Effect of receptor tyrosine kinases inhibitor TFD725 upon risk of death in advanced stage IIIb and IV non-small cell lung cancer patients on second-line chemotherapy

Group 13

Summary

In 2003, a double-blind, randomized controlled Phase IIb trial was initiated to assess whether a second-line chemotherapy regimen containing 50 mg/m^2 of docetaxel every 3 weeks plus 50mg/day of TFD725, a receptor tyrosine kinases inhibitor, was associated with a difference in the instantaneous risk of death as compared with a regimen containing only 75 mg/ m^2 of docetaxel every 3 weeks alone. The patient population of concern was those patients with stage IIIb or IV non-small cell lung cancer. A total of 188 patients were randomized in a 1:1 ratio to receive either docetaxel/TFD725 (treatment, N=98) or docetaxel alone (control, N=90). With most variables, the distribution of baseline characteristics was comparable between the two groups. There were slight differences in the distribution of patients with high lactate dehydrogenase (LDH) and high alkaline phosphatase (Alk Phos); however, the difference in proportions was found to be insignificant when analyzing with the chi-square test (Alk Phos P-value: 0.129; LDH P-value: 0.065). Patients receiving docetaxel/TFD725 were found to be 0.747 times as likely to die at any given time as patients taking docetaxel alone, although this effect was not significant (P=0.084; CI: 0.536, 1.040). Of secondary interest to the study, the probability of survival at one year in the treatment group was 61.2% and in the control group was 54.4%. Thus, there was a 7% increase in probability of survival at one year among those in the treatment group as compared to the control group; however, this was also found to be statistically insignificant (P-value: 0.345; CI: -0.07, 0.21). At the conclusion of the study, 79.3% of patients had died within the treatment group and 84.4% of patients had died within the control group. (Diff: 0.04; P-value: 0.607; CI: -0.12, 0.20). We conclude that adding TFD725 to a docetaxel regimen does not affect instantaneous risk of death with stage IIIb or IV NSCLC.

Comment [A1]: followed for how long?

Comment [A2]: both of these are indicators of more severe disease (as is stage)

Comment [A3]: statistical significance is completely irrelevant here. Instead, we are concerned about the possibility of confounding which is not assessed by hypothesis testing.

Comment [A4]: 10% CI, 80% CI, 95% CI, 99% CI? which?

Comment [A5]: Not a typical way to present the comparison. We like to know a timeframe.

Background

Lung cancer is the most common as well as the leading cause of cancer death worldwide. There are approximately 1.35 million new cases of lung cancer yearly, and the disease causes 1.18 million deaths every year. Mortality rates are 31.2 per 100,000 males and 10.3 per 100,000 females (1, 2). Eighty percent of lung cancer cases are due to non-small cell lung cancer (NSCLC), which is defined as tumors that arise in large, squamous or glandular cells. The primary risk factor associated with disease is tobacco smoking, but additional exposures include radiotherapy, polycyclic aromatic hydrocarbons, asbestos, radon, nutritional and dietary factors and second-hand smoke (3).

Approximately 40% of NSCLC cases present in stage III or higher (4), when surgery is no longer an option, leaving chemotherapy and/or radiotherapy as the only viable treatment methods. Docetaxel has frequently been tested as both a first and second-line agent. In particular, as of 2000, there had been seven phase II trials of single-agent docetaxel used as a second-line treatment after first-line platinum-based therapies against NSCLC (12). Toxicity, response rates and survival associated with different docetaxel regimens have been evaluated to determine the recommended dose specifications for NSCLC patients who failed first-line non-doxcetaxel-based chemotherapy (5-7). One randomized controlled trial recommended a dosage regimen of once every three weeks to evoke a response without causing extreme toxicity (8). Further research has been conducted to assess survival benefits associated with the addition of targeted agents, such as gefitinib and bevacizumab. Thus far, bevacizumab is the only agent to have demonstrated prolonged survival. However, it must be noted that this agent was used in combination with a pacliataxel, carboplatin chemotherapy regimen and not docetaxel-containing therapy (9-11).

TFD725 has recently emerged as a potential second-line combination therapy for advanced NSCLC patients. TFD725 is a receptor tyrosine kinases inhibitor that has been shown in phase I and II clinical

trials to limit tumor growth and spread. In this study, we will compare a docetaxel/TFD725 combination therapy with single-agent docetaxel. Docetaxel was chosen as the taxane for the platinum-based regimen given its proven ability to lead to improved clinical outcomes and its current use as the primary regimen for advanced NSCLC patients.

Questions of Interest

The primary question of interest is assessing the impact of TFD725/docetaxel combination therapy on the instantaneous risk of death as compared to single-agent docetaxel. Secondary aims include assessing the difference in the probability of survival at one year and at the end of the study between treatment and control groups. One year was chosen as an easily interpretable time point that would be useful for clinicians.

Comment [A6]: You had no interest in adjusting for baseline variables or looking for effects within subgroups in this Phase II study?

Source of the Data

This was a double-blind, Phase IIb randomized controlled trial involving 188 patients ranging in age from 46 to 75 years, for whom data collection began in 2003. Clinical site and stage of disease at initial diagnosis determined randomization stratification. Individuals were eligible for inclusion in the study if they had stage IIIb or IV NSCLC, which had progressed on first-line platinum-based therapy. Individuals were ineligible if they were over 80 years of age at time of randomization, had poor performance status, had been treated with a first-line chemotherapy regimen that included docetaxel, or were unable to comply with a required contraception rule while in the study.

Patients were randomly assigned in a 1:1 ratio to receive either combination TFD725/docetaxel therapy (referred to as the "treatment" group) or docetaxel alone (referred to as the "control" group). Patients were instructed to remain on the assigned therapy unless they experienced symptoms caused by unacceptable toxicity, which were not managed by the modified dose specifications outlined in the research protocol. All patients were followed up to monitor clinical outcomes or death, regardless of whether or not they remained on therapy for the full observation time. Baseline data were collected on all 188 patients at the time of randomization, including basic demographic information (age, sex, geographic location), stage of NSCLC, tumor response to first line therapy (yes/no), indicator of abnormal LDH level (yes/no) and indicator of abnormal alkaline phosphatase level (yes/no). Performance status (0, 1, or 2) measured the patient's function using an instrument developed by the Eastern Cooperative Oncology Group (known as ECOG). A value of 3 or higher meant that the patient was capable of limited self-care or worse. At the end of observation time, the final outcome measure for all 188 patients was whether death had occurred or not during the study. There was no missing data for the patients in this study.

Comment [A7]: OK as stated, but censored data is a form of missing data

Statistical Methods

All data was analyzed using the R 2.10.0 statistical software (12). Because this was a randomized trial, we expected that, on average, the groups would be similar in all respects. To assess this assumption, we analyzed the baseline characteristics of the patients according to treatment group. By looking at the distributions of baseline variables, we determined if they were similarly distributed across treatment groups. For binary variables, where the distributions appeared to differ, a chi-square test of difference in proportions was used to determine if there were potential confounders, despite randomization.

Comment [A8]: these analyses are irrelevant. Confounding can occur in the absence of stat significance

The main question of interest for this project was to determine whether patients with advanced NSCLC had higher probability of survival when taking treatment as compared the control group. Specifically, we wished to determine if there was a difference in the overall hazard ratio of dying between the docetaxel/TFD725 and docetaxel groups. We also chose to compare the probability of survival at one year and at the end of follow-up. The time point of one year was chosen given that it is an easily interpretable period for use by clinicians when discussing survival probability with patients.

We wished to compare the overall effect of docetaxel/TFD725 on risk of death at any given time, as measured by the hazard ratio. Because some patients' time to death was censored, we used Kaplan-Meier methods to obtain the estimated survival curves for each chemotherapy group. The hazard ratio between treatment and control groups was then obtained using Cox proportional hazards regression with robust standard errors. Based on the resulting point estimate, 95% confidence intervals, and two-sided p-value, we tested the null hypothesis that the hazard ratio between the two groups is equal to 1. As a secondary analysis, we also compared the probability of survival between groups, using the Kaplan-Meier point estimates and standard errors at one year and at the end of the study. The confidence intervals and p-values associated with the difference in the probability of survival were calculated using Kaplan-Meier estimates at each time period (Appendix 1). All hypothesis tests used 0.05 as the nominal level of significance. It is important to note that we did not correct for multiple comparisons even though we performed several hypothesis tests on the data. While we could have used a Bonferroni correction, the tests we did perform are likely not independent of one another, in which case the Bonferroni correction would have been too conservative.

Finally, the random assignment of patients to chemotherapy groups is assumed to preclude confounding effects, so we did not include covariates in the model evaluating the effect of treatment on the instantaneous risk of death.

Comment [A9]: good to note

Comment [A10]: Bonferroni does not assume independence, it assumes mutually exclusive. Even more conservative, therefore.

Comment [A11]: Usually we do consider secondary adjusted analyses for precision, if nothing else. Also in a Phase II study we would be interested in exploring subgroups, even though such exploration is often not reproducible.

Results

Baseline descriptive statistics for each treatment group are represented in Table 1. A total of 188 patients with advanced NSCLC were randomized into two treatment groups, with 90 patients in the control group and 98 in the treatment group, none of whom had missing data. The mean age (60.5 years in control, 60.4 years in treatment) and mean duration of time between diagnosis and randomization (10.2 months in control, 10.4 in treatment) are similar between treatment groups, as were the percentage of males (52.2% control, 58.2% treatment). The frequencies seen on other demographic and baseline measures, such as the patient's site location, stage IV (or malignant pleural effusion) status, previous tumors with response to chemotherapy, performance status and whether the patient died, were all similar as well between treatment groups. The length of time study patients were observed ranged from 56 to 615 days. We did observe some differences in the proportion of patients with abnormal LDH (17.8% control; 9.2% treatment) and abnormal alkaline phosphatase (32.2% control; 19.4% treatment). However, using the chi-square test analyzing difference in proportions, we found these to not be statistically significant (Diff LDH P-value: 0.129; Diff alkaline phosphatase P-value: 0.065).

Comment [A12]: No, no, no. The patient that died at day 56 was actually observed much longer (he/she is still dead). Give descriptives of the censoring distribution.

Comment [A13]: I point out the irrelevance of this analysis for a third time because it is important that you understand that it is irrelevant.

Table 1: Patient characteristics at randomization by treatment

	Mean (SD)	Min.	Q1	Median	Q3	Max.	P-value
Docetaxel (N=90)							
Age (years)	60.5 (4.79)	50	58	61	63	75	
Time from Diagnosis to Randomization (months)	10.2 (4.34)	3	7	10	13	27	
European sites (%)	18.9						
Male (%)	52.2						
Advanced disease (%)	65.6						
Tumor response to first-line therapy (%)	56.7						
Abnormal LDH (%)	17.8						
Abnormal Alkaline phosphatase (%)	32.2						
Death (%)	80						
	Frequency (%)						
ECOG Performance Status							
0	25.6						
1	68.9						
2	5.6						
Docetaxel/TFD725 (N=98)							
Age (years)	60.4 (5.41)	46	57	60	64	71	
Time from Diagnosis to Randomization (months)	10.4 (4.77)	3	7	10	13	31	
European sites (%)	17.4						
Male (%)	58.2						
Advanced disease (%)	60.2						
Tumor response to first-line therapy (%)	57.1						
Abnormal LDH (%)	9.2						0.129
Abnormal Alkaline phosphatase (%)	19.4						0.065
Death (%)	69.4						
	Frequency (%)						
ECOG Performance Status							
0	34.7						
1	61.2						
2	4.1						

Note: There were no missing data by variable for either treatment or control.

Comment [A14]: Table 1 in an RCT report should not have P values. They are irrelevant (this is the fourth time—you can imagine how important this concept is. Arguably, the randomization imbalance does lessen our belief in the results here.)

Table 2 contains survival descriptive statistics derived from Kaplan-Meier estimates. At one year, 54.4% patients within the control group had survived and 61.2% of patients within the treatment group had survived. At the end of the study, 15.6% of patients within the control group had survived and 19.7% of patients within the treatment group had survived. Table 2 also shows the quartiles for survival time in days between the two treatment groups, with the treatment group having higher survival times at all quartiles (median survival was 370 days for the control group, 411 days for the treatment group). Figure 1 shows the Kaplan-Meier estimated survival curves for patients in the two treatment groups.

Table 2: Survival descriptive statistics by treatment

	Docetaxel (N=90)	Docetaxel/TFD725 (N=98)
Percentage Surviving (%)		
6 months	.93.3	95.9
12 months	54.4	61.2
18 months	19.5	32.0
End of study	15.6	19.7
Quartiles for Survival (days)		
75%	307	312
50%	370	411
25%	500	579

Comment [A15]: months would be a better unit here

Figure 1: Survival Curves by Treatment

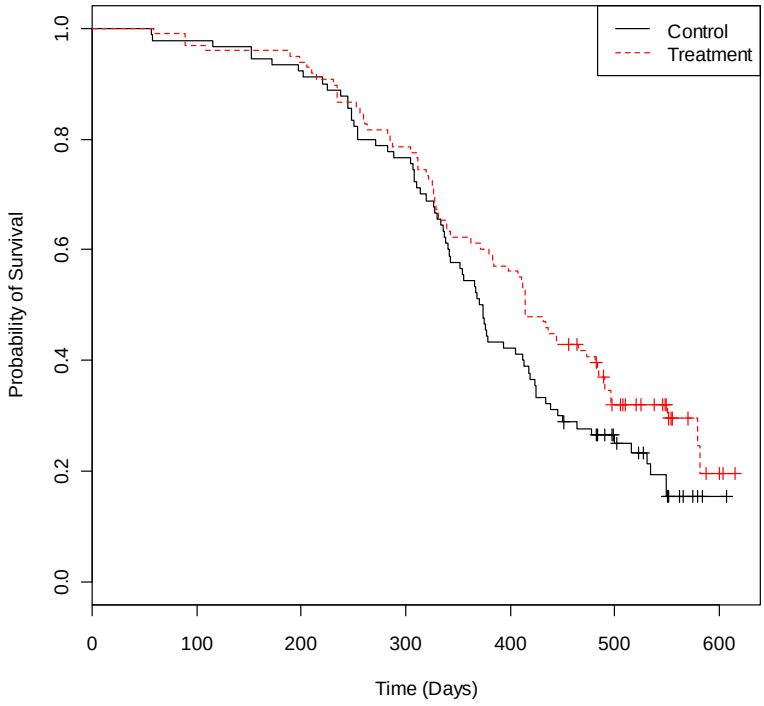

Table 3 contains the results of our survival analyses. In the primary analysis, we found the overall hazard ratio of treatment and control groups to be 0.747 (95% CI: 0.536, 1.04). At any given time, the instantaneous risk of death for patients in the treatment group was estimated to be 0.747 times that of patients in the control group. The two-sided p-value for this hazard ratio is 0.084. Because the p-value is greater than the significance level of 0.05, we cannot reject the null hypothesis of equal risk of death between patients receiving docetaxel/TFD725 and patients receiving docetaxel alone.

As a secondary analysis, we also compared the probability of survival between the two groups at one year and at the end of the study. At one year, a point estimate of the difference in the estimated probability of survival had the treatment group 0.07 higher than the control, with a two-sided p-value of 0.345. Based on the 95% CI, this would be a typical estimate of treatment effect if the true difference in survival probabilities at one year for the population were between 0.07 higher in the control group and 0.21 higher in the treatment group. At the end of the study, a point estimate of the difference in the estimated probability of survival had the treatment group 0.04 higher than the control, with associated two-sided p-value of 0.607. Based on the 95% CI, this would be a typical estimate of treatment effect if the true difference in survival probabilities at 1.7 years (total length of time patients were followed in study) for the population is between 0.12 higher in the control group and 0.20 in the treatment group. Based on the p-values, we cannot reject the null hypothesis of no difference in the probability of survival at either one year or at the end of the study between the two treatment groups.

Comment [A16]: Admittedly secondary, but more descriptive than HR, so I tend to talk about this first and then explicitly mention that HR is primary endpoint.

Table 3: Survival inferential statistics

	Point Estimate (SE)	95% CI	P-value
Hazard Ratio			
Docetaxel / (Docetaxel/TFD725)	0.747 (0.169)	(0.536, 1.040)	0.084
Probability of survival (1 year)			
Docetaxel	0.54 (0.053)	(0.45, 0.66)	
Docetaxel/TFD725	0.61 (0.049)	(0.52, 0.72)	
Difference (Docetaxel/TFD725-Docetaxel)	0.07 (0.072)	(-0.07, 0.21)	0.345
Probability of survival (End of study)			
Docetaxel	0.16 (0.044)	(0.09, 0.27)	
Docetaxel/TFD725	0.20 (0.066)	(0.10, 0.38)	
Difference (Docetaxel/TFD725-Docetaxel)	0.04 (0.080)	(-0.12, 0.20)	0.607

SE=Standard Error

Discussion

Conclusions from our analysis

We found no statistically significant difference in the instantaneous risk of death in patients receiving docetaxel/TFD725 and those receiving docetaxel only. Additionally, our secondary analysis found no difference in the probability of survival between the treatment groups at the time points of one year and at the end of the study. Based on our results, we cannot conclude that adding TFD725, a tyrosine kinases receptor inhibitor, to a stage IIIb or IV NSCLC patient's second-line chemotherapy regimen affects their risk of death at any moment or survival probability.

The Cox regression model we considered here did not include any covariates, but instead simply compared the estimated hazards of each treatment group, averaging across all other variables. While this is a reasonable approach, considering the randomization of the two groups, future studies may consider including covariates for reasons of precision. For example, although abnormal LDH and abnormal alkaline phosphatase levels are known to be predictive of poor survival outcomes (13,14), we did not include them as covariates. These two continuous variables were designed as binary variables within the study, thus they do not add much precision to a regression model with a binary outcome. If a future study measured the level of LDH or alkaline phosphatase over time, such information could be useful in predicting survival outcomes with more precision.

Comment [A17]: You never would have been allowed to not present the adjusted analyses as well. Referees would have forced your hand. That is why it is good to prespecify such secondary analyses in a protocol

Comment [A18]: NO, NO, NO, NO. We are interested in these values at baseline. We would almost never adjust for the post-randomization variables.

Limitations

Our study did not capture whether patients stopped the medication due to unacceptable toxicity. Other randomized controlled trials have tracked change in treatment due to toxicity experienced by the patients, (for example, neutropenia, thrombocytopenia, vomiting, diarrhea, or neuropathy) (15). Capturing the reason that the drug was stopped could provide additional information regarding causes for censoring. Other studies have also measured impact on quality of life using questionnaires. However, this study only measured survival times and outcomes. For a possibly highly toxic chemotherapy treatment, such as the one used in this study, quality of life measures may have provided additional insight. For example, prospective ECOG measurements throughout the study could have provided a way to track impact of the treatment over time on patient's quality of life. Additionally, we did not know the patients' first-line therapy regimen; we only knew whether or not they responded to the therapy. While randomization at baseline reduced the possibility of confounding, we still not know the type of patient's first-line chemotherapy regimen, which would be useful to more accurately describe the patient population for which these results can be generalized.

Comment [A19]: NO. This was a well-done study, meaning that we did not censor subjects just because they stopped therapy. That would be a terrible thing to do.

Future Studies

Though no significant benefit in the instantaneous risk of death was found between patients taking docetaxel/TFD725 compared to patients taking docetaxel alone, there are several possible applications for TFD725 that this study did not explore. Future studies could use a different dose of TFD725. It is also possible that different types of NSCLC may respond to TFD725 better than other types. It would be useful to determine whether TFD725 has a noticeable impact upon risk of death at any moment and survival probability for the different classifications of NSCLC.

Comment [A20]: You could have explored the possibility of these things in the current data. Again, it would generally be expected that you would have explored such things in a real article.

Appendix 1:

The probability of survival at time *t* for a patient in the treatment group (*T*) is asymptotically normally distributed,

$$\hat{\theta}_T \sim N\left(\theta_T, se^2\left(\hat{\theta}_T\right)\right).$$

Likewise, the probability of survival at time *t* for a patient in the control group (*C*) is asymptotically normally distributed,

$$\hat{\theta}_C \sim N\left(\theta_C, se^2\left(\hat{\theta}_C\right)\right).$$

Then the difference in survival probabilities at time *t* is also asymptotically normally distributed,

$$\hat{\theta}_T - \hat{\theta}_C \sim N\left(\theta_T - \theta_C, se^2\left(\hat{\theta}_T\right) + se^2\left(\hat{\theta}_C\right)\right)$$

Using this asymptotic distribution, a Wald confidence interval and p-value can be calculated in the usual way.

References

1. Parkin DM, Bray F, Fertay J, Pisani R. Global cancer statistics, 2002. *CA Cancer J Clin*. 2005;55:74-108.
2. Ries LAG, Eisner MP, Kosary CL, et al: SEER Cancer Statistics Review 1973-1999. Bethesda, MD, National Cancer Institute. http://seer.cancer.gov/csr/1973_1999.
3. Ruano-Ravina A, Figueiras A, Barros-Dios JM. Lung cancer and related risk factors: an update of the literature. *Public Health*. 2003 117(3):149-156.
4. Belani C. Optimizing chemotherapy for advanced non-small cell lung cancer: focus on docetaxel. *Lung Cancer*. Dec 2005;50 Suppl 2: 50S2:S9-16.
5. Kubota K, Watanabe K, Kunitoh H, et al. Phase III randomized trial of docetaxel plus cisplatin versus vindesine plus cisp[atin in patients with stage IV non-small-cel lung cancer: the Japanese Taxotere Lung Cancer Study Group. *J Clin Onco*l. 2004;22:254-61.
6. Fossella F, Pereira JR, yon Pawel et al. Randomized, multinational, phase III study of docetaxel plus platinum combinations versus vinorelbine plus cisplatin for advanced non-small-cell lung cancer: the TAX 326 study Group. *J Clin Oncol*. 2003;21:3016-24.
7. Georgoulias V, Papadakis E, Aexopoulos A, et al. Platinum-based and non-platinum-based chemotherapy in advanced non-small-cell lung cancer: a randomised multicentre trial. *Lancet*. 2001;357:1478-84.
8. Camps C, Massuti B, Jiménez A, Maestu I, Garcia Gómez R, Isla D et al. Randomized phase II study of 3-weekly versus weekly docetaxel in pretreated advanced non-small cell lung cancer: a Spanish Lung Cancer Group trial. *Ann Oncol*. Mar 2006;17(3):467-72.
9. Lynch T Jr, Kim E. Optimizing chemotherapy and targeted agent combinations in NSCLC. *Lung Cancer*. Dec 2005;50 Suppl 2:S25-32.
10. Rigas JR, Lara PN Jr. Current perspectives on treatment strategies for locally advanced, unresectable stage II non-small cell lung cancer. *Lung Cancer*. Dec 2005;50 Suppl 2:S17-24.
11. IASLC 11th World Conference on Lung Cancer. Slides presented during two satellite symposia. 23 March 2007.
12. R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.
13. Koukourakis MI, A Giatromanolaki, E Sivridis, G Bougioukas, V Didilis, KC Gatter, & AL Harris. Lactate dehydrogenase-5 (LDH-5) overexpression in non-small-cell lung cancer tissues is linked to tumour hypoxia, angiogenic factor production and poor prognosis. *British Journal of Cancer*. 2003;89:877 – 885.
14. Brown JE, Cook RJ, Major P, et al. Bone turnover markers as predictors of skeletal complications in prostate cancer, lung cancer, and other solid tumors. J Natl Cancer Inst. 2005 Jan 5;97(1):59-69.
15. Shepard FA, Dancey J, Ramlau R, et al. Prospective Randomized Trial of Docetaxel Versus Best Supportive Care in Patients With Non–Small-Cell Lung Cancer Previously Treated With Platinum-Based Chemotherapy. *J Clin Oncol*, May 2000:18(10): 2095-2103.