Phase II Oncology Trials: Let's Be Positive

Mark J. Ratain

The critical decision in drug development is often at the end of phase II, as phase III trials are an expensive undertaking, with the potential for significant corporate and public consequences in the event of a negative phase III trial. Thus, retrospective analyses of the usefulness of phase II trials are welcomed and valued, as illustrated by the study of Goffin et al. (1). In this analysis of 58 cytotoxic drugs, 46 studied in phase II, the authors show a statistical relationship between objective response rate and probability of approval of the drug, with the important exception of responses in metastatic melanoma and renal cell carcinoma.

Most readers will not be surprised by the authors' findings of a relationship between phase II response rate and subsequent marketing of the drug. A drug may show activity in a specific disease in phase I or II, however, but may not be developed further in that disease because of business concerns. Thus, one would expect a bias towards phase III trials in diseases with larger markets, generally considered to be breast, colorectal, lung, and prostate cancers. Therefore, a lower approval rate in melanoma and renal cell cancer could potentially be affected by a lower number of phase III trials in these indications with smaller markets. In fact, Goffin et al. (1) did not address whether or not phase III trials were conducted with those agents that had phase II response rates of 10% or more. One cannot firmly conclude, therefore, that the phase II results were not predictive for phase III results in melanoma and renal cell cancer, as opposed to an alternative explanation that there was relatively little interest in seeking approval for these indications. What have we learned from this analysis? It is important to emphasize that a correlation, albeit significant, does not necessarily result in usefulness. A more important issue is the usefulness of phase II oncology trials, particularly in the context of predicting marketing approval for the indication studied in the specific phase II trial. To address this question, the data of Goffin et al. (1) have been reanalyzed using standard approaches to studying diagnostic tests. In this context, one normally uses metrics such as sensitivity, specificity, and predictive value (2). To assess this, one can consider various objective response rate cutoffs, as illustrated in Table 1, such as cutoffs of 10% and 20%. Sensitivity, or true positive rate, is defined as the percentage of drugs approved for an indication for which the phase II trial exceeded the threshold response rate. (It should be noted that if no drugs were approved for an indication, it is not possible to calculate the sensitivity). Specificity, or true negative rate, is defined as the percentage of drugs not approved for an indication for which the phase II trial did not exceed the threshold response rate. Positive predictive value is the likelihood of approval (for a disease) given a positive phase II trial (i.e., exceeding threshold response rate). Negative predictive value is the likelihood of not being approved for a disease (regardless of whether a phase III trial was conducted) given a negative phase II trial (i.e., not exceeding threshold response rate).

When the data of Goffin et al. (1) were reanalyzed in this way (Table 1), it is clear that phase II oncology trials have a high negative predictive value but a low positive predictive value. Sensitivity was high at a threshold response rate of 10%, as this would be expected to exclude few active agents. At a higher threshold response rate of 20%, there was greater specificity but with some tradeoff on sensitivity.

Is response rate the right end point for phase II oncology trials? This paradigm has recently been questioned in the context of the development of noncytotoxic agents (3). The data of Goffin et al. suggest that we should consider rejecting our current paradigm for cytotoxic agents as well, particularly if the goal of phase II studies is to predict for phase III success. Oncology has been recently singled out as a therapeutic area for which positive phase II trials have not been predictive of phase III success (4). As noted in Table 1, the best positive predictive value was for non–small-cell lung cancer, in which 75% of agents whose phase II response rate exceeded 20% were subsequently approved. Unfortunately, this threshold would have excluded two of our currently approved cytotoxic agents for non–small-cell lung cancer as only three of the five approved agents in this study met this criterion.

Moving forward, it is critical to consider the purpose of phase II trials. In essentially all other therapeutic areas, such studies are usually randomized, dose-ranging controlled trials, often including a placebo group, but enable determination of the relationship between dose and an end point of clinical interest, and, by extension, provide evidence of activity (5). Most importantly, phase III trials are expected to confirm findings of phase II studies, and, outside of oncology, are generally positive; consequently, phase II trials in other therapeutic areas have a high positive predictive value (4, 6). In contrast, oncology trials have a high negative predictive value but a relatively low positive predictive value (Table 1). These features may be due to the nature of cancer and anticancer agents. On the other hand, these may be due to the differences in the designs used for phase II trials in oncology, in which the classic designs are formally aimed at proving that the drug does not have activity (7). In fact, a trial that is not negative is not necessarily "positive," as borne out by the low positive predictive value for marketing approval, presumably due in a large part to the low phase III success rate.

Author's Affiliation: Section of Hematology/Oncology, Department of Medicine, Committees on Clinical Pharmacology and Pharmacogenomics and Molecular Medicine, and Cancer Research Center, University of Chicago, Chicago, Illinois
Received 5/11/05; accepted 5/25/05.
Requests for reprints: Mark J. Ratain, Department of Medicine, University of Chicago, MC2115, 5841 South Maryland Street, Chicago, IL 60637. Phone: 773-702-4400; Fax: 773-702-3969; E-mail: mratain@medicine.bsd.uchicago.edu.

doi:10.1158/1078-0432.CCR-05-1046

Table 1. Analysis of usefulness of phase II results to predict drug approval within disease category by threshold response rate (using data from Goffin et al.)

	Melanoma (n = 29)	Renal (n = 15)	Breast (n = 26)	NSCLC (n = 25)	Ovarian (n = 22)	Colorectal (n = 28)
ORR > 10%						
SE (%)	?	?	100	80	100	100
SP (%)	84	93	50	70	85	88
PPV (%)	0	0	25	40	33	57
NPV (%)	100	100	100	93	100	100
ORR > 20%						
SE (%)	?	?	100	60	67	50
SP (%)	95	93	77	95	80	96
PPV (%)	0	0	44	75	33	67
NPV (%)	100	100	100	90	80	92

Abbreviations: NSCLC, non – small-cell lung cancer; ORR, objective response rate; SE, sensitivity; SP, specificity; PPV, positive predictive value; NPV, negative predictive value.

In addition, classic phase II designs may not identify a highly active agent, as exemplified by the recent studies of sorafenib. A randomized discontinuation trial of this kinase inhibitor was conducted in 202 patients with metastatic renal cell cancer (8). Although the objective response rate (by independent review) was only 4%, there was an obvious and highly significant difference in failure-free survival after randomization, which was confirmed in a phase III trial of 800 patients (9). If this drug had been developed using classic criteria, it may never have entered phase III for this indication, given its low response rate.

The solution for moving forward is to design trials for success, not for failure. This requires abandonment of current oncology paradigms for early clinical trials and adoption of generally accepted principles of drug development. These principles have been promulgated by the Food and Drug Administration (10) and the U.S. Code of Federal Regulations Title 21 Section 312.21 (11), which states, "Phase 2 includes the controlled clinical studies conducted to evaluate the effectiveness of the drug for a particular indication or indications in patients with the disease or condition under study and to determine the common short-term side effects and risks associated with the drug. Phase 2 studies are typically well controlled, closely monitored, and conducted in a relatively small number of patients, usually involving no more than several hundred subjects." Implementation of such new paradigms will require greater attention to trial design issues in both phase I and II, but with the overriding principle that will require larger trials to prove activity. Such change will hopefully lead to a higher success rate in phase III, thereby allowing a more rapid translation of scientific advances into cost-effective therapy for cancer patients.

References

1. Goffin J, Baral S, Dongsheng T, Nomikos D, Seymour L. Objective responses in patients with malignant melanoma or renal cell cancer in early clinical studies do not predict for regulatory approval. Clin Cancer Res 2005;11:5928–34.
2. Chu K. An introduction to sensitivity, specificity, predictive values and likelihood ratios [review]. Emerg Med 1999;11:175–81.
3. Ratain MJ, Eckhardt SG. Phase II studies of modern drugs directed against new targets: if you are fazed, too, then resist RECIST. J Clin Oncol 2004;22:4442–5.
4. Booth B, Glassman R, Ma P. Oncology's trials. Nat Rev Drug Discov 2003;2:609–10.
5. Shen L. An improved method of evaluating drug effect in a multiple dose clinical trial. Stat Med 2001;20:1913–29.
6. Sheiner LB. Learning versus confirming in clinical drug development. Clin Pharmacol Ther 1997;61:275–91.
7. Ratain MJ, Mick R, Schilsky RL, Siegler M. Statistical and ethical issues in the design and conduct of phase I and II clinical trials of new anticancer agents. J Natl Cancer Inst 1993;85:1637–43.
8. Ratain M, Eisen T, Stadler W, et al. Final findings from a phase II, placebo-controlled, randomized discontinuation trial (RDT) of sorafenib (BAY 43–9006) in patients with advanced renal cell carcinoma (RCC). Proc Am Soc Clin Oncol 2005;23:388s.
9. Onyx Pharmaceuticals [homepage on the Internet]. BAY 43–9006 shown to delay disease progression in phase III study in advanced kidney cancer patients. [cited 2005 Mar 21]. Available from: http://www.onyx-pharm.com/wt/page/pr.1111376865.
10. Food and Drug Administration [homepage on the Internet]. General considerations for the clinical evaluation of drugs. [issued 1978]. Available from: www.fda.gov/cder/guidance/old034fn.pdf.
11. US Government Printing Office-NARA [homepage on the Internet]. Food and Drug Administration, Dept. of Healthy and Human Services, Code of Federal Regulations. Investigational new drug application. Vol. 21, pp. 54, 2003. Available from: http://www.access.gpo.gov/cgi-bin/cfrassemble.cgi?title=200021.