

Is Response Rate Relevant to the Phase II Trial Design of Targeted Agents?

Afshin Dowlati, *Department of Medicine, Case Western Reserve and University Hospitals Ireland Cancer Center and the Developmental Therapeutics Program, Case Comprehensive Cancer Center, Cleveland, OH*

Pingfu Fu, *Department of Epidemiology & Biostatistics, Case Western Reserve University and University Hospitals Ireland Cancer Center and the Developmental Therapeutics Program, Case Comprehensive Cancer Center, Cleveland, OH*

Phase II trials serve an important role in oncologic drug development by screening single agents or combinations of agents in a single disease entity for signs of efficacy that would indicate additional development. The most common design is a single-arm study using a short-term end point such as response rate. Statistically, a test of the null hypothesis versus the alternative hypothesis is formulated in which the null hypothesis $H_0: p = p_0$, where p is the true (but unknown) response rate and p_0 is set at a historical response rate. The alternative hypothesis states that $H_1: p = p_1$, where p_1 is a response rate that would justify additional investigation of the agent. A modification of this approach is the Simon optimal two-stage design, which includes an interim analysis to decide whether or not to proceed with full accrual, thus minimizing the number of patients placed on ineffective therapies.¹ Given the failure of many new agents in phase III clinical trials, phase II trials have come under heavy pressure to better predict phase III outcomes. For this purpose, alternative designs such as randomized phase II studies, randomized discontinuation designs, and alternative short-term end points such as time to progression (TTP) and even longer term end points such as survival² have been put forward. In addition, response rate, the classic short-term end point, has come under criticism and so-called waterfall graphs have been used to show the proportion of patients with minor responses as an additional indication of drug efficacy.³ Another variation is the use of multinomial designs that evaluate efficacy on two short-term end points: tumor response and proportion of patients who do not experience disease progression.⁴

El-Maraghi and Eisenhauer,⁵ in this issue of the *Journal of Clinical Oncology*, confirm that our classic approach continues to be predictive of success. Despite the search for the novel valid surrogate makers of biologic and antitumor activity, tumor response and progression-free survival (PFS) are still the only widely accepted end points of measuring drug efficacy: 87% of the studies use those end points, as indicated by El-Maraghi et al. In their study, 89 phase II trials of 19 targeted agents in six disease types were analyzed. Most agents were small molecule inhibitors; others were antibodies or antisense treatments. They conclude that all new molecular targeted agents receiving regulatory approval showed some measure of response during their single-agent phase II evaluation.

Additional scrutiny of their data leads to other observations. For example, five of six agents that received regulatory approval all had

response rates more than 5% and three of six had response rates more than 10%. This would suggest that, for novel molecular targeted drugs evaluated as single agents, traditional alternate hypothesis response rates used for cytotoxic chemotherapy agents might not be valid. Indeed, when evaluating cytotoxic chemotherapy for non-small-cell lung cancer (NSCLC), a response rate of 20% or higher is considered exciting, whereas epidermal growth factor receptor inhibitors are associated with response rates of approximately 10%.⁶ This suggests that the target response rate for the alternate hypothesis that prompts phase III testing may need to be lowered. This strategy will unfortunately lead to an increase in sample sizes as well as a higher risk of false positivity with phase II trials. It is clear from El-Maraghi et al, however, that in the absence of any response with a single agent, the likelihood of regulatory approval is remote. This conclusion is in concordance with a similar retrospective analysis of chemotherapy studies.⁷ Sponsors that see no response whatsoever run a high risk of failure in the phase III setting. The decision between PFS or response is more complex, but El-Maraghi et al have shown that the median sample size for studies using PFS was 115, compared with 56 for response rate-directed trials. Thus, studies using response rate rather than PFS will likely be more efficient.

The current analysis, however, was limited to trials of single agents. It would be interesting to see if their conclusions remain valid for combinations of molecular targeted agents with cytotoxics. For example, based on a randomized phase III study,⁸ bevacizumab in combination with doublet chemotherapy was recently approved for advanced nonsquamous NSCLC. The rationale to proceed to a phase III study was based on a randomized three-arm phase II study of chemotherapy with two different doses of bevacizumab (7.5 or 15 mg/kg) compared with chemotherapy alone.⁹ There were two primary end points: response rate and TTP. In the high-dose bevacizumab plus chemotherapy arm, a response rate of 40% was observed, compared with a response rate of 31% for chemotherapy alone. This corresponds to the 30% relative increase in response rate as they hypothesized. However, they were unable to observe the anticipated 100% improvement in TTP. This example, in addition to the study combining trastuzumab with chemotherapy,¹⁰ suggests that an increase in response rate is likely to be seen in these combination studies as well. Clearly, the additional complexities raised with combination chemotherapy makes it more difficult in terms of study design and end

point. For example, cytostatic-cytotoxic synergy and synergistic toxicity must be taken into account.

El-Maraghi et al also showed that enrichment of the target patient population was used in only 20% of the studies. This infrequent use of the enrichment process is a testimony to the difficulties in understanding predictive factors for response. For epidermal growth factor receptor inhibitors in NSCLC, only after completion of phase II studies was the importance of DNA sequencing established.¹¹ Enrichment of the enrolling patient population can occur by several methods. Enrolling patients with a particular molecular tumor subtype (population enrichment) is an ideal strategy for many new agents with defined targets; such a strategy, although not commonly used, can yield significant success if the target can be defined and measured, such as human epidermal growth factor receptor 2 expression and the development of lapatinib.¹² Enrichment can also occur by tumor growth characteristics such as the randomized discontinuation design.¹³ In such a design, all patients (heterogeneous population) are given the experimental agent initially. After some fixed period (run-in phase), patients are evaluated; patients who respond to the treatment continue taking the drug, those with disease progression are taken off study, and patients with stable disease are randomly assigned between continued administration of the drug or placebo for another fixed period of time. Although this strategy has had successful applications,¹⁴ its wide application is limited by many drawbacks, such as statistical power, selection bias, sample size, and generalizability of study results.^{15,16} Finally, another form of enrichment relates to simply tightening the inclusion criteria, such as with histologic subtype.

We also find it astonishing that, of the studies that used response as the primary end point, roughly only 50% provided statistical details about their sample size calculation. We previously reported a similar result in published trials of NSCLC.¹⁷ Another important finding is the relative lack of use of randomization during phase II drug development. Although 14 of 65 reports were randomized studies, the majority of randomizations used different doses of the same drug and did not contain a control arm. The argument for the use of a control arm in small randomized phase II trials (ie, to ensure accrual of the appropriate patient population) clearly has not been accepted by the majority of investigators.

In conclusion, El-Maraghi and Eisenhauer have contributed to the field of clinical trial design by showing us that clinical response remains an important end point. Additional research in this field is desperately needed if we are to predict more accurately the outcome for new agents in development.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships marked with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure

Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Employment or Leadership Position: None **Consultant or Advisory Role:** None **Stock Ownership:** None **Honoraria:** Afshin Dowlati, Genentech, Eli Lilly **Research Funding:** Afshin Dowlati, Celgene, GlaxoSmithKline, Genentech **Expert Testimony:** None **Other Remuneration:** None

AUTHOR CONTRIBUTIONS

Conception and design: Afshin Dowlati, Pingfu Fu

Financial support: Afshin Dowlati, Pingfu Fu

Administrative support: Afshin Dowlati, Pingfu Fu

Collection and assembly of data: Afshin Dowlati, Pingfu Fu

Data analysis and interpretation: Afshin Dowlati, Pingfu Fu

Manuscript writing: Afshin Dowlati, Pingfu Fu

Final approval of manuscript: Afshin Dowlati, Pingfu Fu

REFERENCES

- Simon R: Optimal two-stage designs for phase II clinical trials. *Control Clin Trials* 10:1-10, 1989
- Case DL, Morgan TM: Design of phase II cancer trials evaluating survival probabilities. *BMC Med Res Methodol* 3:6, 2003
- Karrison TG, Maitland ML, Stadler WM, et al: Design of phase II cancer trials using a continuous endpoint of change in tumor size: Application to a study of sorafenib and erlotinib in non-small-cell lung cancer. *J Natl Cancer Inst* 99:1455-1461, 2007
- Dent S, Zee B, Dancey J, et al: Application of a new multinomial phase II stopping rule using response and early progression. *J Clin Oncol* 19:785-791, 2001
- El-Maraghi RH, Eisenhauer EA: Review of phase II trial designs used in studies of molecular targeted agents: Outcomes and predictors of success in phase III. *J Clin Oncol* 26:1346-1354, 2008
- Kris MG, Natale RB, Herbst RS, et al: Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: A randomized trial. *JAMA* 290:2149-2158, 2003
- Goffin J, Baral S, Tu D, et al: Objective responses in patients with malignant melanoma or renal cell cancer in early clinical studies do not predict regulatory approval. *Clin Cancer Res* 11:5928-5934, 2005
- Sandler AB, Gray R, Perry MC, et al: Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. *N Engl J Med* 355:2542-2550, 2006
- Johnson DH, Fehrenbacher L, Novotny WF, et al: Randomized phase II trial comparing bevacizumab plus carboplatin and paclitaxel with carboplatin and paclitaxel alone in previously untreated locally advanced or metastatic non-small-cell lung cancer. *J Clin Oncol* 22:2184-2191, 2004
- Slamon DJ, Leyland-Jones B, Shak S, et al: Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* 344:783-792, 2001
- Sequist LV, Bell DW, Lynch TJ, Haber DA: Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. *J Clin Oncol* 25:587-595, 2007
- Geyer CE, Forster J, Lindquist D, et al: Lapatinib plus capecitabine for HER2-positive advanced breast cancer. *N Engl J Med* 355:2733-2743, 2006
- Rosner GL, Stadler W, Ratain M: Randomized discontinuation design: Application to cytostatic antineoplastic agents. *J Clin Oncol* 20:4478-4484, 2002
- Ratain MJ, Eisen T, Stadler WM, et al: Phase II placebo-controlled randomized discontinuation trial of sorafenib in patients with metastatic renal cell carcinoma. *J Clin Oncol* 24:2505-2512, 2006
- Capra WB: Comparing the power of the discontinuation design to that of the classic randomized design on time-to-event endpoints. *Control Clin Trials* 25:168-177, 2004
- Leber PD, Davis CS: Threats to the validity of clinical trials employing enrichment strategies for sample selection. *Control Clin Trials* 19:178-187, 1998
- Bagai RK, Dowlati A: Evolution of statistical methodology and design of phase II/III clinical trials in non-small cell lung cancer. *J Clin Oncol* 24:391s, 2006 (suppl; abstr 7108)