DOI: 10.1093/jnci/djr238
Advance Access publication on June 27, 2011.

More Randomization in Phase II Trials: Necessary but not Sufficient

Lawrence Rubinstein, Michael LeBlanc, Malcolm A. Smith

Correspondence to: Lawrence Rubinstein, PhD, National Cancer Institute, National Institutes of Health, Biometric Research Branch, Division of Cancer Treatment and Diagnosis, Executive Plaza North, Rm 8130, MSC-7434, Bethesda, MD 20892-7434 (e-mail: rubinsteinl@ctep.nci.nih.gov).

Historically, phase II trials in oncology were generally single armed, constructed to distinguish between a tumor response rate felt to indicate a lack of promise (often 5%) and a rate that would indicate potential benefit (often 20%), with a one-sided type I error rate of 5%–10% and a type II error rate of 10%–20% (1). The dominant use of this design was based on the premise that an agent that could not produce a tumor response rate of 20% was not likely to produce a clinically meaningful overall survival (OS) or progression-free survival (PFS) benefit in subsequent phase III testing. Recent trends in oncology drug development have challenged this paradigm. Many phase II trials are now designed to assess the promise of a molecularly targeted agent, given either alone or in combination with another regimen. In many cases, these agents are not anticipated to produce or improve tumor response rates; rather, the desired outcome from their use is improved PFS or OS through means other than direct cell killing as evidenced by tumor shrinkage (2). In general, PFS is the preferred endpoint for such phase II trials. PFS is statistically more efficient than OS because the time to achieve the endpoint of PFS is substantially shorter, and the treatment effect is not diluted by salvage treatment. However, in a situation with no effective salvage therapy and/or a disease with concerns regarding the timing of progression assessment, OS could be chosen as the endpoint. Such trials can be single-arm studies, compared with historical controls, or can be randomized.

The review by Sharma et al (3) in this issue of the Journal is a welcome addition to the growing chorus in favor of increased randomization in phase II trials for agents with little likelihood for single-agent tumor regression and for which endpoints such as PFS are used. This promotion of randomization is already having dramatic effect. Current records of the Cancer Therapy Evaluation Program of the National Cancer Institute (NCI) reveal that only 1.5% (68/4437) of the completed NCI-sponsored phase II studies were randomized. In contrast, 28% (69/243) of the currently active phase II studies are randomized, and of the trials activated after December 31, 2009, 37% are randomized. A primary reason for this increase is the appreciation, in the trial design and review process, that even a modest upward drift in the PFS of the study population compared with historical controls, which is independent of the effect of the new agent being tested, can inflate the type I error rate approximately threefold (4). For example, a drift from 50% to 55% in the control 4-month PFS rate, when not accounted for, will increase the type I error of a single-arm Simon optimal trial (5) targeting a 70% 4-month PFS from 0.10 to 0.26. Coupled with this is the realization that such an upward drift over time is relatively likely for PFS as the standard of care improves (6).

It is widely accepted that a substantial portion of phase II trials will still be appropriately single arm (1,6–8). This includes trials of agents for which tumor regression is anticipated based on mechanism of action, as well as early phase II monotherapy trials to establish a tumor response signal of biological efficacy. Additionally, monotherapy and combination trials with PFS endpoints in diseases with no effective standard therapy and established stable historical controls (eg, recurrent glioblastoma) can be justified. For OS, an historical database for melanoma has proven useful for designing single-arm studies (9). In some situations, adjustment for observed differences in the distribution of known prognostic factors between the historical database and the observed single arm study can

reduce potential bias and strengthen inferences. Examples of adjustment strategies include both regression models and probability reweighting (10).

Importantly, expanding the use of randomization to all phase II situations in which it is appropriate will not by itself maximize the positive predictive value of phase II trials (the probability of a positive phase II trial yielding an agent or combination that is effective in subsequent definitive phase III trials). This value is dependent not only on the type I error rate of the phase II trial but also on the balance between the true- and false-positive rates of the phase II trials for the population of interest, as well as the degree to which the phase II endpoints predict the ultimate phase III endpoints. For example, if the type I and type II error bounds are both .10, then the positive predictive value of a phase II trial will vary between 32% and 61% in the setting in which first, the collection of agents and combinations tested is effective, with probability varying between 5% and 15%, according to the phase II endpoint, and second, the phase II endpoint is a perfect surrogate for the phase III endpoint. Because the second stipulation is never the case, the positive predictive value may be substantially less.

There are four potential approaches to maximizing the effectiveness of phase II trials as predictors for phase III success:

1. The pool of agents and combinations going into phase II testing can be enriched for truly active agents. Enrichment may be possible through the increased use of pharmacodynamic assays in phase I and phase 0 testing (11), allowing for go/no-go decisions before phase II testing. Additional single-arm clinical data (potentially collected at phase I or phase II) may be helpful for screening for agents before undertaking randomization.
2. The subpopulations in which agents and combinations are potentially effective can be better identified so that phase II testing can be limited to such subpopulations. This may be done by increased development and use of pharmacodynamic assays to better characterize the agents (11) and increased development and use of biomarkers to better identify correspondingly sensitive subpopulations of patients (7,8).
3. Phase II endpoints that capture and predict a substantial percentage of the treatment effect reflected in the ultimate phase III endpoints can be identified, established, and used (7,12). Such endpoints, including new imaging endpoints, may vary by class of agent and by disease (13,14).
4. Even if the approaches listed above are only modestly successful in enriching the pool of phase II agents and combinations, so that they are effective, with probability varying between 20% and 40%, according to the phase II endpoint, the positive predictive value of phase II trials (to reflect true efficacy according to the phase II endpoints) could be increased to between 69% and 86%. How well these phase II trials would then predict phase III efficacy would depend upon the proportion of the phase III treatment effect captured by the phase II endpoint. However, in situations in which the above approaches are not so successful in enriching the pool of phase II agents and combinations, conducting phase II trials at the significance level of .05 (rather than the .10) should be considered. In this way, even if the agents and combinations are effective, with probability varying between 10% and 20%, according to the phase II endpoint, the positive predictive value of phase II trials to reflect true efficacy according to the phase II endpoints would vary between 67% and 82%.

A longer-term issue is whether the conduct of randomized phase II trials designed to evaluate PFS endpoints for agents that do not induce tumor regression as single agents represents a productive strategy. It is too early for conclusions, but there is emerging evidence to suggest that future oncologists will view this strategy as having identified only marginally to modestly effective agents. To date, many of the agents requiring phase II randomized trial designs because of their lack of tumor-regressing activity for the disease being studied are angiogenesis inhibitors (eg, agents targeting vascular endothelial growth factor [VEGF/VEGF receptor {R} 2] signaling). This class of agents has been remarkably effective for renal cell carcinoma, and multiple agents in the class have achieved regulatory approval for this indication (15–18). However, this is a tumor type for which many agents targeting VEGFR2 have single-agent tumor-regressing activity (16,19,20). Outside of cancers such as renal cell carcinoma and thyroid cancer, for which substantial single-agent response rates are observed (16,19–22), the track record for VEGF pathway inhibitors has primarily been one of the failed phase III trials or the phase III trials that have resulted in only modest prolongations in PFS and in some cases small increases in OS (23–31), although there are exceptions (32). This limited return on investment is in marked contrast to the improvements in outcome that have been observed for agents prioritized on the basis of their ability to induce substantial rates of regression in classic phase II studies in patients with recurrent/refractory disease. The strategy of prioritizing agents with robust single-agent activity has led to substantial improvements in survival, as demonstrated in phase III trials in a number of settings, such as arsenic trioxide for acute promyelocytic leukemia (33), tretinoin for acute promyelocytic leukemia (34,35), imatinib for chronic myeloid leukemia and Ph^+ acute lymphoblastic leukemia (36,37), rituximab for various types of B-cell non-Hodgkin lymphoma (38,39), trastuzumab for human epidermal growth factor receptor 2–positive breast cancer (40), and lenalidomide for multiple myeloma (41). A separate set of design and prioritization issues that are beyond the scope of this editorial apply to agents lacking single-agent tumor-regressing activity for which there is evidence of synergistic or synthetically lethal interactions when used in combination with other agents.

In summary, Sharma et al. (3) provide a rationale for selecting randomized phase II designs over single-arm designs when evaluating agents that lack single-agent tumor-regressing activity. An open question is whether such agents will be able to provide the level of improvement in survival that patients and their physicians seek.

References

1. Rubinstein L, Crowley J, Ivy P, LeBlanc M, Sargent D. Randomized phase II designs. *Clin Cancer Res.* 2009;15(6):1883–1890.
2. Dhani N, Tu D, Sargent DJ, Seymour L, Moore MJ. Alternate endpoints for screening phase II studies. *Clin Cancer Res.* 2009;15(6):1873–1882.
3. Sharma MR, Stadler WM, Ratain MJ. Randomized phase II trials: a long-term investment with promising returns. *J Natl Cancer Inst.* 2011;103(14):1093–1100.
4. Tang H, Foster NR, Grothey A, Ansell SM, Goldberg RM, Sargent DJ. Comparison of error rates in single-arm versus randomized phase II cancer clinical trials. *J Clin Oncol.* 2010;28(11):1936–1941.
5. Simon R. Optimal two-stage designs for phase II clinical trials. *Control Clin Trials.* 1989;10(1):1–10.

6. Gan HK, Grothey A, Pond GR, Moore MJ, Siu LL, Sargent D. Randomized phase II trials: inevitable or inadvisable? *J Clin Oncol.* 2010; 28(15):2641–2647.
7. Stewart DJ. Randomized phase II trials: misleading and unreliable. *J Clin Oncol.* 2010;28(31):e649–e650.
8. Seymour L, Ivy SP, Sargent D, et al. The design of phase II clinical trials testing cancer therapeutics: consensus recommendations from the clinical trial design task force of the national cancer institute investigational drug steering committee. *Clin Cancer Res.* 2010;16(6):1764–1769.
9. Korn EL, Liu PY, Lee SJ, et al. Meta-analysis of phase II cooperative group trials in metastatic stage IV melanoma to determine progression-free and overall survival benchmarks for future phase II trials. *J Clin Oncol.* 2008;26(4):527–534.
10. Cole SR, Hernan MA. Adjusted survival curves with inverse probability weights. *Comput Methods Programs Biomed.* 2004;75(1):45–49.
11. Kummar S, Kinders R, Rubinstein L, et al. Compressing drug development timelines in oncology using phase '0' trials. *Nat Rev Cancer.* 2007; 7(2):131–139.
12. Sargent DJ, Rubinstein L, Schwartz L, et al. Validation of novel imaging methodologies for use as cancer clinical trial end-points. *Eur J Cancer.* 2009;45(2):290–299.
13. Burzykowski T, Buyse M, Yothers G, Sakamoto J, Sargent D. Exploring and validating surrogate endpoints in colorectal cancer. *Lifetime Data Anal.* 2008;14(1):54–64.
14. Burzykowski T, Buyse M, Piccart-Gebhart MJ, et al. Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. *J Clin Oncol.* 2008;26(12):1987–1992.
15. Summers J, Cohen MH, Keegan P, Pazdur R. FDA drug approval summary: bevacizumab plus interferon for advanced renal cell carcinoma. *Oncologist.* 2010;15(1):104–111.
16. Rock EP, Goodman V, Jiang JX, et al. Food and Drug Administration drug approval summary: sunitinib malate for the treatment of gastrointestinal stromal tumor and advanced renal cell carcinoma. *Oncologist.* 2007;12(1): 107–113.
17. Kane RC, Farrell AT, Saber H, et al. Sorafenib for the treatment of advanced renal cell carcinoma. *Clin Cancer Res.* 2006;12(24):7271–7278.
18. Griffiths C, Hay N, Sutcliffe F, Stevens A. NICE guidance on pazopanib for first-line treatment of advanced renal-cell carcinoma. *Lancet Oncol.* 2011;12(3):221–222.
19. Rixe O, Bukowski RM, Michaelson MD, et al. Axitinib treatment in patients with cytokine-refractory metastatic renal-cell cancer: a phase II study. *Lancet Oncol.* 2007;8(11):975–984.
20. Sternberg CN, Davis ID, Mardiak J, et al. Pazopanib in locally advanced or metastatic renal cell carcinoma: results of a randomized phase III trial. *J Clin Oncol.* 2010;28(6):1061–1068.
21. Ye L, Santarpia L, Gagel RF. The evolving field of tyrosine kinase inhibitors in the treatment of endocrine tumors. *Endocr Rev.* 2010;31(4): 578–599.
22. Bible KC, Suman VJ, Molina JR, et al. Efficacy of pazopanib in progressive, radioiodine-refractory, metastatic differentiated thyroid cancers: results of a phase 2 consortium study. *Lancet Oncol.* 2010;11(10): 962–972.
23. Miles DW, Chan A, Dirix LY, et al. Phase III study of bevacizumab plus docetaxel compared with placebo plus docetaxel for the first-line treatment of human epidermal growth factor receptor 2-negative metastatic breast cancer. *J Clin Oncol.* 2010;28(20):3239–3247.
24. Robert NJ, Dieras V, Glaspy J, et al. RIBBON-1: randomized, double-blind, placebo-controlled, phase III trial of chemotherapy with or without bevacizumab for first-line treatment of human epidermal growth factor receptor 2-negative, locally recurrent or metastatic breast cancer. *J Clin Oncol.* 2011;29(10):1252–1260.
25. Reck M, von PJ, Zatloukal P, et al. Overall survival with cisplatin-gemcitabine and bevacizumab or placebo as first-line therapy for nonsquamous non-small-cell lung cancer: results from a randomised phase III trial (AVAiL). *Ann Oncol.* 2010;21(9):1804–1809.
26. Kindler HL, Niedzwiecki D, Hollis D, et al. Gemcitabine plus bevacizumab compared with gemcitabine plus placebo in patients with advanced pancreatic cancer: phase III trial of the Cancer and Leukemia Group B (CALGB 80303). *J Clin Oncol.* 2010;28(22):3617–3622.
27. Van CE, Vervenne WL, Bennouna J, et al. Phase III trial of bevacizumab in combination with gemcitabine and erlotinib in patients with metastatic pancreatic cancer. *J Clin Oncol.* 2009;27(13):2231–2237.
28. Batchelor T, Mulholland P, Neyns B, et al. The efficacy of cediranib as monotherapy and in combination with lomustine compared to lomustine alone in patients with recurrent glioblastoma—A phase III randomized study. *Neuro-Oncology.* 2010;12(Suppl 4): Abstract 0T-25.
29. Hoff PM, Hochhaus A, Pestalozzi BC, et al. Cediranib + FOLFOX/XELOX versus placebo + FOLFOX/XELOX in patients (pts) with previously untreated metastatic colorectal cancer (MCRC): a randomized, double-blind, phase III study (Horizon II). *Ann Oncol.* 2010;21(Suppl 8):viii9.
30. Hecht JR, Trarbach T, Hainsworth JD, et al. Randomized, placebo-controlled, phase III study of first-line oxaliplatin-based chemotherapy plus PTK787/ZK 222584, an oral vascular endothelial growth factor receptor inhibitor, in patients with metastatic colorectal adenocarcinoma. *J Clin Oncol.* 2011;29(15):1997–2003.
31. Allegra CJ, Yothers G, O'Connell MJ, et al. Phase III trial assessing bevacizumab in stages II and III carcinoma of the colon: results of NSABP protocol C-08. *J Clin Oncol.* 2011;29(1):11–16.
32. Hurwitz H, Fehrenbacher L, Novotny W, et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N Engl J Med.* 2004;350(23):2335–2342.
33. Powell BL, Moser B, Stock W, et al. Arsenic trioxide improves event-free and overall survival for adults with acute promyelocytic leukemia: North American Leukemia Intergroup Study C9710. *Blood.* 2010;116(19): 3751–3757.
34. Tallman MS, Andersen JW, Schiffer CA, et al. All-trans-retinoic acid in acute promyelocytic leukemia. *N Engl J Med.* 1997;337(15):1021–1028.
35. Fenaux P, Chastang C, Chevret S, et al. A randomized comparison of all transretinoic acid (ATRA) followed by chemotherapy and ATRA plus chemotherapy and the role of maintenance therapy in newly diagnosed acute promyelocytic leukemia. The European APL Group. *Blood.* 1999;94(4):1192–1200.
36. Druker BJ, Guilhot F, O'Brien SG, et al. Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *N Engl J Med.* 2006; 355(23):2408–2417.
37. Schultz KR, Bowman WP, Aledo A, et al. Improved early event-free survival with imatinib in Philadelphia chromosome-positive acute lymphoblastic leukemia: a children's oncology group study. *J Clin Oncol.* 2009;27(31): 5175–5181.
38. Coiffier B, Thieblemont C, Van Den Neste E, et al. Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d'Etudes des Lymphomes de l'Adulte. *Blood.* 2010;116(12):2040–2045.
39. Pfreundschuh M, Trumper L, Osterborg A, et al. CHOP-like chemotherapy plus rituximab versus CHOP-like chemotherapy alone in young patients with good-prognosis diffuse large-B-cell lymphoma: a randomised controlled trial by the MabThera International Trial (MInT) Group. *Lancet Oncol.* 2006;7(5):379–391.
40. 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.* 2001;344(11):783–792.
41. McCarthy P, Owzar K, Anderson K, et al.Phase III intergroup study of lenalidomide versus placebo maintenance therapy following single autologous stem cell transplant (ASCT) for multiple myeloma (MM): CALGB ECOG BMT-CTN 100104. *Haematologica.* 2011;96(s1):S23.

Funding

This work was supported, in part, by National Cancer Institute (R01 CA090998 to M.L.).

Affiliations of authors: Biometric Research Branch, National Cancer Institute, Bethesda, MD (LR); Fred Hutchinson Cancer Research Center, Seattle, WA (MLB); Cancer Therapy Evaluation Program, National Cancer Institute, Bethesda, MD (MAS).