VOLUME 26 · NUMBER 4 · FEBRUARY 1 2008

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

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

Edward L. Korn, Ping-Yu Liu, Sandra J. Lee, Judith-Anne W. Chapman, Donna Niedzwiecki, Vera J. Suman, James Moon, Vernon K. Sondak, Michael B. Atkins, Elizabeth A. Eisenhauer, Wendy Parulekar, Svetomir N. Markovic, Scott Saxman, and John M. Kirkwood

From the Biometric Research Branch, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD; Southwest Oncology Group Statistical Office, Seattle, WA; Eastern Cooperative Oncology Group; and Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA; National Cancer Institute of Cancer Clinical Trials Group, Queen's University, Kingston, Ontario, Canada; Cancer and Leukemia Group B Statistical Center, Durham, NC; Mayo Clinic, Rochester, MN; H. Lee Moffitt Cancer Center and Research Institute, University of South Florida, Tampa, FL; Peace Corps, Washington, DC; and Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA.

Submitted May 25, 2007; accepted August 28, 2007.

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

Corresponding author: Edward L. Korn, PhD, Biometric Research Branch, EPN-8129, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD 20892; e-mail: korne@ctep.nci.nih.gov.

0732-183X/08/2604-527/$20.00

DOI: 10.1200/JCO.2007.12.7837

ABSTRACT

Purpose

Objective tumor response rates observed in phase II trials for metastatic melanoma have historically not provided a reliable indicator of meaningful survival benefits. To facilitate using overall survival (OS) or progression-free survival (PFS) as an endpoint for future phase II trials, we evaluated historical data from cooperative group phase II trials to attempt to develop benchmarks for OS and PFS as reference points for future phase II trials.

Patients and Methods

Individual-level and trial-level data were obtained for patients enrolled onto 42 phase II trials (70 trial arms) that completed accrual in the years 1975 through 2005 and conducted by Southwest Oncology Group, Eastern Cooperative Oncology Group, Cancer and Leukemia Group B, North Central Cancer Treatment Group, and the Clinical Trials Group of the National Cancer Institute of Canada. Univariate and multivariate analyses were performed to identify prognostic variables, and between-trial(-arm) variability in 1-year OS rates and 6-month PFS rates were examined.

Results

Statistically significant individual-level and trial-level prognostic factors found in a multivariate survival analysis for OS were performance status, presence of visceral disease, sex, and whether the trial excluded patients with brain metastases. Performance status, sex, and age were statistically significant prognostic factors for PFS. Controlling for these prognostic variables essentially eliminated between-trial variability in 1-year OS rates but not in 6-month PFS rates.

Conclusion

Benchmarks are provided for 1-year OS or OS curves that make use of the distribution of prognostic factors of the patients in the phase II trial. A similar benchmark for 6-month PFS is provided, but its use is more problematic because of residual between-trial variation in this endpoint.

J Clin Oncol 26:527-534. © 2008 by American Society of Clinical Oncology

INTRODUCTION

New agents are needed for the treatment of metastatic melanoma because no evidence of survival prolongation with existing therapy has been established. Phase II clinical trials offer a means to screen therapies for further testing in the phase III setting. However, phase II designs require benchmarks for deciding whether a new therapy is worth pursuing. Historically, phase II trials have tended to use objective response rate (tumor shrinkage) as a benchmark. However, the few therapies that have demonstrated promising response rates in patients with metastatic stage IV melanoma have not had meaningful effects on survival. In addition, tumor shrinkage as a benchmark may not be appropriate for targeted molecular and immunologic therapies that could offer survival benefits for patients manifesting only disease stabilization. Both of these considerations suggest that it might be more appropriate to use an overall survival (OS) or progression-free survival (PFS) endpoint as a benchmark for future phase II trials.[1] It may be possible that such benchmarks derived from historical data will allow more effective selection of new regimens for phase III testing on the basis of results obtained in future cooperative group phase II trials (avoiding the cost and time of concurrent control arms as a reference). The development of these benchmarks has been investigated by performing a meta-analysis of previously collected data from

metastatic melanoma phase II trials performed by the participating cooperative groups.

PATIENTS AND METHODS

Individual and Trial-Level Data

Individual patient data from participants in all metastatic melanoma phase II trials (including randomized phase II trials) that completed accrual in the years 1975 through 2005 were requested. Trials were identified both by a literature search and by searching records in the statistical offices of the participating groups. Individual-level variables considered were sex, age, Eastern Cooperative Oncology Group performance status (PS), presence of visceral metastases (VISC), serum lactate dehydrogenase (LDH) level, OS, and PFS. OS was measured from the time of registration to death (or censoring); PFS was measured from the time of registration to progression (typically, a 25% to 50% increase in the sum of bidimensional products of all measurable lesions or appearance of new lesions) or death from any cause (or censoring). Trial-level variables considered were exclusion of patients with brain metastases (BRAIN-METS), exclusion of patients with liver metastases, exclusion of patients with visceral metastases, previous treatment for metastatic disease, and the year during which accrual was completed. For the purposes of this study, all treatments were considered inactive.

Statistical Methods

OS and PFS distributions were estimated using Kaplan-Meier plots. Prognostic variables for OS and PFS were assessed using likelihood-ratio tests from proportional hazards modeling. Prognostic variables for event rates (1-year OS rate and 6-month PFS rate) were assessed using logistic regression models; the small number of observations that were censored before the relevant time point (eg, 1 year) were omitted. The reported P values for statistical significance are all two-sided and unadjusted for multiple comparisons, with $P \leq .05$ taken as being statistically significant.

Exploration of between trial-arm variability in event rates was examined by two methods (detailed in Appendix A, online only). In one method, the event rate for each of the treatment arms was compared with the overall event rate to look for outliers. The second method used a logistic-normal model, where the variance of the normal effect is an estimate of the underlying between trial-arm variability not accounted for by differing distributions of prognostic variables. If this variability is large, then comparisons to a historical control rate are problematic[2] and may lead to too many false-positive trials (Appendix A).

All analyses were performed using SAS (Version 9.1; SAS Institute, Cary, NC).

RESULTS

Individual-level data were obtained from 42 trials (70 trial arms) involving 2,100 patients (Appendix B, online only). The years of closure to accrual ranged from 1977 through 2005. LDH level was available for only 136 patients and was not considered further. Two trial-level variables were not considered because all the trials included patients with visceral disease, and only two trials (n = 55) excluded patients with liver metastases.

Prognostic Variables for OS

For the 2,100 patients, 73 patients had censored observations for OS (either lost to follow-up or still alive at the time of this data collection), 25 patients of whom had censored observations before 365 days. The median survival time was 6.2 months (95% CI, 5.9 months to 6.5 months), with 25.5% (95% CI, 23.6% to 27.4%) alive at 1 year (Fig 1A). In the univariate analyses, PS was the most important individual-level prognostic variable (Fig 1B), followed by presence of visceral disease (Fig 1C) and sex (Fig 1D); age was not prognostic (Fig 1E). Whether the trial excluded patients with brain metastases (Fig 1F) and the year during which the trial closed (Fig 1G) were important trial-level prognostic variables; whether the trial excluded patients who had previous treatment was not prognostic (Fig 1H). Details of the univariate analyses are given in Appendix Table A1, online only.

Multivariate analyses were performed using the variables that had prognostic ability in the univariate analyses (Table 1). Because a patient could not be included in these multivariate analyses if data were missing or unavailable on any of these variables, only 1,278 patients were used in this analysis. The majority of the missing or unavailable data was due to visceral disease status not being available (n = 790), in particular, for all trials closed before 1982 (n = 576), two trials in 1984 (n = 36), and single trials in 1987 (n = 33), 1990 (n = 19), and 1991 (n = 15). There were only nine patients with PS of 3 with available multivariate data; these patients were pooled with the 100 patients with PS of 2 for all multivariate analyses. To enable a better comparison of the univariate and multivariate results, we have also included in Table 1 a repeat of the univariate analyses restricted to the patients with complete data. The multivariate analyses show that PS, VISC, sex, and BRAIN-METS are statistically significant prognostic factors for OS. When pairwise interactions of these four variables were considered in addition to their main effects, there were no statistically significant interactions.

Prognostic Variables for PFS

Five trials had no progression data available (n = 103), and three additional patients had missing progression data, leaving a sample size of n = 1,994. Of these 1,994 patients, 1,971 had a PFS event (progression or death) and 23 had censored observations, two of which occurred before 6 months. The median PFS was 1.7 months (95% CI, 1.6 months to 1.8 months), with 14.5% (95% CI, 12.9% to 16.1%) progression free at 6 months (Fig 2A). In the univariate analyses, PS was the most important individual-level prognostic variable (Fig 2B). Age, sex, whether the patient had visceral disease, and whether the trial excluded patients with brain metastases were also prognostic to a lesser degree, with only sex additionally having statistically significant associations with 6-month PFS rates (Appendix Table A2, online only; Figs 2C through 2H).

Multivariate analysis using the variables that were statistically significant in any of the univariate analyses are listed in Table 2. PS, sex, and age are statistically significant, although the hazard ratios associated with age and sex are relatively small. When pairwise interactions of these three variables were considered in addition to their main effects, there was a statistically significant interaction between PS and age ($P = .031$) for the 6-month PFS rates and no statistically significant interactions for the PFS distributions.

Trial-Arm Variability in 1-Year OS Rates and 6-Month PFS Rates

Two trial arms with sample sizes of fewer than 10 patients were pooled, leaving 68 trial arms (Appendix B, online only). Figure 3A shows the 1-year OS rates for the trial arms plotted against the sample size for the trial arm. The 95% confidence bounds suggest that no trial arm has a statistically different rate from the overall mean 1-year rate of 25% (524 of 2,075 patients). The logistic-normal model results for OS rates demonstrate a statistically significant between trial-arm variance component that is essentially eliminated with control for the prognostic variables (Appendix Table A3, online only).

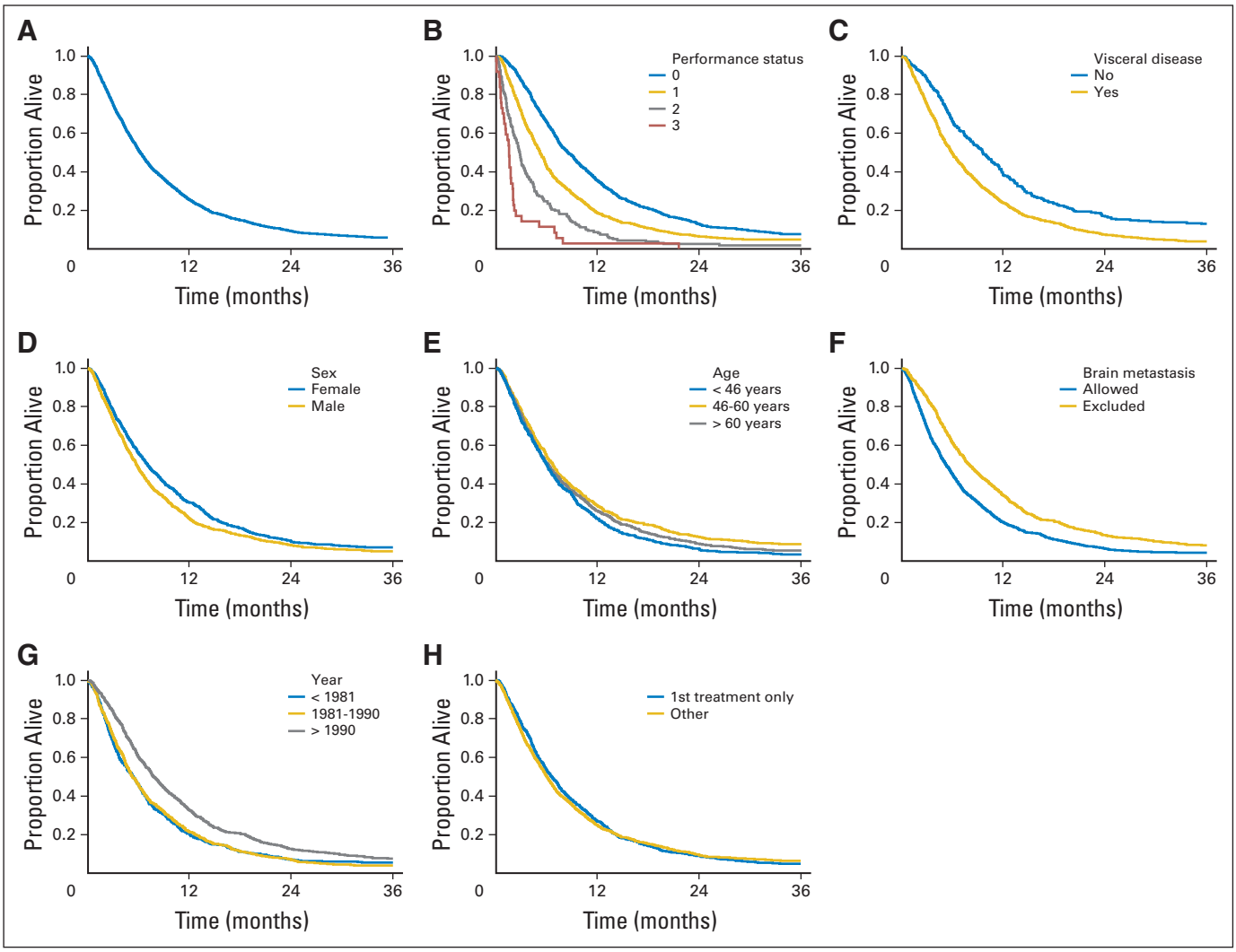

Fig 1. Overall survival (A) categorized by (B) performance status, (C) presence/absence of visceral disease, (D) sex, (E) age, (F) whether the trial excluded/allowed patients with brain metastases, (G) year trial closed, and (H) whether the trial excluded/allowed previous treatment.

Figure 3B shows the 6-month PFS rates for the trial arms plotted against the sample size for the trial arm. The 95% confidence bounds suggest that one trial (Southwest Oncology Group S9348[3]) has a 6-month PFS rate (30%) that differs from the overall mean 6-month PFS rate of 15% (298 of 1,992 patients). The favorable PS distribution of the 79 patients on this trial (59 patients with PS of 0; 20 patients with PS of 1) does not alone explain the high rate. The logistic-normal model results for PFS rates demonstrate a statistically significant between trial-arm variance component that is not eliminated when controlling for PS or the other variables, even when S9348 is omitted from the analysis (Appendix Table A4, online only). The implications of this residual between-trial variation are discussed below.

Benchmarks for Future Phase II Trials

Regardless of whether previous trials showed between-trial variation in survival rates, future trials may have different rates than in the past because of patient mixes that differ in terms of prognostic variables. To address this, we consider defining the null hypothesis target for a phase II trial based on the prognostic variables recorded in the trial. Table 3 contains the relevant information for a trial using a 1-year OS rate as the endpoint. These predicted values are based on a logistic regression analysis with effects included for PS, sex, VISC, and BRAIN-METS.

We recommend the following to analyze a phase II trial using Table 3. For each patient on the trial, obtain his or her predicted 1-year OS rate from the top half of the table if patients with brain metastases are excluded in the trial, or from the bottom half of the table if patients with brain metastases are allowed in the trial. Let π be the average of these predicted values for the patients in the trial (ie, the historical control rate). After the trial is complete, calculate the proportion of patients alive at 1 year. Declare the treatment worthy of further study if null hypothesis that the 1-year OS rate $\leq \pi$ can be rejected with a P value less than .10. A CI for the difference between the observed proportion and π should also be calculated.

What should the sample size be? If an expected 1-year OS rate for phase II trials conducted at the participating institution(s) is available, we recommend the following: Let π_0 be the expected rate. Choose the sample size (using the binomial distribution) so that a trial testing the null hypothesis that the 1-year OS rate $\leq \pi_0$ will have 90% power to

Table 1. Multivariate Analysis of Overall Survival and Comparison With Univariate Analyses

Variable	No. of Patients*	Overall Survival Distribution			1-Year Overall Survival Rates		
		Univariate† HR	Multivariate		Univariate‡ OR	Multivariate	
			Adjusted HR§	*P*‖		Adjusted OR¶	*P*#
Performance status							
0	639	1.00**	1.00	< .0001	1.00††	1.00	< .0001
1	530	1.56	1.55		2.53	2.59	
2-3	109	2.90	2.58		5.79	4.68	
Visceral disease							
No	277	1.00	1.00	< .0001	1.00	1.00	< .0001
Yes	1,001	1.54	1.53		2.00	1.95	
Sex							
Female	496	1.00	1.00	< .0001	1.00	1.00	< .0001
Male	782	1.22	1.28		1.66	1.78	
Brain metastases							
Excluded	705	1.00	1.00	.0012	1.00	1.00	< .0001
Allowed	573	1.46	1.33		1.96	2.36	
Year closed (continuous)	1,278	0.76‡‡	0.97‡‡	NS	0.74‡‡	1.37‡‡	NS

Abbreviations: HR, hazard ratio; OR, odds ratio; NS, not significant.
*Sample sizes for overall survival distribution comparisons; sample sizes for 1-year survival rate comparisons are slightly smaller.
†Analyses restricted to 1,278 individuals who have data available for all the variables listed.
‡Analyses restricted to 1,257 individuals who have data available for all the variables listed and whose data was not censored before the 1-year time point.
§Adjusted HR is adjusted for the other variables listed.
‖*P* value is testing the association of the variable and overall survival in a multivariate analysis that controls for the other variables listed.
¶Adjusted OR is adjusted for the other variables listed.
#*P* value is testing the association of the variable and the overall survival rate at 1 year in a multivariate analysis that controls for the other variables listed.
**First listed category for categorical variables is always the reference category for HRs.
††First listed category for categorical variables is always the reference category for ORs.
‡‡Reported HR here is for a difference in year of closure of 12 years, with a value less than 1 suggesting that more recent trials have better survival.

detect the alternative hypothesis that the 1-year OS rate is more than π_0+15%. If no expected rate is available, use $\pi_0 = 35\%$ (yielding a sample size of 72 patients). A trial with 72 patients will have 85% to 90% power to detect an increase of 15 percentage points in the 1-year OS rate over the historical control rate (with one-sided type 1 error $\leq 10\%$).

Alternatively, one can calculate an historical OS survival curve (Appendix C, online only), which can then be used for comparison with the observed phase II OS data on the new trial, again using a *P* value of less than .10 to decide whether the new regimen should be pursued. This latter approach will lead to a smaller sample size. For example, with 1 year of accrual and 1 year of follow-up, a sample size of 63 patients (instead of 72 patients) would be required to detect a hazard ratio of 1.51, which corresponds to an improvement in 1-year OS from 35% to 50%.

For 6-month PFS rates, the same approach can be used, except that the calculation of the benchmark 6-month PFS rate depends only on the PS of the patients on the trial (this being by far the most important prognostic variable). In particular, one calculates the average π of the predicted values for the patients in the trial using the predicted rates of 18.0% for PS 0 patients, 12.3% for PS 1 patients, 7.4% for PS 2 patients, and 2.9% for PS 3 patients (Appendix Table A2, online only). The sample size of the trial can again be chosen to detect a 15 percentage point improvement over the historical rate π_0 of 6-month PFS. If no historical rate is available, use $\pi_0 = 15\%$, yielding a sample size of 53 patients. However, because of the between-trial variability in PFS rates, the true type 1 error for phase II trials using this approach may be larger than the nominal 10%. For example, if the between-trial variance were 0.191 (Appendix Table A4, online only), then the actual type 1 error could possibly be as high as 80%, although there is the possibility of using a value larger than π for the null hypothesis to lessen the type 1 error (Appendix A). We do not recommend comparisons with the whole historical control PFS curve, as assessment frequencies may unduly influence this curve.[5]

DISCUSSION

Combinations of the prognostic variables for OS found in this study (PS, VISC, BRAIN-METS, and sex) have been noted in other studies of metastatic melanoma,[6-17] including some studies[18-21] whose trials partially overlap with the trials considered here. Additional variables not available for analysis here have been found to be prognostic for overall survival, including LDH and other laboratory biomarkers,[11,12,16,17,19,22] number of metastatic sites,[10,14,18,19,20] and time from diagnosis to metastases.[7,10,13,15] To our knowledge, prognostic variables for PFS have not been studied, so that the finding of the prognostic ability of PS and lack of important prognostic ability of the other variables considered is new. However, there may be prognostic variables not considered in this study, which in the future could be incorporated into the modeling. In any event, what is important for determining an historical control benchmark is not that one has controlled for all important prognostic variables, but that it is unlikely that (1) there will be a large effect of unmeasured prognostic variables when the known prognostic variables are accounted for, or (2) levels of the unmeasured variables in future trials will be different than in the historical trials.

The choice of the time points of 1 year for OS rates and 6 months for PFS rates were somewhat arbitrary. We wanted to choose a time

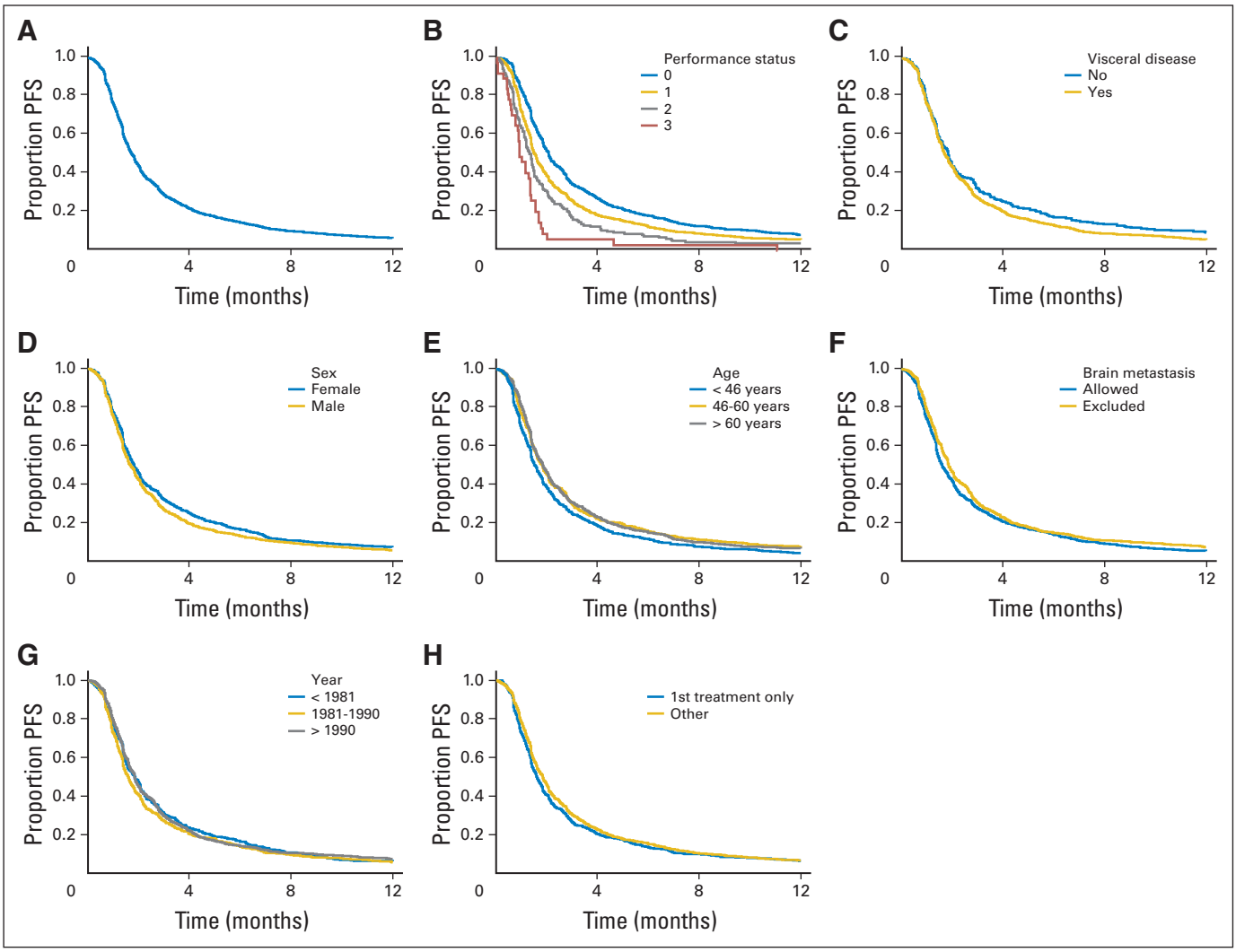

Fig 2. Progression-free survival (PFS) (A) categorized by (B) performance status, (C) presence/absence of visceral disease, (D) sex, (E) age, (F) whether the trial excluded/allowed patients with brain metastases, (G) year trial closed, and (H) whether the trial excluded/allowed previous treatment.

that was long enough so that the rates would be clinically meaningful and relatively low (to increase statistical power), but short enough to avoid a lengthy trial. For OS, one can use an historical control OS curve (as described in Appendix C) to avoid choosing a single time point for comparison of rates. The choice of a 15 percentage point improvement in rates for sample size determination was also arbitrary and can easily be changed. However, this improvement over a 35% 1-year OS rate corresponds to a hazard ratio of 1.51. This is not an insubstantial benefit to be targeting, so one would not want to target a larger difference. On the other hand, because the decision point for declaring a treatment worthy of further study in a phase II trial is approximately one half the targeted difference, only an observed improvement of 7 to 8 percentage points over the historical control would be required for a so-called positive trial that had a targeted 15 percentage point difference. Therefore, targeting a smaller difference may not be wise, because a positive trial may then not be convincing enough evidence to proceed with further development. If one had a series of trials with effective agents in addition to the trials considered in this study, then one could choose time points and targeted differences to maximize the ability of the design to identify effective agents.[23,24]

The choice of whether to use OS or PFS as the primary endpoint is not straightforward. OS has the most unequivocal relevance, although it is not necessary to show clinically relevant benefit in a phase II trial (whose raison d'être is to show sufficient activity to begin a phase III trial). In addition, 6-month PFS results are obtained 6 months earlier than 1-year OS results, and a trial with a PFS endpoint may require a smaller sample size if one is willing to hypothesize a larger treatment effect (hazard ratio) for PFS than OS. Conversely, the use of OS does allow one to compare the whole survival curve with a historical curve, which can offer some benefits in a smaller sample size. In addition, the residual trial-arm variation seen with the PFS endpoint may lead to an excessive number of false-positive phase II trials. Therefore, we recommend OS as the primary endpoint. Six-month PFS rates could be used for an early assessment of the agent in the trial, after which the assessment of mature OS data would validate the decision.

Phase II trials with objective response endpoints frequently use a two-stage design in which the trial is stopped after the first stage if a minimal number of responses is not seen.[25] In principle, these two-stage designs can be applied to the binary endpoints of alive/dead at 1

Table 2. Multivariate Analysis of PFS and Comparison With Univariate Analyses

Variable	No. of Patients*	PFS Distribution			6-Month PFS Rates		
		Univariate†	Multivariate		Univariate‡	Multivariate	
		HR	Adjusted HR§	*P*‖	OR	Adjusted OR¶	*P*#
Performance status							
0	636	1.00**	1.00	< .0001	1.00††	1.00	.0021
1	530	1.30	1.32		1.47	1.50	
2-3	109	1.85	1.83		3.08	2.99	
Visceral disease							
No	277	1.00	1.00	NS	1.00	1.00	NS
Yes	998	1.15	1.11		1.31	1.24	
Sex							
Female	495	1.00	1.00	.026	1.00	1.00	NS
Male	780	1.11	1.14		1.36	1.36	
Age (continuous)	1,275	0.88‡‡	0.86‡‡	.0006	0.78‡‡	0.77‡‡	.043
Brain metastases							
Excluded	703	1.00	1.00	NS	1.00	1.00	
Allowed	572	1.16	1.07		1.18	1.05	NS

Abbreviations: PFS, progression-free survival; HR, hazard ratio; OR, odds ratio; NS, not significant.
*Sample sizes for PFS distribution comparisons; sample sizes for 6-month PFS rate comparisons are slightly smaller.
†Analyses restricted to 1,275 individuals who have data available for all the variables listed.
‡Analyses restricted to 1,273 individuals who have data available for all the variables listed and whose data was not censored before the 6-month time point.
§Adjusted HR is adjusted for the other variables listed.
‖*P* value is testing the association of the variable and PFS in a multivariate analysis that controls for the other variables listed.
¶Adjusted OR is adjusted for the other variables listed.
#*P* value is testing the association of the variable and the 6-month PFS rate in a multivariate analysis that controls for the other variables listed.
**First listed category for categorical variables is always the reference category for HRs.
††First listed category for categorical variables is always the reference category for ORs.
‡‡Reported HR here is for a difference in age of 22 years (the interquartile range of the age distribution), with a value less than 1 suggesting that older patients have better PFS.

year or progression free at 6 months. However, this would require temporarily stopping accrual after the first stage while the survival data mature. One possibility is to have multiple trials ongoing, so that while awaiting first-stage results from one agent, one could be accruing patients to a trial of a second agent.

The historical control benchmarks for OS developed in this article allow one to perform single-arm phase II trials. An alternative strategy is to conduct a randomized phase II screening trial[26] in which patients are randomly assigned to the experimental or control treatment. The advantages of this approach are that there are no questions

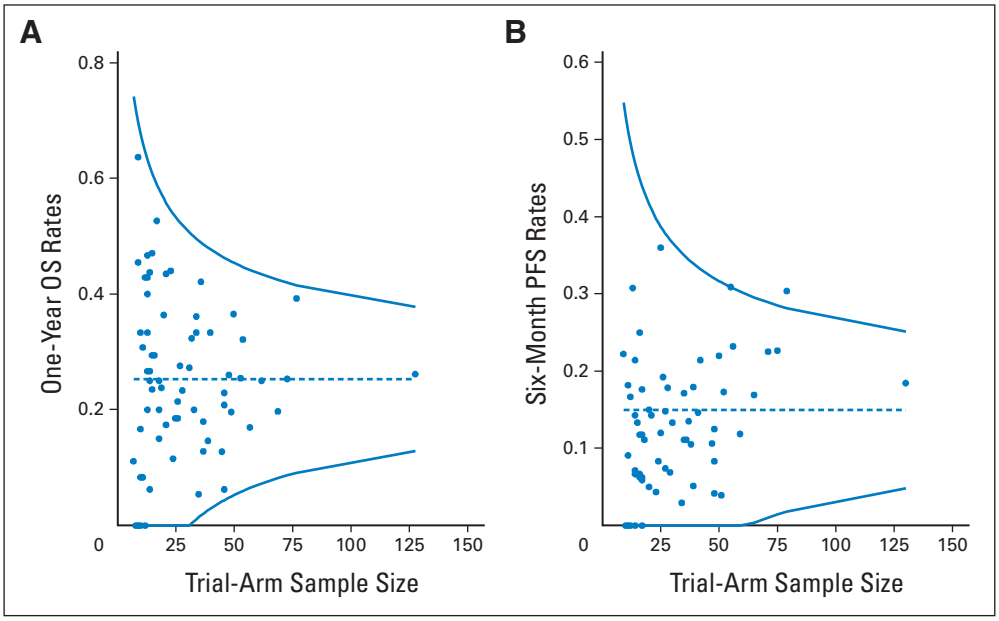

Fig 3. Event rates for each trial arm versus the sample size in the trial arm: (A) overall survival (OS) rates at 1 year, (B) progression-free survival (PFS) rates at 6 months. The solid lines are 95% confidence bounds. The dotted line is the overall 1-year survival rate (25%) or the overall 6-month PFS rate (15%). (A small number of plotted points have been slightly jittered to avoid complete overlap.)

Table 3. Predicted 1-Year Survival Rates and Observed 1-Year Rates (± SE) From Logistic Regression Model

	Trials With Brain Metastases Excluded							
	Men				Women			
	No Visceral Disease		Visceral Disease		No Visceral Disease		Visceral Disease	
Performance Status	Predicted Survival (%)	Observed ± SE (%)	Predicted Survival (%)	Observed ± SE (%)	Predicted Survival (%)	Observed ± SE (%)	Predicted Survival (%)	Observed ± SE (%)
0	49.6	54 ± 7	33.5	36 ± 3	63.8	59 ± 9	47.4	44 ± 5
1	27.6	31 ± 8	16.4	11 ± 3	40.6	44 ± 10	25.9	28 ± 5
2-3	17.4	—*	9.8	18 ± 12	27.4	—	16.2	0 ± 0

	Trials With Brain Metastases Allowed							
	Men				Women			
	No Visceral Disease		Visceral Disease		No Visceral Disease		Visceral Disease	
Performance Status	Predicted Survival (%)	Observed ± SE (%)	Predicted Survival (%)	Observed ± SE (%)	Predicted Survival (%)	Observed ± SE (%)	Predicted Survival (%)	Observed ± SE (%)
0	34.8	28 ± 8	21.5	18 ± 4	48.8	56 ± 9	32.8	34 ± 5
1	17.1	18 ± 7	9.6	13 ± 3	27.0	14 ± 8	15.9	18 ± 5
2-3	10.3	—	5.5	4 ± 3	17.0	—	9.5	18 ± 8

NOTE. SEs (standard errors) of observed rates are calculated using the binomial distribution. For comparison purposes, the 1-year survival rates that formed the basis of the American Joint Committee on Cancer staging system[4] were 59%, 57%, and 41% for stage M1a (no visceral disease), M1b (lung metastases), and M1c (other visceral metastases), respectively.

*Cells with fewer than 10 observations do not have observed rates shown.

concerning the adequacy of an historical control experience, and PFS can safely be used as the endpoint (in a blinded trial). The major disadvantage is its larger required sample size. For example, to detect a hazard ratio of 1.51 (which corresponds to 1-year OS rates of 35% *v* 50%), 220 patients would need to be accrued (over 1 year, with 1 year of additional follow-up) to have the same type 1 and 2 errors as an historical control comparison with 63 patients (or 74 patients if only the 1-year survival rates are used and not the complete survival distributions). It is possible to reduce the size of the randomized screening trial by relaxing the type 1 and 2 errors. Another disadvantage of the randomized screening design is that patients may be reluctant about having their randomized treatment be an older control therapy. Whether a randomized screening design or an historically controlled phase II trial is used, a trial demonstrating promising activity will need to be followed by a definitive randomized phase III trial.

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:** Vernon K. Sondak, Schering-Plough (C), Pfizer (C), Bayer-Onyx (C); Michael B. Atkins, Novartis (C), Bayer (C), Bristol-Myers Squibb Co (C); John M. Kirkwood, Eleos, Inc (C), Intrexon, Inc (C), Antigenics, Inc (C) **Stock Ownership:** None **Honoraria:** Michael B. Atkins, Novartis; John M. Kirkwood, Pharmadura CME, Stategic Edge **Research Funding:** John M. Kirkwood, Pfizer, Bristol-Myers Squibb Co, AstraZeneca **Expert Testimony:** None **Other Remuneration:** None

AUTHOR CONTRIBUTIONS

Conception and design: Edward L. Korn, Ping-Yu Liu, Vernon K. Sondak, Michael B. Atkins, Elizabeth A. Eisenhauer, Scott Saxman, John M. Kirkwood

Provision of study materials or patients: Vernon K. Sondak, Elizabeth A. Eisenhauer, Svetomir N. Markovic, John M. Kirkwood

Collection and assembly of data: Edward L. Korn, Ping-Yu Liu, Sandra J. Lee, Judith-Anne W. Chapman, Donna Niedzwiecki, Vera J. Suman, James Moon, Vernon K. Sondak, Elizabeth A. Eisenhauer, Wendy Parulekar, Svetomir N. Markovic, Scott Saxman, John M. Kirkwood

Data analysis and interpretation: Edward L. Korn, Sandra J. Lee, Judith-Anne W. Chapman, Vera J. Suman, Vernon K. Sondak, John M. Kirkwood

Manuscript writing: Edward L. Korn, Judith-Anne W. Chapman, Vera J. Suman, Vernon K. Sondak, Michael B. Atkins, Svetomir N. Markovic, John M. Kirkwood

Final approval of manuscript: Edward L. Korn, Ping-Yu Liu, Sandra J. Lee, Judith-Anne W. Chapman, Donna Niedzwiecki, Vera J. Suman, Vernon K. Sondak, Michael B. Atkins, Elizabeth A. Eisenhauer, Wendy Parulekar, Svetomir N. Markovic, Scott Saxman, John M. Kirkwood

REFERENCES

1. Korn EL, Arbuck SG, Pluda JM, et al: Clinical trial designs for cytostatic agents: Are new approaches needed? J Clin Oncol 19:265-272, 2001

2. Taylor JMG, Braun TM, Li Z: Comparing an experimental agent to a standard agent: Relative merits of a one-am or randomized two-arm Phase II design. Clin Trials 3:335-348, 2006

3. Margolin KA, Liu PY, Flaherty LE, et al: Phase II study of carmustine, dacarbazine, cisplatin, and tamoxifen in advanced melanoma: A Southwest Oncology Group study. J Clin Oncol 16:664-669, 1998

4. Balch CM, Buzaid AC, Soong S-J, et al: Final version of the American Joint Committee on Cancer staging system for cutaneous melanoma. J Clin Oncol 19:3635-3648, 2001

5. Panageas K, Ben-Porat L, Dickler MN, et al: When you look matters: The effect of assessment schedule on progression-free survival. J Natl Cancer Inst 99:428-432, 2007

6. Einhorn LH, Burgess MA, Vallejos C, et al: Prognostic correlations and response to treatment in advanced metastatic malignant melanoma. Cancer Res 34:1995-2004, 1974

7. Amer MH, Al-Sarraf M, Vaitkevicius VK: Clinical presentation, natural history and prognostic factors in advanced malignant melanoma. Surg Gynecol Obstet 149:687-692, 1979

8. Rampen F: Malignant melanoma: Sex differences in survival after evidence of distant metastases. Br J Cancer 42:52-57, 1980

9. Presant CA, Bartolucci AA: Prognostic factors in malignant melanoma: The Southeastern Cancer Study Group experience. Cancer 49:2192-2196, 1982

10. Balch CM, Soong S, Murad TM, et al: A multifactorial analysis of melanoma: IV. Prognostic factors in 200 melanoma patients with distant metastases. J Clin Oncol 1:126-134, 1983

11. Heimdal K, Hannisdal E, Gundersen S: Regression analyses of prognostic factors in metastatic malignant melanoma. Eur J Cancer Clin Oncol 25:1219-1223, 1989

12. Sirott MN, Bajorin DF, Wong GYC, et al: Prognostic factors in patients with metastatic malignant melanoma. Cancer 72:3091-3098, 1993

13. Barth A, Wanek LA, Morton DL: Prognostic factors in 1,521 melanoma patients with distant metastases. J Am Coll Surg 181:193-201, 1995

14. Brand CU, Ellwanger U, Stroebel W, et al: Prolonged survival of 2 years or longer for patients with disseminated melanoma. Cancer 79:2345-2353, 1997

15. Falkson CI, Falkson HC: Prognostic factors in metastatic malignant melanoma. Oncology 55:59-64, 1998

16. Keilholz U, Conradt C, Legha SS, et al: Results of interleukin-2-based treatment in advanced melanoma: A case record-based analysis of 631 patients. J Clin Oncol 16:2921-2929, 1998

17. Eton O, Legha SS, Moon TE, et al: Prognostic factors for survival of patients treated systemically for disseminated melanoma. J Clin Oncol 16:1103-1111, 1998

18. Ryan L, Kramar A, Borden E: Prognostic factors in metastatic melanoma. Cancer 71:2995-3005, 1993

19. Manola J, Atkins M, Ibrahim J, et al: Prognostic factors in metastatic melanoma: A pooled analysis of Eastern Cooperative Oncology Group trials. J Clin Oncol 18:3782-3793, 2000

20. Unger JM, Flaherty LE, Liu PY, et al: Gender and other survival predictors in patients with metastatic melanoma on Southwest Oncology Group trials. Cancer 91:1148-1155, 2001

21. Balch CM, Soong S, Gershenwald JE, et al: Prognostic factors analysis of 17,600 melanoma patients: Validation of the American Joint Committee on Cancer melanoma staging system. J Clin Oncol 19:3622-3634, 2001

22. Deichmann M, Benner A, Bock M, et al: S100-beta, melanoma-inhibiting activity, and lactate dehydrogenase discriminate progressive from non-progressive American Joint Committe on Cancer stage IV melanoma. J Clin Oncol 17:1891-1896, 1999

23. Van Glabbeke M, Verwij J, Judson I, et al: Progression-free rate as the principal end-point for phase II trials in soft-tissue sarcomas. Eur J Cancer 38:543-549, 2002

24. Francart J, Legrand C, Sylvester R, et al: Progression-free survival rate as primary endpoint for phase II cancer clinical trials: Application to mesothelioma—The EORTC Lung Cancer Group. J Clin Oncol 24:3007-3012, 2006

25. Simon R: Optimal two-stage designs for phase II clinical trials. Control Clin Trials 10:1-10, 1989

26. Rubinstein LV, Korn EL, Freidlin B, et al: Design issues of randomized phase II trials and a proposal for phase II screening trials. J Clin Oncol 23:7199-7206, 2005

27. Kish JA, Kopecky K, Samson MK, et al: Evaluation of fludarabine phosphate in malignant melanoma. Invest New Drugs 9:105-108, 1991

28. Amato DA, Bruckner H, Guerry D, et al: Phase II evaluation of dibromodulcitol and actinomycin D, hydroxyurea, and cyclophosphamide in previously untreated patients with malignant melanoma. Invest New Drugs 5:293-297, 1987

29. Breslow NE: Analysis of survival data under the proportional hazards model. Int Stat Rev 43:45-58, 1975

Acknowledgment

We thank Lesley Seymour, MD, PhD, and Wendy Walsh, BSc, for supplying NCIC Clinical Trials Group data and Michael Smylie, MD, for providing developmental input to the meta-analysis.

Appendix

The Appendix is included in the full-text version of this article, available online at www.jco.org. It is not included in the PDF version (via Adobe® Reader®).