target protein activity or expression, and thus, efficacy of the targeted agent may vary according to which subpopulation is evaluated. Enriching the population to maximize possible activity could be achieved through restriction of study entry to those with a specific tumor histology or those whose tumors (over)express a molecular target. However, at the time of early clinical studies, such predictors of activity are by definition hypothetical, unless the agent affects a target that has had validated predictors identified through earlier clinical trials of other drugs affecting the same target.

The appropriate end point for phase II trials of targeted agents has also been debated.[8-12] Because many of these agents may affect tumor cells by reducing proliferation, rather than by causing cell death, the impact on tumor growth may be stabilization of disease or minor tumor shrinkage. Thus, it is argued that focusing only on objective response could result in overlooking some agents that could improve survival by causing disease stabilization. Indeed, in lung cancer, attainment of stable disease, in addition to responses, has been shown to contribute to improved survival.[13]

The design of phase II trials and the choice of which end point is to be used are closely linked; the use of a nonrandomized design in the traditional phase II trial is reasonable because objective responses are unlikely to be due to natural disease processes, rendering a control group unnecessary. However, even if response is believed to be a reasonable end point for trials of particular agents, the sample size of a nonrandomized trial may need adjustment. This is necessary if the hypothesized response rate of interest for the targeted agent is lower than what might be considered standard for cytotoxic agents, as would be the case if the trial population is unselected for predictive markers.

When objective response alone is not foreseen to be a useful end point, novel designs have been proposed for phase II screening trials. Included is a multinomial design in which decisions about early stopping and conclusions on activity are based not only on the number of responses seen, but also on the proportion of patients demonstrating early disease progression.[14,15] Another approach of interest is the randomized discontinuation trial where patients are treated with a new agent for a specified period of time, after which those with stable disease are randomly assigned to continue or discontinue therapy.[16,17] The end point for this type of trial is either time to event (eg, progression) or the proportion of patients progression free at a specific time point after random assignment. This design also provides information on the activity of the drug in terms of rates of response and progression at the end of the run-in phase. The randomized discontinuation design has been promoted to be of particular help in screening cytostatic drugs, such as molecular targeted agents, by permitting early assessment of whether delays in progression are related to treatment or disease and whether they are of sufficient magnitude to suggest that the drug may be effective.

The issues we have just highlighted illustrate the fact that numerous proposals regarding changes to traditional phase II trial design and end points have been made in recent years. However, there is little information on the impact of these discussions on the actual end points and designs used for the evaluation of novel targeted agents in phase II trials. In this article, we report the results of a review of phase II designs, end points, and outcomes of a series for targeted agents that have been studied in clinical trials in the last few years. In particular, we were interested in determining the planned design and end point(s) in these trials, what the observed results were, and, when possible, whether the results of phase II trials of individual agents were useful in predicting which agents achieve regulatory approval. Our interest in this last point speaks to the ultimate goal of phase II screening trials regardless of design: Is the output of the trials useful in identifying those agents likely to succeed in phase III trials?

METHODS

Agents

For this review, we focused on the 31 targeted agents that were the basis of a review of phase I trial design published in 2004.[18] These agents have distinct intracellular or extracellular targets in pathways and can be grouped into major classes on the basis of the chemistry of the agent (small molecule, antisense oligonucleotide, or antibody) as well as the expected molecular target (Table 1).[19-83] All of these agents completed phase I investigation long enough in the past that phase II trials, if undertaken, should now be reported.

Tumor Types and Search Strategy

We defined six common solid tumor types (breast, colorectal, lung, ovarian, prostate, and renal cell carcinomas) on which to focus. We performed a MEDLINE search for published articles of completed phase II single-agent studies for all of the 31 drugs limited to the six tumor types noted. Studies using the agents in combination and studies in progress were excluded from the review. The search terms included the name of the agent (including trade name, if applicable) and the molecular target, with the limit of clinical trials, phase II. A cutoff date of April 30, 2006 was used for retrieving publications.

In addition to full publications, we also attempted to identify final reports of phase II studies published as abstracts that were not yet reported in full article form. Abstracts were identified using an electronic search of the proceedings of the American Society of Clinical Oncology, the American Association for Cancer Research, and the San Antonio Breast Cancer Symposia meetings until the end of 2005. Identified abstracts were excluded if they indicated continuing patient accrual or incomplete efficacy results. Abstract data were verified against data presented at the actual scientific meeting, when possible.

Definition of Trial Versus Report

Each publication, whether abstract or full article, was considered a report. A small number of reports included more than one trial. Examples were reports in which multiple tumor types were included (each analyzed separately), multiple targeted agents were evaluated (in a noncomparative randomized design), or multiple doses of the same targeted agent were tested. To describe the outcomes of each active arm of targeted therapy appropriately, we considered a study with n prospective active treatment groups in which there was no intent to conduct a formal statistical comparison to be n trials within a single report. For those few reports where random assignment was to standard therapy or placebo, the study arms with no active targeted therapy or placebo were not considered as separate trials, and the reported outcome was that which was planned for the overall study.

Data Abstraction

All data were abstracted independently for each trial by both authors, with any discrepancies resolved by consensus. A summary of abstract data is found in Appendix Table A1 (online only). All planned primary and secondary

Table 1. Targeted Agents From 2004 Phase I Review Subjected to Search for Phase II Results

	Phase I Review Agents			This Phase II Review		
Target Category	Target	Agent	Drug Type	No. of Single-Agent Phase II Reports Identified	No. of Single-Agent Phase II Trials in Selected Tumors* Reviewed	Reference No.
EGFR/HER-2 signaling pathways	Farnesyltransferase	BMS-214662	SM			
		R115777	SM	5	6	19-23
		L778, 123	SM			
		SCH 66336	SM	1	1	24
	MEK	CI-1040	SM	1	3	25
	mTOR	CCI-779 (temsirolimus)	SM	3	7	26-28
	Raf kinase	ISIS 5132	AS	4	5	29-32
		BAY 43-9006 (sorafenib)	SM	1	2	33
Cell surface receptors	EGFR	ZD1839 (gefitinib)	SM	14	17	34-47
		OSI-774 (erlotinib)	SM	5	5	48-52
		C225 (cetuximab)	AB	3	3	53-55
		MAb225	AB			
		EMD 72000	AB			
		EKB 569	SM			
		RG83852	AB			
	HER-2	Trastuzumab	AB	8	9	56-63
	c-kit	STI-571 (imatinib)	SM	7	9	64-70
Angiogenesis	VEGF	Bevacizumab	AB	2	5	71,72
	VEGFR (plus other targets)	ZD6474	SM	1	2	73
		PTK787	SM			
		SU6668	SM			
		SU5416 (semaxanib)	SM	2	2	74,75
		SU11248 (sunitinib)	SM	2	2	76,77
	Other	Endostatin	Other			
Extracellular matrix	Matrix metalloproteinase	BB-2516 (marimastat)	SM	1	3	78
		BAY 12-9566 (tanomastat)	SM			
		COL-3	SM			
		BMS-275291	SM	1	2	79
Other	BCL-2	G3139	AS			
	PKCα	ISIS 3521 (aprinocarsen)	AS	5	5	29,32,80-82
	DNA methyltransferase	MG98	AS	1	1	83
Total No.		31†		65‡	89	

Abbreviations: EGFR, epidermal growth factor receptor; HER-2, human epidermal growth factor receptor 2; SM, small molecule; MEK, mitogen-activated protein-Erk kinase; mTOR, mammalian target of rapamycin; AS, antisense oligonucleotide; AB, antibody; VEGF, vascular endothelial growth factor; VEGFR, vascular endothelial growth factor receptor; PKCα, protein kinase C alpha.

*Selected tumors were breast, lung (small-cell and non–small-cell), prostate, colorectal, renal, and ovary.

†Nineteen agents identified for this review.

‡Actual total is 67, but two reports have been counted twice because they each included two agents and thus are found twice in this table (ISIS 5312 and ISIS 3521).

end points (eg, response, change in target expression, progression-free survival, either individually or in a multinomial combination) were abstracted, as were the actual outcome measures (eg, number of patients enrolled/eligible, number of patients with complete response, partial response, stable disease, progressive disease [PD], median time to progression, and overall survival). Data were not collected with respect to sex, age, or nature of the prior systemic treatment of patients.

Calculations

Once data abstraction was completed, several computations were undertaken. To determine the total response rate for a given trial, the total number of patients achieving either complete or partial response was divided by all eligible patients. Trials without response reported or collected were classified as not reported for this outcome. Although we initially planned to report stable disease rate, because substantially differing duration requirements were used to define stable disease across the trials reviewed, we elected to calculate the nonprogression rate as a means of better standardizing the output of our review. To determine the percentage of patients with nonprogression, patients with PD as best response were subtracted from the number of eligible patients entered to give the total number of nonprogressors (non-PD). This figure was then divided by the total number of eligible patients. Although we recognized that assignment of PD as a best response was somewhat dependent on the timing of follow-up (usually between 6 and 12 weeks), this was less variable than the duration used to define stable disease (which ranged from a few weeks to $>$ 6 months).

Once the response rate and PD rate for each trial was calculated, additional summary information was generated for presentation in tabular form. This included the calculation of overall rates of response (and nonprogression), which were determined for each drug by grouping all patients in all trials of a given agent with either complete or partial response (or nonprogression) and then dividing by the total number of eligible patients.

Response rates for trials and for drugs were categorized in the following range groupings: 0%, more than 0% to $\leq$ 10%, more than 10% to $\leq$ 20%, and more than 20%. Nonprogression rates were categorized as follows: $\leq$ 30%, more than 30% to $\leq$ 50%, more than 50% to $\leq$ 65%, and more than 65%. The numbers of trials with response rates or non-PD rates in the ranges noted were displayed in table format according to a variety of groupings of the trial data.

This included tables representing the number of trials in the various response categories by disease type, by individual drugs, by drugs grouped by target, and by population enrichment.

Drug Approval

Information on whether the agents under evaluation had received accelerated or full US Food and Drug Administration (FDA) approval for use in any of the six tumor types as of June 2007 was also identified. A table was then created that listed the number of drugs with overall response rates in the ranges described earlier versus the number of drugs in each grouping that had achieved regulatory approval. The relationship between the four response categories and the probability of FDA approval was assessed by the exact Cochran-Armitage linear trend test.

Agents With No Phase II Trials

There were 12 agents from the original list of 31 for which no phase II trials in any of the six tumor types were found. Every reasonable effort was made to determine the reasons for this (eg, drug stopped development, skipped phase II altogether, was evaluated in other diseases than those we focused on, or went into combination phase II immediately after phase I). However, not all agents could be traced.

RESULTS

Agents and Trials

Of the 31 agents surveyed, reports on single-agent phase II evaluation were retrieved for 19 in at least one of the prespecified tumor types (Table 1).[19-83] Altogether, 65 reports were identified (53 articles and 12 abstracts). Several reports contained results of evaluation of more than one tumor type or involved several different agents or dose levels; thus, the final tally of trials was 89. These were spread across all six tumor types, with the largest numbers in breast (21 trials) and lung cancers (13 trials in non–small-cell lung cancer and nine trials in small-cell lung cancer), followed by renal (15 trials), prostate (14 trials), and ovarian cancers (six trials; Appendix Table A2, online only).

Trial Design

Randomized versus nonrandomized designs. In the majority of reports (51 of 65 reports; 78%), the investigational agent was evaluated using a nonrandomized single-arm design (Table 2). Randomization was used in 14 (22%) of 65 reports. However, in only two reports was there random assignment to a placebo (this included the randomized discontinuation phase of the BAY 43-9006 renal cell carcinoma trial[33] and a randomized phase II study of gefitinib *v* placebo in prostate cancer[45]). One study in prostate cancer randomly assigned patients between an experimental agent and an active corticosteroid control arm.[74] In two reports,[29,32] patients were randomly assigned between two different investigational drugs; in both, the random assignment was between ISIS 3521 (aprinocarsen) and ISIS 5132 in a noncomparative phase II design. Finally, in nine reports, the random assignment was between various dose levels of the same investigational drug.

Table 2. Trial Design and End Points

Design and End Point	No. of Reports (N = 65)	Trials (N = 89) No.	Trials (N = 89) %
Nonrandomized	51	62	70
Randomized	14	27	30
Comparator arms:			
Placebo/standard		3	
Other investigational drug		4	
Other dose of same agent		20	
Primary end point			
Objective response		51	57
Multinomial (response and progressive disease)		10	11
Proportion progression free		8	9
Progression-free survival		8	9
Other		12	13

End points. The primary end point on which the trial design was based was most commonly objective response (51 trials; 57%; Table 2). In addition, 10 studies used a multinomial end point incorporating both response and nonprogression. Thus, objective response was the primary or coprimary end point in 61 trials. Only 16 trials were designed with end points of progression-free survival or the proportion of patients progression free at a prespecified time point. Some of the remaining studies were in prostate cancer and used measures of prostate-specific antigen (PSA) change (eg, PSA response, change in slope of PSA increase) as the primary end point. In addition, in two reports, toxicity was the primary end point.

Population enrichment. Efforts to enrich the population under evaluation by restricting entry to patients of a particular molecular or histologic tumor subtype were undertaken in 18 (20%) of 89 trials. As can be seen in Table 3, in 14 studies, enrichment was on the basis of a molecular marker assessed in tumor.

Hypotheses used in design. For those trials in which objective response was the primary end point, we attempted to identify the hypotheses used to derive the sample size from the methods sections of articles because we postulated that response rates of interest might be lower for targeted agents than those that have been traditionally used for phase II trials of cytotoxic drugs. Unfortunately, of the 51 trials in which response was indicated to be the primary end point, only 27 described the hypotheses that had led to design and the planned sample size. Of these, 20 based the sample size and, thus, stopping rules on response rates of interest of 20% or higher. In seven trials, response rates in the 10% to 15% range were targeted.

Sample size. Not surprisingly, the mean planned sample size for the trials reviewed depended on the design. For trials in which objective response was the end point, the mean maximum sample size planned was 56 patients (based on the data reported for 35 trials). When progression-free survival or non-PD was the end point used, the mean maximum sample size planned was 115, and for the multinomial design, the mean maximum sample size was 41.

Trial Results: Response and Nonprogression

Response rates. Seventy-six of 89 trials reported objective response outcomes. In total, 38 trials had overall response rates of 0%. In the other 38 trials, objective responses were seen; in 19 trials, response rates were more than 10%, and in eight trials, response rates were more than 20%.

Appendix Table A3 (online only) shows the reported response rates categorized in the ranges shown for all trials sorted by agent. Appendix Table A4 (online only) displays the same data but sorted by tumor type. Trial response results for all agents affecting the same target are shown in Appendix Table A5 (for example, all epidermal growth factor receptor inhibitors trials are displayed in one line in the

Table 3. Population Enrichment

Basis for Enrichment	No. of Trials (n = 18)	Agent*	Tumor Type*	Reference No.
EGFR expression	3	Erlotinib (n = 2); cetuximab (n = 1)	Lung (n = 1); ovary (n = 1); CRC (n = 1)	50,51,53
HER-2 expression	7	Trastuzumab	Breast (n = 5); lung (n = 1); ovary (n = 1)	56-59,61,62
Histology				
BAC	1	Erlotinib	Lung	49
Clear cell	2	Bevacizumab	Renal cell carcinoma	72
c-kit	4	Imatinib	Lung (n = 3); ovary (n = 1)	65,67,68
Other (SD at 12 weeks)	1	Sorafenib	Renal cell carcinoma	33
Total	18	Trastuzumab (n = 7); imatinib (n = 4); erlotinib (n = 3); bevacizumab (n = 2); sorafenib (n = 1); cetuximab		

Abbreviations: EGFR, epidermal growth factor receptor; CRC, colorectal cancer; HER-2, human epidermal growth factor receptor 2; BAC, bronchioloalveolar carcinoma; SD, stable disease.
*Numbers in parentheses represent No. of trials.

table; table online only). Finally, Table 4 provides the overall response rate by agent, pooling results for all trials (across all tumor types).

Nonprogression rates. Results for nonprogression rates were also tabulated, but not all are shown. Table 4 lists the non-PD rates by agent (pooling across all trials for each particular agent). As can be seen, non-PD rates were variable, but several agents (sorafenib, cetuximab, temsirolimus, trastuzumab, and gefitinib) had non-PD rates of 50% or more overall.

Although in most tumor types the ranking of agents by response rates or non-PD rates was similar (data not shown), renal cell carcinoma trials seemed to display a different pattern. Appendix Table A6 (online only) shows overall response rate and non-PD rates in renal cell carcinoma studies by agent. High non-PD rates were seen with four agents (sunitinib, sorafenib, temsirolimus, and imatinib), but only one of these, sunitinib, had a response rate that was more than 20%; the remainder had observed response rates less than 10%.

Response and Non-PD Rates by Disease Type and Regulatory Approval

To identify whether there were tumor-related patterns in the response or non-PD results and regulatory approval (as of June 2007), we examined overall rates of non-PD and response by agent in each tumor type in the trials reviewed, as shown in Appendix Tables A7 and A8 (online only). Numbers within each tumor type are too small to apply statistics, but it was observed that no agent with a 0% response rate in a given tumor type received approval in that tumor type. Similarly, no agent with non-PD rates less than 30% in a given tumor type received approval in that tumor type.

Table 4. Overall Response and Non-PD Rates by Agent

Agent	No. of Trials	Overall Response Rate (%)	Total No. of Patients in Response Rate Denominator*	Overall Non-PD Rate (%)	Total No. of Patients in Non-PD Rate Denominator*
Sorafenib	2	4	202	75	202
Marimastat	3	NR	NA	NR	NA
Bevacizumab	5	5.3	97	NR	NA
BMS-275291	2	0	80	36	80
Cetuximab	3	3	169	60	58
Temsirolimus	7	8	190	66	190
CI-1040	3	0	52	12	52
Aprinocarsen	5	0	87	13	60
ISIS 5132	5	0	71	27	71
MG98	1	0	15	40	15
Erlotinib	5	12	200	39	150
R1155777	6	5	208	23	208
SCH 66336	1	0	21	14	21
Imatinib	9	0	112	21	112
Sunitinib	2	28	127	0	63
Semaxanib	2	5	45	0	29
Trastuzumab	9	18	562	53	420
Gefitinib	17	10	698	50	627
ZD6474	2	0	46	NR	NA
Total	89				

Abbreviations: PD, progressive disease; NR, not reported in any trial; NA, not applicable.
*Response and non-PD rate denominators may not match if some trials had one or the other not reported. Only trials with data reported were used to calculate rates.

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.