Association of second line chemotherapy using docetaxel and TFD725 with prolonged survival

ABSTRACT

Long-term survival for patients with non-small cell lung cancer (NSCLC) is poor, and second-line therapies are often needed as a result of disease progression following first-line therapies. While docetaxel has proven to be an adequate second-line regimen, the tyrosine kinase inhibitor TFD725 has been found in Phase I and IIa trials to be safe and efficacious. This study examined whether second-line chemotherapy using docetaxel plus TFD725 was associated with improved survival over the use of docetaxel alone in patients initially diagnosed with advanced stage NSCLC.

In a double blind phase IIb clinical trial, a total of 188 men (n = 104) and women (n = 84) aged 46 to 71 years were randomized to receive solely docetaxel (n = 90) or docetaxel plus TFD725 (n = 98). Participants had stage IIIb or higher disease and had progressed on therapy. Randomization was stratified by stage at diagnosis and site. The primary endpoint was all-cause mortality. The association between mortality and treatment arm was assessed using an unadjusted Cox proportional hazards regression model. A subgroup analysis stratified by stage of disease at diagnosis was also performed.

The risk of death for patients receiving TFD725 plus docetaxel was found to be 25% lower compared to patients receiving docetaxel alone (HR: 0.75; 95% CI: 0.54, 1.04; p = 0.08). There was no decrease in risk of death in participants with stage IV disease or malignant pleural effusion (n: 118; HR: 0.99, 95% CI: 0.67, 1.46; p = 0.95). However, among those with stage IIIb disease the risk of death for patients in the TFD725 plus docetaxel arm was 47% less than that of patients in the docetaxel alone arm (n: 70; HR: 0.53; 95% CI: 0.29, 0.99; p = 0.05). Of participants in the docetaxel alone group, 32.2% had an abnormal alkaline phosphatase level at baseline, compared with only 19.4% in the TFD725 plus docetaxel group (p = 0.04).

In conclusion, there was no difference in overall survival between the treatment arms. However, among patients with stage IIIb disease, there was improved survival in participants receiving TFD725 plus docetaxel in comparison to docetaxel alone. Despite randomization, though, participants in the TFD725 plus docetaxel group may have had less severe disease as indicated by the lower proportion with abnormal alkaline phosphatase levels.

Comment [A1]: what was length of follow-up

Comment [A2]: was it prespecified? was it the only such subgroup analysis?

Comment [A3]: Could give some descriptive statistics about median survival and/or 1 year survival

Comment [A4]: It is very important to know whether this question existed before looking at the data, or only after. Did you adjust for multiple comparisons?

Comment [A5]: If this is important to mention, it should probably be mentioned before the results or at least before the subgroup analyses.

Comment [A6]: your wording is too strong for an exploratory subgroup analysis whose p value is not adusted for multiple comparisons

Comment [A7]: also worse LDH, worse stage, worse ECOG—all these things go together

In real life we would likely have prespecified a secondary analysis adjusting for known important baseline variables

BACKGROUND

Lung cancer is the leading cause of cancer-related death in both men and women. Treatment options are determined by type (small cell vs. non small cell) and stage of cancer at diagnosis.[1-3] As of 2005, the 5-year relative survival for all stages combined is only 15%.

Although long-term survival for metastatic non-small cell lung cancer (NSCLC) remains poor, chemotherapy provides modest survival improvement and reductions in symptoms. Past studies have shown that cisplatin-based chemotherapy regimens improve patient survival.[2,4] Despite improvements to first-line treatments, NSCLC will often continue to progress resulting in a need for additional therapy. Positive results in phase III clinical trials have established docetaxel as a standard second-line agent.[2,5-6]

Unfortunately, due to the aggressive nature of the disease, second-line regimens aimed at treating NSCLC remain only modestly successful at best.[3,5] Researchers have recently turned to compounds that selectively target molecular pathways relevant to cancer development and progression.[7,8] The most promising such agents in NSCLC have been epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (EGFR-TKIs).[3] Tyrosine kinase pathways are triggered by some tumor growth factors known to cause angiogenesis. Therefore, blocking these kinases could impair a tumor's ability to grow. Several pre-clinical studies have examined the effectiveness of EGFR-TKIs, including gefitinib and erlotinib, in

enhancing the anti-tumor activity of concurrent first-line chemotherapy.[3,8] Results from these studies are poor, and EGFR-TKIs seems to do little to improve the effectiveness of first-line chemotherapy alone.

Nonetheless, preliminary studies indicate that EGFR-TKIs may be effective as a maintenance therapy following initial chemotherapy.[8] Further research is needed to understand the potential benefits of EGFR-TKIs as components of second-line chemotherapy for NSCLC. TFD725 is an experimental molecule that has exhibited activity against several receptor tyrosine kinases in vitro and in animal experiments. Phase I and IIa clinical trials have also provided encouraging initial safety and efficacy data. Additional research into the effectiveness of TFD725 in treating NSCLC is warranted.

QUESTIONS OF INTEREST

The specific aim of this study was to assess whether second line chemotherapy using docetaxel plus TFD725 was associated with improved survival over the use of docetaxel alone in patients initially diagnosed with late stage NSCLC.

Since the stage of cancer at diagnosis is the current standard by which treatment of NSCLC is assigned[1-3], this study also analyzed whether any association between survival and second line treatment regimen differed by disease stage at initial diagnosis.

Comment [A8]: good to motivate in this way. But still need to consider whether primary, secondary, exploratory

SOURCE OF THE DATA

The study consisted of a randomized double-blind phase IIb clinical trial conducted at multiple clinical sites in the United States and Europe starting in 2003. Subjects were NSCLC patients with stage IIIb or IV disease at diagnosis that had progressed on first-line chemotherapy. Exclusion criteria included the following: 1) first line therapy that included docetaxel; 2) ECOG score of 3 or higher; 3) age older than 80 years at randomization; 4) unwillingness to use adequate contraception during trial. Subjects were randomized to receive docetaxel (50mg/m2 every three weeks) plus TFD725 (50mg/day) or docetaxel alone (dose 75 mg/m2 every three weeks). Randomization was stratified by stage and site location. Baseline data were collected from patients at enrollment. Demographic data included site location, age, and gender. Data collected on disease state included stage at diagnosis, response to first-line therapy, and whether two proxy measures of disease severity, lactic acid dehydrogenase (LDH) and alkaline phosphatase, were abnormal at baseline. Lastly, ECOG performance status scores were collected. This was a complete case analysis, as there was no missing data in this study.

Comment [A9]: 1:1

Comment [A10]: except for follow-up which was sometimes censored

As this was a randomized trial, we expected potential confounding to be minimized. However, the classification of disease stage was made prior to randomization and a subject's clinical status could have changed substantially by the time of study enrollment. Randomization by stage at diagnosis may have failed to account for this change. The time from initial diagnosis to randomization, response to first-line therapy, level of LDH and level of alkaline phosphatase all relate to degree of disease progression and, by a priori consideration, these could have been unbalanced between the treatment arms. Regarding ECOG performance scores, past studies found individuals with a score of 2 to have improved survival over those with a performance score of 0 or 1.[2-3,9]

Comment [A11]: but is there any reason it would have differed across the randomized groups in the way it changed?

Comment [A12]: Most unusual, don't you think?

Unfortunately, some measurements were not collected in this study that might also confound the relationship between survival and treatment arm, although randomization should have prevented this. Survival rates from NSCLC are less among certain racial and ethnic groups, particularly African-Americans.[1] Data on race would have provided useful information regarding the risk of death in this study. Smoking history, current smoking status, disease histology, and past treatments would have aided in assessing the comparability of treatment arms at baseline, as these have been found to impact EGFR-TKI effectiveness in past analyses.[2-3,8,10] Measurements of drug tolerance and side effects would also have

Comment [A13]: This should be in the discussion, not in the methods

been reasonable to collect; differences in treatment adherence between arms would not be accounted for by randomization and could impact survival.

Comment [A14]: And because it would not be accounted for by randomization, we would NOT want to do an analysis based on patient adherence (except in a very limited fashion to try to gain insight into mechanism of action). It has many times been found that placebo patients who take their medication survive better than placebo patients who do not. Impending death is likely the cause of lack of compliance, rather than vice versa

STATISTICAL METHODS

Descriptive statistics were used to assess the comparability of baseline characteristics and treatment groups at the time of randomization. Differences between treatment arms and dichotomous or categorical measurements (site location, gender, disease stage, response to first line therapy, normality of LDH or alkaline phosphatase levels, and ECOG status) were determined through chi-square tests for independence, without using a Yates' continuity correction. To examine if the continuous measurements of age and the time from initial diagnosis to randomization (in months) varied between treatment arms, two sample t-tests for unequal variance were used to compare differences in the means.

Comment [A15]: hypothesis testing is irrelevant here.

Survival estimates and 95% confidence intervals were calculated among the treatment arms using the Kaplan-Meier method with Greenwood's formula to approximate the variance. Survival curves were compared between the two arms using Kaplan-Meier plots. Differences in survival distributions between the treatment groups were evaluated using the logrank test. A Kaplan-Meier plot was also used to compare survival estimates between the treatment arms by disease stage.

Comment [A16]: of the estimates

In order to assess how the risk of death differed by treatment arm, an unadjusted Cox proportional hazards regression model with robust variance estimates was used to estimate hazard ratios and 95% confidence intervals. A subgroup analysis was used to examine any potential differences in the risk of death between treatment arms by disease stage; hazard ratios and 95% confidence intervals were estimated among disease stage groups from unadjusted Cox proportional hazards regression models with robust variance estimates.

Comment [A17]: this is a bit of Stata terminology. More generally we would say "Huber-White sandwich estimator"

All p-values computed in these analyses were two-sided, and all analyses were conducted using either Stata 10 or R version 2.8.1. A p-value at or below 0.05 was considered statistically significant.

RESULTS

The total sample consisted of 188 men and women aged 46 to 71 years, of which 90 (47.9%) were randomized to receive solely docetaxel and 98 (52.1%) were randomized to receive docetaxel plus TFD725. The median follow-up time was 18.1 months (95% CI: 16.4, 18.9) in the group receiving solely docetaxel and 17.9 months (95% CI: 16.6, 18.1) in the group receiving TFD725 plus docetaxel. During the time of the study, 72 deaths were observed in the docetaxel alone group and 68 deaths were observed in the TFD725 plus docetaxel group. No patients were excluded due to missing data.

Comment [A18]: very nicely summarized

Baseline characteristics suggested that the treatment arms were mostly comparable at the time of randomization (Table 1). The distribution of age was similar across groups, with a mean age of approximately 60 (SD: 5) for both arms. The majority of patients in both the TFD725 plus docetaxel and docetaxel alone arms were from North America (82.7% and 81.1%), with each group consisting of a slightly higher proportion of men (58.2% and 52.2%). Most patients were initially diagnosed as having stage IV disease or malignant pleural effusions: 60.2% of the TFD725 plus docetaxel group and 65.6% of the docetaxel alone group. Although all patients had stage IIIb or higher disease, at least 94% of patients in each treatment arm had a high ECOG performance status (grade 0 or 1). This measurement along with stage at diagnosis, patient demographics, and response to first line treatments were not found to differ between the treatment arms, which was evidenced through p-values greater than 0.05 when testing for differences between the arms.

Comment [A19]: I think there is atrend toward worse disease in the placebo group

Time from initial diagnosis to randomization was between 6 and 13 months for most patients (at least 60% in each arm), but the range was from 3 to 31 months. At the time of randomization most patients in

the two arms had normal LDH (90.8% and 82.2%) and alkaline phosphatase levels (80.6% and 67.8%). These proportions, though, were somewhat unbalanced between the treatment arms. Tests for whether the distribution of these measurements varied by treatment arm provided evidence that abnormal alkaline phosphatase levels were not equally distributed (p = 0.04). There was not strong evidence that abnormal LDH levels were unbalanced between the treatment arms (p = 0.08). More patients with stage IV disease or malignant pleural effusions had abnormal LDH and alkaline phosphatase levels at baseline compared to those with stage IIIb disease (80.0% vs. 20.0% and 77.1% vs. 22.9%).

Using Kaplan-Meier estimates, survival probability did not appear to differ between the treatment arms (Figure 1) and was found to be similar between the groups at 6, 12, and 18 months post randomization (Table 2). A logrank test also indicated no difference in survival between the treatment arms (p = 0.08). However, the probability of survival associated with treatment appeared to differ by disease stage (Figure 1). Participants with stage IIIb disease and no malignant pleural effusions treated with TFD725 plus docetaxel appeared to have improved survival over participants with higher stage disease treated with TFD725 plus docetaxel as well as over those with either stage of disease treated with docetaxel alone.

The risk of death for patients in the TFD725 plus docetaxel arm was 25% less (HR: 0.75; 95% CI: 0.54, 1.04) than that in the docetaxel alone arm (Table 3). There was not strong evidence, though, that the risk of death differed by treatment arm (p = 0.08). Among those with stage IIIb disease without malignant pleural effusion, the risk of death for patients in the TFD725 plus docetaxel arm was 47% less than that of patients in the docetaxel alone arm (HR: 0.53; 95% CI: 0.29, 0.99). We found some indication that among those with stage IIIb disease the risk of death differed by treatment arm (p = 0.05). There was not sufficient evidence to show that the risk of death differed by treatment arm among those with later stage disease (HR: 0.99; 95% CI: 0.67, 1.46; p = 0.95).

Comment [A20]: so you did worry about the comparibility

Comment [A21]: hypothesis testing is irrelevant here

Comment [A22]: You should have presented estimates of median survival and/or 1 year survival probabilities. Without those, the interpretation of HR is difficult. We don't know if we care. (Doubling survival from 3 minutes to 6 minutes is rarely of interest)

Comment [A23]: I presume this is the Wald test from PH regression. The logrank test you quote above is the score test. One of them probably would have sufficed

Comment [A24]: explicitly mention that this is not adjusted for multiple comparisons, so this is merely exploratory

DISCUSSION

This phase IIb trial of second line therapy for stage IIIb or higher NSCLC compared mortality in participants receiving either the tyrosine kinase inhibitor TFD725 plus docetaxel or docetaxel alone. We did not find evidence that the risk of death was lower in the TFD725 plus docetaxel group compared to the docetaxel alone group. There was some evidence of improved survival for those with less advanced disease at time of diagnosis. Specifically, those with stage IIIb disease receiving TFD725 plus docetaxel had a lower risk of death compared to those receiving docetaxel alone. This finding suggests a possibly beneficial treatment effect for individuals with less advanced disease. No effect was seen in participants with stage IV disease or malignant pleural effusion.

This study has several limitations. Typically NSCLC is only staged at the time of first diagnosis, which is reasonable for making treatment decisions about second-line therapies but complicates the randomization procedure.[1,2,5] Since disease staging only occurred at the time of diagnosis, it is unclear exactly how advanced participants' disease was at the time of randomization. In particular, the time from initial diagnosis to randomization ranged between 3 and 31 months. A patient with a greater time span between initial diagnosis and randomization was perhaps more likely to experience a worsening of disease compared to a patient with a shorter time span. This aspect of the study design may have contributed to the observation that subjects with elevated markers of disease severity (alkaline phosphatase) were unequally distributed between the treatment groups. Subjects in the TFD725 plus docetaxel group may have had less severe disease at time of randomization. Thus, randomization by stage at diagnosis may have failed to account for differing rates of disease progression and resulted in unbalanced groups. As we did not compensate for this possibility, future studies may benefit from stratifying based on better markers of disease severity at enrollment in order to ensure the comparability of treatment arms.

Further, we lacked data on smoking history, race, and disease histology, all of which have been associated with disease progression or treatment response.[1-3,8,10] Although we would expect our randomization to

Comment [A25]: and luckily there was no suggestion of harm in this group

Comment [A26]: it doesn't really. The randomization process should ensure that comparable patients are in each group, even if we don't know their current metastasis status

Comment [A27]: Nope. Just random chance in a randomized study

Comment [A28]: "Failed" is a strong word. But in real-life we would have done stratified analyses at least as a secondary outcome

Comment [A29]: not a big problem

have eliminated the possibility of confounding, we have no way to confirm that our procedure resulted in equal distribution of these factors. Thus, residual confounding may have prevented us from detecting a true treatment effect. Also, a lower dose of docetaxel was used in the TFD725 plus docetaxel group. If there was a corresponding decrease in the treatment effect of docetaxel, this may have masked some potential benefit of the experimental agent. Finally, we lacked data on treatment tolerability and side effects.

Comment [A30]: or it could be that if we gave the higher dose of DOC with TFD, we would have killed everyone

Advanced NSCLC has a poor prognosis, and patients who progress on first-line platinum based chemotherapy have few options for additional treatment. The current standard second-line chemotherapy is docetaxel alone. While tyrosine kinase inhibitors have shown promise in treating NSCLC, in this study the use of TFD725 plus docetaxel as a second-line regimen for advanced NSCLC was not found to have improved survival over the standard regimen of docetaxel alone. The effect seen in lower stage participants was intriguing, however. We hesitate to make definitive conclusions about the treatment effect based on this subgroup analysis, but do find it encouraging. A larger clinical trial designed specifically to test this therapeutic strategy in patients with stage IIIb disease without malignant effusion would be a natural next step.

Comment [A31]: good way to state it

TABLES AND FIGURES

Comment [A32]: P values are usually irrelevant in Table 1. All theydo is indicate type I errors. It is the means we care about

Table 1. Descriptive statistics of the patient population by treatment arm

	TOTAL (N = 188)		Treatment arm: TFD725 plus docetaxel (N = 98, 52.1%)		Treatment arm: Docetaxel alone (N = 90, 47.9%)		p-value *
	N	%	*N*	%	*N*	%	
Site location							0.78
North America	154	81.9%	81	82.7%	73	81.1%	
Europe	34	18.1%	17	17.3%	17	18.9%	
Gender							0.41
Male	104	55.3%	57	58.2%	47	52.2%	
Female	84	44.7%	41	41.8%	43	47.8%	
Stage at diagnosis							0.45
Stage IIIb without malignant pleural effusion	70	37.2%	39	39.8%	31	34.4%	
Stage IV or malignant pleural effusion	118	62.8%	59	60.2%	59	65.6%	
Achieved tumor response to first line therapy	107	56.9%	56	57.1%	51	56.7%	0.95
Abnormal LDH level at time of randomization	25	13.3%	9	9.2%	16	17.8%	0.08
Abnormal alkaline phosphatase level at time of randomization	48	25.5%	19	19.4%	29	32.2%	0.04
Performance status on ECOG scale							0.38
0 - Fully active	57	30.3%	34	34.7%	23	25.6%	
1 - Restricted in physically strenuous activity	122	64.9%	60	61.2%	62	68.9%	
2 - Ambulatory and capable of all selfcare but unable to carry out any work activities	9	4.8%	4	4.1%	5	5.6%	
Age at randomization (years) *Mean (SD), Min Median Max*	60.4 (5.4), 46 60 71		60.5 (4.8), 50 61 75		60.4 (5.1), 46 60.5 75		0.86
Time from initial diagnosis to randomization (months) *Mean (SD), Min Median Max*	10.4 (4.8), 3 10 31		10.2 (4.3), 3 10 27		10.3 (4.6), 3 10 31		0.82

** For categorical variables, the p-value is from a chi-squared test for independence between treatment arm and the variable. For continuous variables, the p-value is from a two-sample t-test with unequal variance comparing the means of the variable between treatment arms.*

Figure 1. Kaplan-Meier (KM) plots of survival probability comparing treatment arms overall and by disease stage at diagnosis (N = 188)

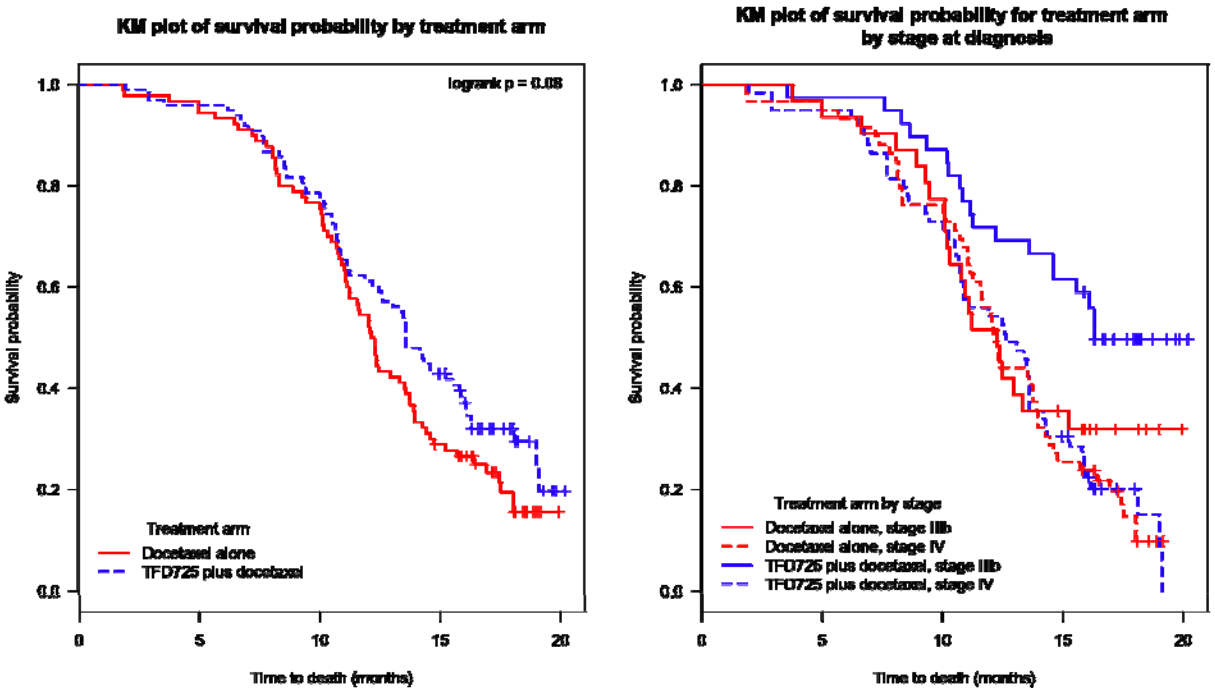

Table 2. Survival probability by treatment arm at 6, 12, and 18 months post-randomization

Treatment arm	6 months	95% CI	12 months	95% CI	18 months	95% CI
TFD725 plus docetaxel	95.9%	89.5% , 98.4%	61.2%	50.8% , 70.1%	32.0%	22.7% , 41.6%
Docetaxel alone	93.3%	85.8% , 96.9%	54.4%	43.6% , 64.0%	19.5%	11.4% , 29.1%

Table 3. Hazard ratios from unadjusted Cox proportional hazards regression models comparing treatment arms overall and by disease stage at diagnosis

	N	HR	95% CI	p-value
TFD725 plus docetaxel vs. docetaxel alone	188	0.75	0.54 , 1.04	0.08
TFD725 plus docetaxel vs. docetaxel alone by disease stage *				
Stage IIIb without malignant pleural effusion	70	0.53	0.29 , 0.99	0.05
Stage IV or malignant pleural effusion	118	0.99	0.67 , 1.46	0.95

* *Results from subgroup analysis by disease stage*

REFERENCES

1. American Cancer Society. *Cancer Facts & Figures 2005*. Atlanta: American Center Society; 2005.

2. Socinski MA. Clinical issues in the management of non-small-cell lung cancer and the role of platinum-based therapy. *Clin Lung Cancer*. 2004;5(5):274-289.

3. Penne K, Bohlin C, Schneider S, Allen D. Gefitinib (Iressa™, ZD1839) and tyrosine kinase inhibitors: The wave of future in cancer therapy. *Cancer Nursing*. 2005;28(6):481-486.

4. Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials. Non-small Cell Lung Cancer Collaborative Group. *BMJ*. Oct 7 1995;311(7010):899-909.

5. Huisman C, Smith EF, Giaccone G, Postmus PE. Second-line chemotherapy in relapsing or refractory non-small-cell lung review. *J Clin Oncol*. Nov 1 2000;18(21):3722-3730.

6. Shepherd FA, Dancey J, Ramlau R, et al. Prospective randomized trial of docetaxel versus best supportive care in with non-small-cell lung cancer previously treated with platinum-based chemotherapy. *J Clin Oncol*. May 2000;18(10):2095-2103.

7. Socinski MA, Morris DE, Masters GA, Lilenbaum R, American College of Chest P. Chemotherapeutic management of stage IV non-small cell lung cancer. *Chest*. Jan 2003;123(1 Suppl):226S-243S.

8. Di Maio M, Gridelli C, Normanno N, Peroone F, Ciardiello F. Epidermal growth factor receptor tyrosine kinase inhibitors in non-small cell lung cancer. *J of Cell Physiol*. 2005;205:355-363.

9. Gebbia V, Galetta D, De Marinis F. Non small cell lung cancer patients with ECOG PS2: unsolved questions and lessons from clinical trials. *Annals of Oncol*. 2005;16(S4):iv123-iv131.

10. Gritz ER, Dresler C, Sarna L. (2005). Smoking, the missing drug interaction in clinical trials: ignoring the obvious. *Cancer Epidemiol Biomarkers Prev*. 2005;14(10):2287-2293.

11. Hoang T, Schiller JH. Advanced NSCLC: from cytotoxic systemic chemotherapy to molecularly therapy. *Expert Rev Anticancer Ther*. Aug 2002;2(4):393-401.

12. Gridelli C, Massarelli E, Maione P, et al. Potential role of molecularly targeted therapy in the management of advanced nonsmall cell lung carcinoma in the elderly. *Cancer*. Oct 15 2004;101(8):1733-1744.