

Randomized Phase II Trial Designs With Biomarkers

Boris Freidlin, Lisa M. McShane, Mei-Yin C. Polley, and Edward L. Korn

All authors: National Cancer Institute, Bethesda, MD.

Submitted April 4, 2012; accepted June 19, 2012; published online ahead of print at www.jco.org on August 6, 2012.

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

Corresponding author: Boris Freidlin, PhD, Biometric Research Branch, EPN-8129, National Cancer Institute, Bethesda, MD 20852; e-mail: freidlinb@ctep.nci.nih.gov.

© 2012 by American Society of Clinical Oncology

0732-183X/12/3026-3304/\$20.00

DOI: 10.1200/JCO.2012.43.3946

A B S T R A C T

Efficient development of targeted therapies that may only benefit a fraction of patients requires clinical trial designs that use biomarkers to identify sensitive subpopulations. Various randomized phase III trial designs have been proposed for definitive evaluation of new targeted treatments and their associated biomarkers (eg, enrichment designs and biomarker-stratified designs). Before proceeding to phase III, randomized phase II trials are often used to decide whether the new therapy warrants phase III testing. In the presence of a putative biomarker, the phase II trial should also provide information as to what type of biomarker phase III trial is appropriate. A randomized phase II biomarker trial design is proposed, which, after completion, recommends the type of phase III trial to be used for the definitive testing of the therapy and the biomarker. The recommendations include the possibility of proceeding to a randomized phase III of the new therapy with or without using the biomarker and also the possibility of not testing the new therapy further. Evaluations of the proposed trial design using simulations and published data demonstrate that it works well in providing recommendations for phase III trial design.

J Clin Oncol 30:3304-3309. © 2012 by American Society of Clinical Oncology

INTRODUCTION

The primary purpose of phase II trials is to screen whether new therapies are worthy of definitive testing in randomized phase III trials. In the development of a targeted therapy, there is often a biomarker that can potentially identify patients for whom the therapy will work. In this case, it is desirable that the phase II trial address not only whether the therapy should move forward to phase III definitive evaluation, but also whether this definitive evaluation should include the biomarker. There are various ways one can incorporate a biomarker into a phase III trial design.¹ The phase III design that provides the most information about the biomarker and the therapy is the biomarker-stratified design, in which the treatment assignment is randomized between the new and standard treatments for all patients, with a separate assessment of the treatment effect in the biomarker-positive and biomarker-negative subgroups. Because the trial needs to provide a sufficiently accurate estimate of the effect of the therapy in the biomarker-positive subgroup (and possibly the biomarker-negative subgroup), the required sample size will be larger than that for a standard phase III trial, which does not have a biomarker component. When the preliminary data suggest strongly that the therapy will not work in the biomarker-negative subgroup, an enrichment phase III trial that randomly assigns only biomarker-positive patients and assesses the effect of the therapy

in this subgroup is appropriate. Although an enrichment design has a smaller number of patients randomly assigned as compared with the biomarker-stratified design or standard phase III trial design (with no biomarker), it will need to screen for biomarker status approximately the same number of patients randomly assigned in a biomarker-stratified trial.

The purpose of this article is to propose a specific randomized phase II trial design that can be used to guide decision making for further development of an experimental therapy. In particular, the proposed approach is developed to optimize making one of four possible decisions after analysis of the phase II trial results (Table 1): one, perform a randomized phase III trial with a biomarker-enrichment design; two, perform a randomized phase III trial with a biomarker-stratified design; three, perform a randomized phase III trial without using the biomarker; and four, drop consideration of the new therapy. We evaluate the proposed design via simulation under different sets of hypothesized treatment effects for the biomarker-positive and biomarker-negative subgroups. We also evaluate the proposed designs by retrospectively assessing how they would have performed if they had been applied to some recently published clinical situations with completed randomized trials that considered treatment effects in biomarker-defined subgroups. We end with a discussion of some extensions and modifications of the proposed design.

Table 1. Possible Recommendations From Proposed Randomized Phase II Biomarker Trial Design for Future Phase III Trial Design

Recommendation
Biomarker-enrichment design
Biomarker-stratified design
Drop biomarker (standard phase III trial design)
No further testing of new therapy

PROPOSED TRIAL DESIGN

Randomized phase II trials typically use an intermediate end point like progression-free survival (PFS) to obtain the results earlier than overall survival (OS) and to be able to target larger treatment effects than one would target for OS. The proposed phase II design uses this end point. We assume that at the time the study is designed, there is some preliminary rationale suggesting that the biomarker-positive patients are likely to derive the most benefit from the new therapy, but benefit for the biomarker-negative patients cannot be ruled out. Patients are randomly assigned to the experimental and control treatments, and the biomarker status of each patient is recorded. The idea is to use the observed treatment effects (hazard ratios) in the biomarker-positive and -negative subgroups, as well as overall, to guide decision making.

With a specified sample size and after sufficient follow-up, a decision is made concerning further development of the new treat-

ment and biomarker as described in Figure 1. In step 1, the null hypothesis that PFS is the same for both treatment arms in the biomarker-positive subgroup is tested (at the one-sided .10 significance level). If this test in step 1 does not reject (ie, we have not demonstrated that the experimental treatment is better than the control in the biomarker-positive subgroup), then in step 2A, the null hypothesis that PFS is the same for both treatment arms for all randomly assigned patients is tested (at the one-sided .05 significance level). If this test in step 2A does not reject, the design recommends no further testing of the new therapy. If this test in step 2A does reject (ie, the new treatment is better than the control in the whole population), then the conclusion is that the new treatment is potentially active, but the biomarker is not useful, and the design recommends dropping the biomarker and performing a standard randomized phase III trial.

If the test rejects the null hypothesis in step 1 (ie, the new treatment is better than the control in the biomarker-positive subgroup), then the recommendation (step 2B) is based on a two-sided 80% CI for the hazard ratio (control over new treatment hazard) in the biomarker-negative subgroup. If the whole CI is below 1.3, then there is strong evidence that the new treatment is, at best, only marginally helpful for the biomarker-negative patients, and a biomarker-enrichment phase III trial is recommended. If the whole CI is above 1.5, then there is evidence that the treatment works sufficiently well in the biomarker-negative patients (and therefore the biomarker is not useful), and the recommendation is to drop the biomarker and perform a standard randomized phase III trial. If neither of these conditions holds, then a phase III biomarker-stratified design is recommended,

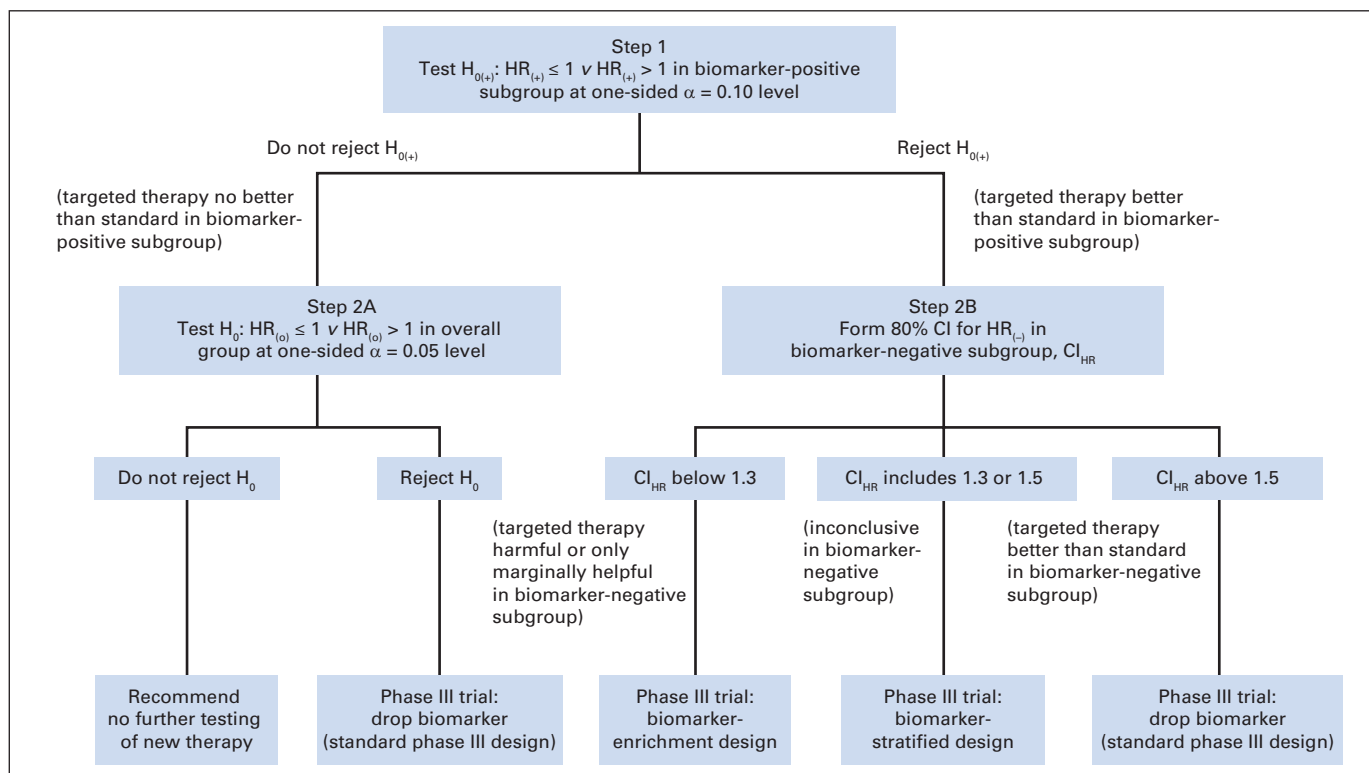


Fig 1. Decision algorithm for recommendation of phase III trial design based on the outcome of the proposed phase II biomarker trial design. H₀, null hypothesis in the overall group; H₀₍₊₎, null hypothesis in the biomarker-positive subgroup; HR, hazard ratio; HR₍₊₎, HR of standard therapy relative to targeted therapy in biomarker-positive subgroup; HR₍₋₎, HR of standard therapy relative to targeted therapy in biomarker-negative subgroup; HR₍₀₎, HR of standard therapy relative to targeted therapy in overall group.

because there is insufficient information on how well the treatment works in the biomarker-negative subgroup.

Sample Size Considerations

In addition to the decision rule at analysis time, as delineated in Figure 1, part of trial design is the specification of sample size to have sufficient power to make appropriate decisions. Because the underlying assumption is that the benefit of the new therapy is greatest among biomarker-positive patients, sample size considerations are driven by the biomarker-positive subgroup. Consider a clinical setting where the median PFS with standard treatment is relatively low (eg, < 6 months), and we hope that the experimental therapy will double the median PFS in the biomarker-positive subgroup, (ie, a hazard ratio of 2 in this subgroup). Detecting this effect with 90% power at the one-sided .10 significance level would require 56 PFS events (progressions or deaths) in the biomarker-positive subgroup. We would also want to have approximately at least this many events in the biomarker-negative subgroup to help with the decision making. There are various combinations of accrual and follow-up that will yield a specified number of events for a given median PFS (Appendix, online only).

Simulations

We perform simulations using the design for accrual and follow-up described in the Appendix (online only) and Figure 1 for decision making. We consider the eight scenarios for the true median PFS (and hazard ratios) in the biomarker-positive and biomarker-negative subgroups, and four possible prevalences for biomarker positivity (Table 2). For the different biomarker-positivity prevalences, the average simulated sample sizes in the biomarker-positive and -negative subgroups are 70 and 140 (20% prevalence), 70 and 133 (33% prevalence), 75 and 75 (50% prevalence), and 133 and 65 (67% prevalence), respectively.

When the treatment does not work in either the biomarker-positive or biomarker-negative subgroups (scenario 1), we drop consideration of using the new therapy 87% to 88% of the time (depending on the prevalence of positivity), which is the correct decision. When the experimental agent works only in the biomarker-positive subgroup (hazard ratio, 2; scenario 2), the best recommendation would be an enrichment design, with an acceptable recommendation being to perform a biomarker-stratified design (in which the utility of the biomarker would become apparent in that larger trial). For this scenario, the probability of either of these recommendations is $\geq 89\%$.

For scenarios 3 and 4, the experimental therapy works (hazard ratios, 1.5 and 1.75, respectively), but the biomarker is not useful. Under these scenarios, the best recommendation would be a phase III trial design with no biomarker, with an acceptable recommendation being to perform a biomarker-stratified design. For scenario 3, the probability of making one of these recommendations is $\geq 75\%$, and for scenario 4, the probability is $\geq 93\%$. Note that under scenario 3 (with hazard ratio of 1.5), the probability of no further testing is as high as 22%. This is a consequence of targeting a 2.0 hazard ratio in the biomarker-positive subgroup and the limited sample size of a phase II trial, as described in Discussion.

For scenarios 5 and 6, the experimental therapy works in the biomarker-positive subgroup (hazard ratio, 2.0), but not as well in the biomarker-negative subgroup (hazard ratios, 1.5 and 1.25, respectively). For scenario 5, the best recommendation would be a

biomarker-stratified design or a phase III trial with no biomarker; the probability of making one of these recommendations is $\geq 93\%$. For scenario 6, the best recommendation would be an enrichment design, with an acceptable recommendation being to perform a biomarker-stratified design; the probability of making one of these recommendations is $\geq 89\%$.

Scenario 7 represents a situation in which the experimental therapy works in the biomarker-positive subgroup (but with a hazard ratio of 1.75 and not 2.0), and is slightly harmful in the biomarker-negative subgroup (hazard ratio, 0.75). The best recommendation would be an enrichment design, with an acceptable recommendation being a biomarker-stratified design. (Presumably, the biomarker-stratified design would include the possibility of stopping early for futility/inefficacy in the biomarker-negative subgroup.¹) The probability of making one of these recommendations is 78% to 93%. There is a fairly high probability (approximately 20%) of abandoning the experimental treatment when the positivity prevalence is $\leq 50\%$, similar to scenario 3.

Finally, when the biomarker is prognostic but the treatment does not work in either biomarker subgroup (scenario 8), then the recommendation is for no further testing of the new therapy 87% to 89% of the time.

RETROSPECTIVE EVALUATION OF PUBLISHED DATA

To illustrate application of the proposed design, we use summary data from previously published randomized phase II and III trials that evaluated biomarker-subgroup treatment effects. Table 3 contains the summary statistics for the biomarker-subgroup treatment effects along with the hypothetical phase III that would have been recommended using our proposed approach. The evaluation was performed by simulating 50,000 phase II trial data sets, treating the observed summary statistics in Table 3 as true values, and then using the sample sizes and follow-up as described in the Appendix (online only) and the decision-making algorithm as described in Figure 1.

In trials 1 to 4, the hazard ratio in the biomarker-positive subgroup is ≥ 2 (in contrast to trials 5 and 6), corresponding to the treatment effect exceeding the targeted effect for the design. For these four settings, application of the proposed trial design recommended no future testing $\leq 2\%$ of the time. The probabilities of the other three recommendations depend on the hazard ratio in the biomarker-negative subgroup in a logical way (eg, recommending an enrichment design almost always when the hazard ratio is 0.35 in trial 2).

Trial 6 illustrates a situation in which there is a clear biomarker-by-treatment interaction (ie, the hazard ratios are different for the biomarker subgroups), but with only a modest benefit of the new therapy in the biomarker-positive patients and no benefit in the biomarker-negative patients. The correct decision here is arguably an enrichment design or to have no further testing, depending on whether one considers a PFS hazard ratio of 1.47 clinically meaningful. However, the proposed design recommends a biomarker-stratified design 37% of the time, reflecting the limitations of a relatively small sample size. A similar limitation of a small sample size is seen in the probabilities of the recommendations in trial 7, where because the new treatment is potentially not

Randomized Phase II Designs With Biomarkers

Table 2. Simulated Probabilities of Four Decision Outcomes of Proposed Randomized Phase II Biomarker Design Under Different Scenarios

Scenario	Biomarker Subgroup						Probability of Recommendation for Trial Design of Future Phase III Testing				
	Positive			Negative			Enrichment Design	Biomarker Stratified	No Biomarker	No Further Testing	
	Experimental Treatment Median PFS	Control Median PFS	HR	Experimental Treatment Median PFS	Control Median PFS	HR					
Prevalence of biomarker positive, 20%											
1	4	4	1.0	4	4	1.0	0.06	0.04	0.03	0.87	
2	8	4	2.0	4	4	1.0	0.53	0.36	0.01	0.10	
3	6	4	1.5	6	4	1.5	0.01	0.52	0.38	0.10	
4	7	4	1.75	7	4	1.75	< 0.01	0.51	0.48	0.01	
5	8	4	2.0	6	4	1.5	0.02	0.79	0.18	0.02	
6	8	4	2.0	5	4	1.25	0.13	0.76	0.06	0.06	
7	7	4	1.75	3	4	0.75	0.76	0.02	< 0.01	0.22	
8	6	6	1.0	4	4	1.0	0.06	0.04	0.02	0.87	
Prevalence of biomarker positive, 33%											
1	4	4	1.0	4	4	1.0	0.06	0.05	0.03	0.87	
2	8	4	2.0	4	4	1.0	0.50	0.40	0.01	0.09	
3	6	4	1.5	6	4	1.5	0.01	0.52	0.34	0.14	
4	7	4	1.75	7	4	1.75	< 0.01	0.54	0.44	0.02	
5	8	4	2.0	6	4	1.5	0.02	0.79	0.17	0.03	
6	8	4	2.0	5	4	1.25	0.13	0.76	0.06	0.06	
7	7	4	1.75	3	4	0.75	0.75	0.04	< 0.01	0.22	
8	6	6	1.0	4	4	1.0	0.06	0.05	0.02	0.88	
Prevalence of biomarker positive, 50%											
1	4	4	1.0	4	4	1.0	0.04	0.06	0.02	0.88	
2	8	4	2.0	4	4	1.0	0.38	0.53	< 0.01	0.08	
3	6	4	1.5	6	4	1.5	0.02	0.52	0.23	0.22	
4	7	4	1.75	7	4	1.75	0.01	0.60	0.33	0.06	
5	8	4	2.0	6	4	1.5	0.03	0.79	0.14	0.04	
6	8	4	2.0	5	4	1.25	0.12	0.77	0.05	0.07	
7	7	4	1.75	3	4	0.75	0.66	0.15	< 0.01	0.20	
8	6	6	1.0	4	4	1.0	0.04	0.06	0.02	0.88	
Prevalence of biomarker positive, 67%											
1	4	4	1.0	4	4	1.0	0.04	0.06	0.02	0.88	
2	8	4	2.0	4	4	1.0	0.38	0.61	< 0.01	0.01	
3	6	4	1.5	6	4	1.5	0.03	0.68	0.16	0.14	
4	7	4	1.75	7	4	1.75	0.01	0.71	0.25	0.02	
5	8	4	2.0	6	4	1.5	0.03	0.85	0.11	0.01	
6	8	4	2.0	5	4	1.25	0.13	0.83	0.03	0.01	
7	7	4	1.75	3	4	0.75	0.73	0.20	< 0.01	0.06	
8	6	6	1.0	4	4	1.0	0.04	0.06	0.01	0.89	

NOTE. Simulations are based on 50,000 replications. Abbreviations: HR, hazard ratio; PFS, progression-free survival.

useful, the recommendation for no further testing should ideally be higher than 45% in this setting.

It is interesting to note that for the two phase II studies in Table 3 (studies 1 and 5), the most likely recommendation from the proposed design was an enrichment trial. In fact, study 5 was followed by an enrichment phase III trial in Met-positive patients (NCT01456325), and study 1 was followed by an enrichment phase II trial in KRAS-mutated patients (NCT01395758).

DISCUSSION

We have targeted a doubling of median PFS (hazard ratio of 2) in the biomarker-positive subgroup, a value that is larger than one would

typically use for a screening randomized phase II trial in an unselected population. However, for a targeted agent that only benefits a fraction of the population, a hazard ratio of 2 in the targeted population corresponds to a smaller hazard ratio in the unselected population. Moreover, when median PFS values are low, recent experience with targeted therapies demonstrates the feasibility of achieving such effects.^{2,10-15} Therefore, we believe that targeting a hazard ratio of 2 in a biomarker-positive subgroup is appropriate when the median PFS is relatively short (eg, < 6 months). It is possible to target a smaller PFS hazard ratio than 2 in the biomarker-positive subgroup (eg, in settings with median PFS > 6 months), but this would require a larger phase II sample size. One can use the computer program (available at <http://brb.nci.nih.gov/Data/FreidlinB/RP2BM>) to modify the proposed

Table 3. Retrospective Evaluation of Proposed Randomized Phase II Biomarker Design Using Historical Trial Data

Trial	Biomarker (positivity prevalence)	Historical Data				Simulation/Evaluation of Proposed Phase II Design					
		Biomarker Subgroup				Average Sample Size		Probability of Recommendation for Trial Design for Future Phase III Testing			
		Positive		Negative		Positive	Negative	Enrichment Design	Biomarker Stratified	No Biomarker	No Further Testing
		Control Median PFS (months)	HR*	Control Median PFS (months)	HR*						
Erlotinib ± tivantinib ^{2,3}	<i>KRAS</i> mutation (14%)	1.0	5.56	2.3	1.01	70	140	0.62	0.38	< 0.01	0.00
Gefitinib v chemotherapy ⁴	<i>EGFR</i> mutation (60%)	6.0	2.08	2.0	0.35	105	70	0.98	< 0.01	0.00	0.02
Radiation ± temozolomide ⁵	MGMT (45%)†	5.9	2.08	4.4	1.61	71	86	0.01	0.75	0.21	0.02
BSC ± cetuximab ⁶	<i>KRAS</i> wild type (58%)	1.9	2.50	1.8	1.01	97	70	0.38	0.61	< 0.01	0.01
Erlotinib ± MetMab ⁷	Met (54%)	1.6	1.79	2.9	0.50	83	71	0.84	0.01	0.00	0.15
FOLFIRI ± cetuximab ⁸	<i>KRAS</i> wild type (64%)	8.7	1.47	8.1	0.93	122	69	0.36	0.37	0.01	0.26
Capecitabine ± bevacizumab ⁹	VEGF-A (37%)	2.9	1.39	4.11	1.12	70	118	0.13	0.34	0.08	0.45

Abbreviations: BSC, best supportive care; FOLFIRI, fluorouracil, leucovorin, and irinotecan; HR, hazard ratio; MGMT, methylguanine DNA-methyltransferase; PFS, progression-free survival; VEGF-A, vascular endothelial growth factor A.

*HRs > 1 represent the experimental therapy performing better than the control therapy.

†Methylation of MGMT promoter.

design to target a smaller hazard ratio by adjusting the sample size and CI cutoffs in step 2B of Figure 1. Another computer program allows use of response rates instead of PFS as the phase II trial end point (Appendix, online only).

In practice, randomized evaluation of a new therapy is implemented either by adding the new therapy (A) to the standard of care (B) using the so-called add-on design that randomly assigns patients to either A + B or B (eg, trials 1 and 3 to 7 in Table 3), or by a head-to-head comparison of the new therapy against the standard of care, in which case patients are randomly assigned to either therapy A or therapy B (eg, trial 2 in Table 3). In settings where the standard of care has proven clinical benefit, the design should incorporate an aggressive interim inefficacy/futility look, especially in biomarker-negative patients.¹⁶ For example, we recommend the following rule in each biomarker subgroup: if after half of the required events are observed, the estimate of the hazard ratio (control over experimental) is ≤ 1 , then accrual to the subgroup stops (if the biomarker-positive subgroup is stopped for futility, then the entire study is stopped).¹ Another practical concern is that the biomarker status may not be available for a fraction of patients. In theory, when this fraction is relatively low, investigators may consider randomly assigning these patients and including them in the overall comparison in step 2A.

The recommended phase II trial design with biomarkers can be incorporated into a phase II/III design strategy.¹⁷ This strategy allows one to use phase II patients in the phase III evaluation and streamlines transition from phase II to III components of drug development. After the patients have been accrued on the phase II trial and have been observed for the required number of phase II events in the biomarker-positive subgroup, the decision is made as to whether to proceed to phase III evaluation and, if so, what phase III trial design to recommend. If an enrichment phase III design is recommended, accrual

begins again but only in the biomarker-positive subgroup. If a biomarker-stratified or a phase III design without the biomarker is recommended, the trial continues accruing all patients. Note that in the biomarker setting, because of the data-driven phase III trial design selection, the phase II/III design strategy needs to adjust for inflation in the nominal phase III type I error.

We have assumed that a randomized phase II trial will be performed before embarking on a phase III trial. In some situations, early nonrandomized data in a targeted subgroup will be so dramatic as to lead directly to enrichment phase III trials restricted to the targeted subgroup. For example, a dramatic response rate for crizotinib in early studies in patients with ALK-positive non-small-cell lung cancer led to additional phase II and III trials restricted to ALK-positive patients.^{18,19} However, there is the possibility that without sufficient testing of biomarker-negative patients, further testing of the therapy may be unnecessarily restricted. For example, there was a 26% to 35% response rate seen for 23 patients with ALK-negative non-small-cell lung cancer,¹⁹ leading to a question as to whether the phase III trials should have been restricted to ALK-positive patients. In this case, results of a randomized phase II trial using the proposed design could have helped to inform the choice of the phase III trial design.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Manuscript writing: All authors

Final approval of manuscript: All authors

REFERENCES

1. Freidlin B, McShane LM, Korn EL: Randomized clinical trials with biomarkers: Design issues. *J Natl Cancer Inst* 102:152-160, 2010
2. Sequist LV, von Pawel J, Garney EG, et al: Randomized phase II study of erlotinib plus tivantinib versus erlotinib plus placebo in previously treated non-small-cell lung cancer. *J Clin Oncol* 29:3307-3315, 2011
3. Sequist LV, Akerley WL, Brugger W, et al: Final results from arq 197-209: A global randomized placebo-controlled phase 2 clinical trial of erlotinib plus arq 197 versus erlotinib plus placebo in previously treated egfr-inhibitor naïve patients with advanced non-small cell lung cancer. Presented at the European Society of Clinical Oncology Congress, Milan, Italy, October 8-12, 2010 (abstr 3630)
4. Mok TS, Wu YL, Thongprasert S, et al: Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 361:947-957, 2009
5. Hegi ME, Diserens AC, Gorlia T, et al: MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med* 352:997-1003, 2005
6. Karapetis CS, Khambata-Ford S, Jonker DJ, et al: K-ras mutations and benefit from cetuximab in advanced colorectal cancer. *N Engl J Med* 359:1757-1765, 2008
7. Spiegel DR, Ervin TJ, Ramlau R, et al: Randomized, phase II multicenter, double-blind, placebo-controlled study evaluating MetMAB, an antibody to Met receptor, in combination with erlotinib, in patients with advanced non-small-cell lung cancer (OAM4558g). Presented at European Society of Clinical Oncology Congress, Milan, Italy, October 8-12, 2010 (abstr LBA15)
8. Van Cutsem E, Köhne CH, Hitre E, et al: Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med* 360:1408-1417, 2009
9. Jubb AM, Miller KD, Rugo HS, et al: Impact of exploratory biomarkers on the treatment effect of bevacizumab in metastatic breast cancer. *Clin Cancer Res* 17:372-281, 2011
10. Demetri GD, van Oosterom AT, Garrett CR, et al: Efficacy and safety of sunitinib in patients with advanced gastrointestinal stromal tumour after failure of imatinib: A randomised controlled trial. *Lancet* 368:1329-1338, 2006
11. Escudier B, Eisen T, Stadler WM, et al: Sorafenib for treatment of renal cell carcinoma: Final efficacy and safety results of the phase III treatment approaches in renal cancer global evaluation trial. *J Clin Oncol* 27:3312-3318, 2009
12. Dematteo RP, Ballman KV, Antonescu CR, et al: Adjuvant imatinib mesylate after resection of localised, primary gastrointestinal stromal tumour: A randomised, double-blind, placebo-controlled trial. *Lancet* 373:1097-1104, 2009
13. 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* 28:1061-1068, 2010
14. Yao JC, Shah MH, Ito T, et al: Everolimus for advanced pancreatic neuroendocrine tumors. *N Engl J Med* 364:514-523, 2011
15. Chapman PB, Hauschild A, Robert C, et al: Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Med* 364:2507-2516, 2011
16. Freidlin B, Korn EL: Monitoring for lack of benefit: A critical component of a randomized clinical trial. *J Clin Oncol* 27:629-633, 2009
17. Korn EL, Freidlin B, Abrams JS, et al: Design issues in randomized phase II/III trials. *J Clin Oncol* 30:667-671, 2012
18. Kwak EL, Bang YJ, Camidge DR, et al: Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med* 363:1693-1703, 2010
19. Center for Drug Evaluation and Research: Summary review for regulatory action: Crizotinib. http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/202570Orig1s000SumR.pdf

