Published in final edited form as:
Cancer. 2012 November 1; 118(21): 5358–5365. doi:10.1002/cncr.27528.

Impact of Disease Progression Date Determination on Progression-free Survival Estimates in Advanced Lung Cancer

Yingwei Qi, M.D. M.S.[1], **Allen Ziegler**, **L. Katie, B.S.**[1], **Shauna L. Hillman, M.S.**[1], **Mary W. Redman, Ph.D.**[2], **Steven E. Schild, M.D.**[3], **David R. Gandara, M.D.**[4], **Alex A. Adjei, M.D. Ph.D.**[5], and **Sumithra J. Mandrekar, Ph.D.**[1]

[1] Mayo Clinic Rochester, Rochester, MN 55905

[2] Fred Hutchinson Cancer Research Center, Seattle, WA 98109

[3] Mayo Clinic Phoenix, Scottsdale, AZ 85259-5404

[4] University of California, Davis, Sacramento, CA; 95817

[5] Roswell Park Cancer Institute, Buffalo, NY 14263

Abstract

PURPOSE—Progression-free survival (PFS) based endpoints are controversial; however in advanced lung cancer, overall survival is largely influenced by the progression status. We thus evaluated the impact of progression date (PD) determination approach on PFS estimates.

METHODS—Individual patient data from 21 trials (14 NCCTG; 7 SWOG) were used. Reported progression date (RPD) was defined as either the scan date or the clinical deterioration date. PD was determined using 4 methods (M): RPD (M1), one day after last progression-free scan (M2), midpoint between last progression-free scan and RPD (M3), and using an interval censoring approach (M4). PFS was estimated using Kaplan-Meier (M1, M2, M3), and maximum likelihood (M4). Simulation studies were performed to understand the impact of the length of time elapsed between the last progression-free scan and the PD on time to progression (TTP) estimates.

RESULTS—PFS estimates using RPD were the highest, with M2 being the most conservative. M3 and M4 were similar due to majority of progressions occurring during treatment (i.e., frequent disease assessments). M3 was less influenced by the length of the assessment schedules (%difference from true TTP <1.5%) compared to M1 (11% to 30%) and M2 (-8% to -29%). The overall study conclusion was unaffected by the method used for randomized trials.

CONCLUSION—The magnitude of difference in the PFS estimates is large enough to alter trial conclusions in advanced lung cancer. Standards for PD determination, use of sensitivity analyses, and randomized trials are critical when designing trials and reporting efficacy using PFS based endpoints.

Introduction

Progression-free survival (PFS) is a common outcome measure for phase II oncology trials for various diseases, including non-small cell and small cell lung cancer (NSCLC, SCLC). Among many appeals of PFS, it is a measure of both efficacy and tumor growth associated with initial therapy, unaffected by any subsequent treatment upon disease progression.[1-4] In

Corresponding author: Sumithra J. Mandrekar, Ph.D. Division of Biomedical Statistics and Informatics, Mayo Clinic Rochester 200 1st ST SW, Rochester, MN 55905 mandrekar.sumithra@mayo.edu.

Presented at the 47th Annual Meeting of the American Society of Clinical Oncology, June 3-7, 2011, Chicago, IL.

a disease with poor prognosis such as advanced NSCLC or SCLC, the true endpoint of overall survival is mostly determined by the progression status of the disease.[1] In addition, with the increase in molecularly targeted therapies, patients typically experience stable disease rather than tumor shrinkage, PFS has thus become an accepted alternate endpoint in assessing treatment efficacy, as it includes a patient who achieves stable disease for an extended period of time as a success, in addition to those who achieve complete or partial response.[1]

PFS-based endpoints in phase II trials remain controversial. There are ongoing discussions of whether PFS can fully capture the potential benefit of a treatment or if a regimen could be approved based on PFS analysis.[5,6] Moreover, the accuracy and validity of PFS as an endpoint is impacted by several 'in-born' factors, including ascertainment bias in an open label trial, timing and scheduling issues such as imbalance in assessment schedules across different arms, treatment holidays, missed assessments, and occurrence of progression in the middle of a long disease evaluation interval.[7-9] With the advent of the Response Evaluation Criteria in Solid Tumors (RECIST),[10,11] ascertainment bias has been reduced or minimized. However, there is no well-accepted standard for addressing tumor assessment timing and scheduling issues, within and across disease types.

Panageas et al.[8] elegantly discussed the problems related to disease assessment scheduling and progression date determination. Tumor assessment is usually scheduled/repeated after a fixed number of treatment cycles, e.g., every other cycle (usually every 6 or 8 weeks depending on cycle length), while patients are on active treatment. Following this paradigm, disease progression can happen anytime between a progression-free scan and the following scan indicating disease progression, a situation called interval censoring in survival data analysis. A common practice is to use the first documented progression date as a surrogate for the true progression date, which likely yields an "overestimation" of the median PFS associated with the study regimen. In extreme cases, this magnitude of difference in PFS estimate may be large enough to alter trial conclusions (i.e., resulting in false positive conclusions). In Panageas et al.'s study, a close relationship between cycle length and progression date declaration was also observed, which highlights the biased inference in practice. This is particularly of concern in single-arm phase II trials where the trial design is based on historical data, instead of concurrent controls.

When the length of the surveillance intervals vary, the validity of comparisons of median PFS across multiple single arm phase II trials within a disease group, or across different treatment arms within one randomized trial, become questionable. In a randomized trial, due to treatment delay from adverse events, say for example, progression date may be reported later in the more toxic arm, even if the true PFS is identical between treatment arms.[9] In addition, with evaluations performed between scheduled assessments (out of concern for lack of efficacy, or signs of clinical progression), patients on the arm with inferior efficacy may have more frequent tumor measurements compared to those receiving treatment of superior efficacy, the so-called evaluation-bias. Consequently, the efficacy of the superior arm is overestimated to some extent.[9] Simulation studies have also provided evidence that variations in the timing of tumor assessment can significantly bias trial conclusions based on PFS analyses and yield misleading results.[7,8]

The goal of our study was to systematically assess the impact of progression date (PD) determination on PFS estimates in phase II trials of advanced non-small cell lung cancer (NSCLC), and extensive stage small cell lung cancer (SCLC). Based on individual patient data from the North Central Cancer Treatment Group (NCCTG) and the Southwest Oncology Group (SWOG) advanced lung cancer trials and simulation studies, we compared

the PFS estimates using different approaches for determining PD, as well as assessed the impact of the length of disease assessment schedules on the time to progression estimates.

Methods

Trial identification

An initial analysis was performed on individual patient data from consecutive North Central Cancer Treatment Group (NCCTG) phase II trials, including 10 NSCLC trials and 7 SCLC trials. To further investigate the consistency and validate the findings on NCCTG trials, we subsequently performed the same analysis on individual patient data from SWOG phase II trials, including 4 NSCLC and 3 SCLC trials. All trials included in this study were activated in the year 2000 or beyond and utilized the RECIST criteria[10] for tumor assessment. To be eligible, a patient must have had a baseline scan and either a follow-up scan, or died before the first scheduled post-baseline scan.

Statistical analysis

In this study, we explored four approaches for PD determination, specifically: Method 1-reported progression date (RPD), which was defined as either the scan date for radiographic progression or the clinical deterioration date (determined by treating physician) for non-radiographic progression; Method 2-one day post last progression-free scan date; Method 3-midpoint between the last progression-free scan date and the RPD; Method 4- multiple imputation method based on nonparametric approach for interval censored data.

For Methods 1-3, PFS was defined as the time from registration to earlier of disease progression or death from any cause, and was estimated using Kaplan-Meier survival analysis. For patients alive and with no documented disease progression, the progression-free survival time was right censored as of the last radiographic scan or clinical assessment date that documented no disease progression. The expectation-maximization (EM) iterative convex minorant algorithm (ICM) approach (EM-ICM), was used for Method 4 to compute non-parametric maximum likelihood estimator (NPMLE).[12] The EM algorithm is an iterative procedure to compute the NPMLE of the distribution function when missing data are present (interval censored data in this case), where the missing data are estimated using the conditional expectation given the observed data. The ICM algorithm is another iterative procedure to compute NPMLE of interval censored data, where only the diagonal elements of Hessian matrix are used. In the hybrid method of EM-ICM, EM and ICM algorithm steps alternate. Specifically, the ICM step searches for the NPMLE in the self-consistent estimate set defined by the EM step until convergence is met. Furthermore, in order to explore the impact of differential progression assessment dates across arms on randomized trials, we applied these 4 methods to the 4 randomized trials. PFS estimates between arms in one trial were compared using log-rank test for progression date assessment Methods 1-3, and generalized log-rank test for Method 4.[13]

Simulation studies to understand the impact of the length of time elapsed between the last non-progression free scan and the reported disease progression date. Since death is an unambiguous endpoint, our simulation study only considered time-to-progression (TTP), and not PFS. The simulations assumed an exponential distribution for the TTP distribution where the rate parameter was calculated based on the median TTP of NCCTG NSCLC data (considered as true TTP) with a uniform censoring distribution (2-60 months). One thousand samples f 50 and 100 observations were generated respectively using various tumor assessment schedules (every 4, 6, 8, 10, and 12 weeks) for the simulated progression times. Methods 1-3 were then applied to analyze these data. Due to the nonparametric nature of Method 4 and the parametric nature of the simulated data, Method 4 was not considered in

the simulation study. The average of the 1000 estimated median TTP derived from each of these methods was compared to the true median TTP. The percent difference was calculated for each scenario as follows: percent difference = [True median TTP - average of estimated median TTP] / true median TTP.

All statistical analyses were conducted in SAS version 9.2,[14] and R version 2.13.0.[15]

Results

Data Description

NCCTG data—The NCCTG data were frozen in November 2009. A total of 660 NSCLC patients (10 trials) and 116 SCLC patients (4 trials) were included. All trials were negative for the protocol defined primary endpoint. Disease assessment schedules followed either every 6 (5 NSCLC, 3 SCLC trials) or 8 weeks (5 NSCLC, 1 SCLC trials) during treatment; and, every 3 months (3 NSCLC, 1 SCLC trials), 3 months for a specified amount of time followed by 6 months thereafter (2 NSCLC, 2 SCLC trials), or every 6 months (5 NSCLC, 1 SCLC trial) during post-treatment. See Table 1 for trial characteristics.

For NSCLC (SCLC), 23% (7%) of patients are alive at the time of this analysis with a median follow-up of 1 year (1.5 years); and 67% (66%) of patients experienced disease progression during treatment. For NSCLC (SCLC) patients whose disease progressed, the median time from the last progression-free scan to reported progression was 1.4 (1.4) months during treatment, and 3.0 (3.0) months during follow-up. See Table 2 for the follow-up and progression summary.

SWOG data—The SWOG data were frozen in December 2010. A total of 297 NSCLC patients (4 trials) and 131 SCLC patients (3 trials) were included. Disease assessment schedules of the SWOG trials are similar to NCCTG trials: every 6 weeks (3 NSCLC, 2 SCLC trials), 8 weeks (1 SCLC trial), or 10 weeks (1 NSCLC) during treatment; and, every 3 months (2 SCLC trials), 3 months for a specified amount of time followed by 6 months thereafter (3 NSCLC), or every 6 months (1 NSCLC, 1 SCLC trial) during post-treatment.

The SWOG data also demonstrated good maturity. For NSCLC (SCLC), compared to the NCCTG data, 7% (4%) of patients are alive at the time of this analysis with a median follow-up of 3.7 years (2.2 years); and 78% (87%) of patients reported progression during treatment. For NSCLC (SCLC) patients whose disease progressed, the median time from the last progression-free scan to progression during and post treatment was 1.3 (1.1) months, and 1.7 (1.3) months respectively. Follow-up and progression status was summarized in Table 2.

Overall Progression-free Survival Estimates

NCCTG data—The median PFS across all NSCLC trials was 4.3 months (Method 1), 1.8 months (Method 2), 3.3 months (Method 3), and 3.45-3.52 months (Method 4, lower and upper limit). The median PFS across all SCLC trials was 2.7 months (Method 1), 0.03 (Method 2), 1.8 months (Method 3), and 1.45-1.45 months (Method 4). For Method 2 compared to Method 1, the percent difference in the PFS estimates among NSCLC (SCLC) was 52% (98%); likewise, for Method 3 compared to Method 1, the percent difference was 29% (46%). See table 3 for the summary of PFS estimates using the 4 methods. For both NSCLC and SCLC, PFS estimates using the reported progression date was the highest (as expected), with Method 2 being the most conservative. Methods 3 and 4 yielded similar results since majority of disease progression occurred during treatment where frequent disease assessments were performed.

SWOG data—The median PFS across all NSCLC trials was 3.1 months (Method 1), 1.5 months (Method 2), 2.2 months (Method 3), and 2.79-2.86 months (Method 4). The median PFS across all SCLC trials was 1.3 months (Method 1), 0.03 (Method 2), 0.7 months (Method 3), and 0.03-0.03 months (Method 4). The percent difference in the PFS estimates among NSCLC (SCLC) was 58% (99%) for Method 2 compared to Method 1; for Method 3 compared to Method 1, the percent difference was 23% (33%). See table 3 for summaries of PFS estimates using the 4 methods. Conclusions from the SWOG data regarding differences in PFS estimates are similar to what was observed in the NCCTG data reported above.

Results of PFS estimates by trial, although not presented, were similar to the results of the overall PFS estimates. The percent difference in the PFS estimates ranged from as low as 25% to as high as 99% among NSCLC trials (NCCTG and SWOG combined), 43%-98% among SCLC trials for Method 2 compared to Method 1; similarly, for Method 3 compared to Method 1, the percent difference in the PFS estimates were between 15%-47% among NSCLC, 19%-50% among SCLC trials.

Outcomes for randomized trials

For randomized trials (3 NCCTG trials and 1 SWOG trial), the 4 methods resulted in the same overall conclusions, and the comparisons across arms in each trial was consistent across the four methods. Refer to Table 4 for detailed summaries.

Simulation results

Eleven percent and seven percent of patients died without progression for NCCTG NSCLC and SCLC respectively; and 24% (16%) of patients died without progression for SWOG NSCLC (SCLC). Thus, death without progression in both NCCTG and SWOG data represented relatively low percentages. Our simulated time-to-progression data, therefore, represents the majority of patients in our trials: patients alive with disease progression (both during treatment and follow-up) and those who progressed at the time of death. The median time-to-progression for NCCTG NSCLC patients was 4.3 months which was used as the true median TTP for the simulation study.

Method 1 (RPD) consistently overestimated the true median TTP across assessment schedules, and Method 2 (where the progression date was one-day post the last progression - free scan) consistently underestimated the true median TTP. Method 3 (midpoint) provided TTP estimates closest to the true median TTP, as reflected by the <1.5% difference in the average median TTP estimates from the true median TTP of 4.3 months. This index also increased with increase in the length of the disease assessment schedule for Method 1 and Method 2. The length of the disease assessment schedule did not impact the results of Method 3. See Table 5 for the details of the simulation results for sample size of 50. The results were similar for trials with a sample size of 100 (data not shown).

Discussion

In this study, we systematically studied the impact of approach used for the declaration of the progression date on PFS estimates in phase II trials of advanced NSCLC, and extensive stage SCLC. Although NCCTG and SWOG trials had different data collection schedules for monitoring disease progression, the results are strikingly consistent. PFS estimates using reported progression date were the highest as a consequence of the length of the assessment interval. Method 2 (one day post the last progression-free scan) was the most conservative. Method 3 (midpoint between the last progression-free scan and reported progression date) and Method 4 (non-parametric interval censoring) yielded similar results, which were in between the estimates using Method 1 and 2, since majority of the disease progression

occurred during treatment when frequent disease assessments were performed (i.e., every 4, 6, 8 or 10 weeks). Analysis of randomized trials revealed that the trial conclusions remained unaffected by the method used to determine progression date. The simulation study also confirmed these findings. Although the true median PFS is unknown in reality, our simulation results provide reasonably convincing evidence that the traditional PFS analysis approach of using the reported progression date may yield misleading results (overestimation), especially in the case of single trials.

Panageas et al. demonstrated the bias associated with the PFS analyses using simulations with progression dates marked by lower limit, midpoint, and upper limit of interval censored data.[8] These three methods utilized are similar to our approach, except in the definition of lower limit, which was defined as the date of the last non-progression scan (and we used one day past the last progression-free scan). Regardless, their study reported that, compared to the pre-specified true median PFS, the PFS estimates using Kaplan-Meier method based on the above three approaches showed a certain pattern: the upper limit based method consistently overestimates the true median PFS and the lower limit consistently underestimates it. Our findings are thus consistent with those of Panageas et al.

Several possible solutions to address the timing and scheduling issue when PFS is used as the primary endpoint have been proposed in the literature. Whether in single arm or randomized trial designs, one simple way to reduce PFS assessment bias introduced by timing and scheduling is to carefully consider the relationship between the evaluation frequencies relative to the median PFS. Panageas et al. recommended the interval censoring approach as the solution for PFS analysis.[8] The need for appropriate interval censoring survival analysis is also evident based on our results. Methods 1-3 examined in our study essentially utilize an imputation approach: right, left and midpoint imputation respectively. It is then fairly easy to apply a well accepted time-to-event approach such as the Kaplan-Meier method to analyze these data. However, based on our results and those of Panageas et al., ignoring the interval censored nature of the PFS data could lead to erroneous conclusions. With advances in the methodology and the availability of easy-touse statistical software packages, the application of interval censoring in survival analysis has become relatively simple.[14] We however note that the interpretation of the interval censoring results and the associated survival curves remains largely uncommon in the oncology literature. In addition, the FDA has recommended sensitivity analyses for evaluating the robustness of PFS as an end point in oncology clinical trials.[16] In a trial in which PFS is the primary end point, Bhattacharya et al. also recommended performing sensitivity analyses to explore the impact of assessment time imbalances, nonradiologically confirmed PFS events, and missing data.[9]

Another approach is to use a binary endpoint for diseases which have a short median PFS such as advanced NSCLC and SCLC (for example, 6-month PFS rate). This was also recommend by the studies of Panageas et al.[8] and Friedlin et al.[7] With careful attention to trial design, a tumor assessment date could coincide with the primary endpoint PFS assessing date, for example, 6 months for 6-month PFS rate. This is especially important for patients who have not progressed before 6 months, as all patients would be evaluated by the fixed time point of 6 months, which increases the homogeneity of results across trials. Other proposals include a placebo-controlled double-blinding design (when feasible) to limit the evaluation time bias between arms in randomized trials.[7] Independent central review of progression-free survival endpoint has also been proposed and implemented.[17] In addition, since percentage change of tumor burden is less subject to the issue of timing and scheduling, combining it with PFS endpoint may be another direction worthy of exploration.[18]

In conclusion, when progression-free survival is used as the primary endpoint for phase II trials in advanced lung cancer, considerations should be taken to minimize the bias caused by timing and scheduling disease progression assessment. Appropriate statistical analysis, including interval censoring approach and sensitivity analyses, should be performed to reduce the potential biases. Interpretation of trial outcomes based on PFS merits caution. Standards for progression date determination, use of sensitivity analyses, and randomized trials are critical when designing trials and reporting efficacy using PFS based endpoints.

Reference

1. Mandrekar SJ, Qi Y, Hillman SL, et al. Endpoints in phase II trials for advanced non-small cell lung cancer. J Thorac Oncol. 2010; 5(1):3–9. [PubMed: 19884856]
2. Buyse M, Squifflet P, Laporte S, et al. Prediction of survival benefits from progression-free survival in patients with advanced non small cell lung cancer: Evidence from a pooled analysis of 2,838 patients randomized in 7 trials. J Clin Oncol. May.26(20 suppl):2008. abstr 8019.
3. Lara PN Jr, Redman MW, Kelly K, et al. Disease control rate at 8 weeks predicts clinical benefit in advanced non-small-cell lung cancer: results from Southwest Oncology Group randomized trials. J Clin Oncol. 2008; 26:463–467. [PubMed: 18202421]
4. Foster NR, Qi Y, Shi Q, et al. Tumor response and progression-free survival as potential surrogate endpoints for overall survival in extensive stage small-cell lung cancer: findings on the basis of North Central Cancer Treatment Group trials. Cancer. 2011; 117(6):1262–1271. [PubMed: 20960500]
5. Tuma R. Progression-free survival remains debatable endpoint in cancer trials. J Natl Cancer Inst. 2009; 101(21):1439–1441. [PubMed: 19828873]
6. Broglio KR, Berry DA. Detecting an overall survival benefit that is derived from progression-free survival. J Natl Cancer Inst. 2009; 101(23):1642–1649. [PubMed: 19903805]
7. Freidlin B, Korn EL, Hunsberger S, et al. Proposal for the use of progression-free survival in unblinded randomized trials. J Clin Oncol. 2007; 25:2122–2126. [PubMed: 17513819]
8. Panageas KS, Ben-Porat L, Dickler MN, et al. When you look matters: The effect of assessment schedule on progression-free survival. J Natl Cancer Inst. 2007; 99:428–432. [PubMed: 17374832]
9. Bhattacharya S, Fyfe G, Gray RJ, et al. Role of sensitivity analyses in assessing progression-free survival in late-stage oncology trials. J Clin Oncol. 2009; 27(35):5958–5964. [PubMed: 19826121]
10. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. J Natl Cancer Inst. 2000; 92:205–216. [PubMed: 10655437]
11. Eisenhauer EA, Therasse P, Bogaert J, et al. New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1). Eur J Cancer. 2009; 45(2):228–247. [PubMed: 19097774]
12. Wellner JA, Zhan Y. A hybrid algorithm for computation of the nonparametric maximum likelihood estimator from censored data. JASA. 1997; 92:945–959.
13. Zhao Q, Sun J. Generalized log-rank Test for mixed interval-censored failure time data. Stat Med. 2004; 23:1621–1629. [PubMed: 15122741]
14. So, Y.; Johnston, G.; Kim, SH. Analyzing interval-censored survival Data with SAS® Software. 2010. SAS Institute & UNC-Chapel Hill. Macros downloaded from SAS Institute Inc.; (http://support.sas.com/kb/24980). Downloaded on January 28, 2011
15. R interval censoring package—Icens—downloaded from http://cran.r-project.org/web/packages/Icens/index.html. Accessed on May 10, 2011
16. Food and Drug Administration. Guidance for industry: clinical trial endpoints for the approval of cancer Drugs and biologics. 2007. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM071590.pdf
17. Amit O, Bushnell W, Dodd L, Roach N, et al. Blinded independent central review of the progression-free survival endpoint. Oncologist. 2010; 15(5):492–495. [PubMed: 20489186]
18. Fridlyand J, Kaiser LD, Fyfe G. Analysis of tumor burden versus progression-free survival for Phase II decision making. Contemp Clin Trials. 2011; 32(3):446–452. [PubMed: 21266203]

Table 1

Trial Characteristics

Trial	N	Year Closed	Study Agents	Disease Assessment Schedule	
				During Treatment	Post Treatment
NSCLC					
N0022	46	2002	oral vinorelbine	every 8 weeks	every 3 months
N0026	132	2003	gemcitabine, pemetrexed disodium	every 6 weeks	every 6 months
N0222	47	2006	paclitaxel, carboplatin, gefitinib	every 8 weeks	every 3 months
N0323	52	2006	temsirolimus	every 8 weeks	every 6 months
N0326	25	2006	sorafenib	every 8 weeks	every 6 months
N0426	48	2007	pemetrexed disodium, bevacizumab	every 6 weeks	every 6 months
N0528	93	2008	cedirinib, gemcitabine, carboplatin	every 6 weeks	every 6 months
N0626	110	2010	pemetrexed, sorafenib	every 6 weeks	every 3, then 6 months
N0821	65	2010	pemextrexed, carboplatin, bevacizumab	every 6 weeks	every 3, then 6 months
N9921	42	2005	carboplatin, paclitaxel	every 8 weeks	every 3 months
S0027	57	2003	vinorelbine, docetaxel	every 10 weeks	every 3, then 6 months
S0339	53	2004	gemcitabine, carboplatin, bortezomib	every 6 weeks	every 6 months
S0341	55	2005	erlotinib	every 6 weeks	every 3, then 6 months
S0342	132	2005	carboplatin, paclitaxel, cetuximab	every 6 weeks	every 3, then 6 months
SCLC					
N0027	26	2003	topotecan, carboplatin, G-CSF	every 6 weeks	every 3, then 6 months
N0124	21	2004	imatinib mesylate	every 8 weeks	every 3, then 6 months
N0423	46	2008	pemetrexed disodium, carboplatin	every 6 weeks	every 6 months
N0621	23	2008	saracatinib	every 6 weeks	every 3 months
S0119	31	2003	gemcitabine, irinotecan	every 6 weeks	every 6 months
S0327	42	2004	bortezomib	every 6 weeks	every 3 months
S0435	58	2007	sorafenib	every 8 weeks	every 3 months

Table 2

Follow-up and Progression Status Summary

	NCCTG		SWOG	
	NSCLC (n=660)	**SCLC (n=116)**	**NSCLC (n=297)**	**SCLC (n=131)**
Alive (%)	22.6%	6.9%	7.1%	3.8%
Follow-up Time in months for Alive Patients Median (Range)	1.0 (0.0-6.0)	1.5 (0.1-6.3)	3.7 (0.8-5.2)	2.2 (0.2-2.8)
Progression Occurrences*, N (%)				
On Treatment	346 (66.5)	69 (66.3)	175 (78.1)	93 (86.9)
Long Term Follow-up	174 (33.5)	35 (33.7)	49 (21.9)	14 (13.1)
Time in Months from Last Progression-free Scan to Disease Progression* Median (Range)				
During Treatment	1.4 (0.1-4.1)	1.4 (0.3-3.1)	1.3 (0.03-13.1)	1.1 (0.0-2.1)
During Follow-up	3.0 (0.0-48.3)	3.0 (0.1-14.1)	1.7 (0.2-24.3)	1.3 (0.4-5.3)

*Reported for patients who progressed; NCCTG NSCLC (n=520), NCCTG SCLC (n=104), SWOG NSCLC (n=123), SWOG SCLC (n=50).

Table 3

Overall PFS Estimates by Method

Progression Date Determination Method	NCCTG		SWOG	
	NSCLC (n=660)	SCLC (n=116)	NSCLC (n=297)	SCLC (n=131)
Method 1				
Median PFS (95% CI)*	4.3 (3.8-4.7)	2.7 (1.6-4.2)	3.1 (2.5-4.1)	1.3 (1.2-1.6)
6-month estimate (%)	36.5 (32.9-40.4%)	21.5 (15.1-30.6%)	31.0 (26.1-36.7%)	10.7 (6.5-17.5%)
1-year estimate (%)	12.3 (9.9-15.3%)	3.8 (1.4-9.8%)	12.5 (9.2-16.8%)	3.8 (1.6-9.0%)
Method 2				
Median PFS (95% CI)*	3.3 (2.7-3.6)	1.8 (0.8-3.3)	2.2 (1.6-2.8)	0.7 (0.6-0.8)
6-month estimate (%)	28.1 (24.7-31.8%)	13.6 (8.6-21.7%)	23.6 (19.1-29.1%)	6.7 (3.4-13.1%)
1-year estimate (%)	8.9 (6.8-11.7%)	1.0 (0.1-6.9%)	8.6 (5.8-12.6%)	3.4 (1.3-8.8%)
Method 3				
Median PFS (95% CI)*	1.8 (1.6-2.0)	0.03 (0.03-1.4)	1.5 (1.2-1.7)	0.03 (n/a)
6-month estimate (%)	17.1 (14.4-20.3%)	4.7 (2.0-11.0%)	17.2 (13.4-22.0%)	5.3 (2.6-11.0%)
1-year estimate (%)	7.9 (5.9-10.6%)	1.3 (0.2-8.1%)	6.4 (4.1-9.9%)	3.1 (1.2-8.0%)
Method 4				
Lower, Upper*	3.45, 3.52	1.45, 1.45	2.79, 2.86	0.03, 0.03

*Estimates are in months.

Table 4

PFS Comparisons for Randomized Trials

Trial	Progression Date Determination Method			
	Method 1 (median PFS)*	Method 2 (median PFS)*	Method 3 (median PFS)*	Method 4 (lower, upper)*
Trial 1				
Arm A	4.6	2.9	3.7	3.94, 4.24
Arm B	3.6	1.6	2.6	2.46, 3.02
Arm C	4.2	1.6	3.4	4.14, 4.20
P value †	0.59	0.24	0.32	0.67
Trial 2				
Arm A	5.7	3.1	4.6	4.47, 4.60
Arm B	4.9	1.6	2.9	4.63, 4.99
P value †	0.25	0.19	0.18	0.25
Trial 3				
Arm A	3.1	1.3	2.1	2.60, 2.60
Arm B	3.8	1.5	3.4	3.48, 4.11
P value †	0.38	0.40	0.50	0.67
Trial 4				
Arm A	6.7	4.0	5.5	4.07, 4.11
Arm B	5.2	3.3	4.2	3.25, 3.84
P value ‡	0.43	0.35	0.39	0.77

† P value for arm comparison: Log-rank test

‡ P value for arm comparison: Generalized Log-rank test

* Estimates in months.

Table 5

Simulation Results of the TTP Estimates (N=50, 1000 simulation runs)

Disease Assessment Schedule	Average Time-to-Progression Estimate in Months (% Difference from the True Median TTP)		
	Method 1	Method 2	Method 3
4 weeks	5.1 (10.9%)	4.3 (-8.3%)	4.6 (1.1%)
6 weeks	5.3 (15.5)	4.0 (-12.8%)	4.6 (0.3%)
8 weeks	5.6 (21.1%)	3.8 (-18.0%)	4.7 (1.2%)
10 weeks	5.8 (26.3%)	3.5 (-23.6%)	4.6 (0.3%)
12 weeks	6.0 (30.2%)	3.3 (-28.5%)	4.6 (-0.2%)