**Effect of the Addition of TFD725
to Second-Line Chemotherapy on Survival in Non-Small Cell Lung Cancer**

Summary

Backgound: Lung cancer is the principal cause of cancer related mortality in both men and women in the United States, and non-small cell lung cancer accounts for the majority of lung tumors. Although first-line treatment with platinum–based chemotherapy has increased survival for late-stage patients, effective second-line therapies are still under development.

Objective: This multicenter, Phase IIB, double-blind, randomized trial was conducted to investigate the use of an experimental second-line chemotherapeutic agent, TFD725, in combination with an established second-line agent, docetaxel, versus docetaxel alone to improve survival of non-small cell lung cancer patients. The primary endpoint of interest was the hazard ratio comparing instantaneous risk of death between the treatments arms, and the secondary endpoint of interest was 12-month proportion surviving.

Methods: Patients (n=188) were randomized in 1:1 ratio to the treatment arm (50 mg/m^2 of docetaxel per 3 weeks and 50 mg/m^2 TFD725 per day, n=98), or the control arm (75 mg/ m^2 docetaxel per 3 weeks, n=90). Randomization was stratified by clinical site and disease stage at diagnosis (Stage IIIB without malignant pleural effusion vs. Stage IIIB with malignant pleural effusion or Stage IV). The median duration of follow-up was 551 days in the control group and 546 days in the treatment group. The Kaplan-Meier method was used to estimate the proportion of patients who survived to 9, 12, 15, and 18 months in the control and treatment arms of the study. The Cox proportional hazard regression was utilized to estimate the hazard ratio comparing the instaneous risk of death between the TFD725+docetaxel group and the docetaxel group. Exploratory analyses using additional Cox regression models examined potential effect modifiers and precision variables in order to generate hypotheses to test in future research.

Results: The Cox proportional hazards regression estimated a hazard ratio of 0.75, which compares instantaneous risk of death between patients in the TFD725+ docetaxel group to the docetaxel group (95% confidence interval, 0.54 to 1.04; p=0.08). No statistically significant difference in hazard of death between treatments arms was observed. The 12-month survival probability was 6.8% higher in the TFD+docetaxel group than the control group (95% confidence interval, 7.3% lower to 20.9% higher; p=0.35). Stratification by abnormal LDH levels and abnormal alkaline phosphatase levels did not produce any statistically significant associations in Cox regressions. Stratifying by stage of disease at diagnosis produced a significant hazard ratio between the treatment groups among those with Stage IIIB disease at diagnosis; among these patients, those taking TFD725+docetaxel had a 0.53 times the instantaneous risk of death as those in the control group (95% confidence interval, 0.29 to 0.99; p=0.05).

Conclusion: There was no significant difference in the hazard ratio comparing instantaneous risk of death or 12-month proportion surviving between the treatments arms. Overall, the current study does not provide evidence that second-line therapy of TFD725 and docetaxel improves survival of late-stage NSCLC patient, compared to docetaxel alone. The data suggest that survival may be improved by the addition of TFD725 to second-line therapy, for patients initially diagnosed with Stage IIIB NSCLC without malignant pleural effusion, but not for patients diagnosed with Stage IIIB with malignant pleural effusion or Stage IV NSCLC. Stage of disease at initial diagnosis may be a potentially important effect modifier and warrants further investigation.

Background

In the United States, lung cancer is the principal cause of cancer related mortality in both men and women. The current study focuses on non-small cell lung cancer (NSCLC), including adenocarcinomas, squamous cell carcinomas and large cell carcinomas. NSCLC accounts for about 84% of all lung tumors *(1)*. Treatment options depend upon initial stage at diagnosis *(2)*, and stage of disease at diagnosis is a critical factor in NSCLC prognosis (*3*). For patients who present with advanced disease (defined as stage IIIB or IV), surgery is not a viable option. Stage IIIB cancer is treated with chemotherapy and radiation, whereas stage IV cancer is treated with chemotherapy and supportive care, or supportive care alone *(2)*. The prognosis for patients with advanced NSCLC is poor, with 5-year survival rates of 5-15% for stage IIIB disease and <5% for those with stage IV disease. Chemotherapy is an essential option for cancer mitigation *(1)*. For NSCLC, first-line chemotherapy is typically platinum-based, with cisplatin and carboplatin being the most commonly-used agents *(4)*. Additionally, serum lactate dehydrogenase

Comment [A1]: You could have given brief comments about the age, sex, and stage of disease distributions

Comment [A2]: instantaneous

Comment [A3]: statistical jargon If the list is not too long (and it never should be too many) list the variables considered. E.g. "Exploratory analyses adjusted for stage of disease, abnormal LDH and alkaline phosphatase levels, sex and age. Additionally, treatment effects were examined within subgroups defined by those variables.

Comment [A4]: In the context of PH regression, this usually means adjustment for those variables in a very special way. "Subgroup analyses" would be clearer here.

Comment [A5]: You have a huge multiple comparison issue here. I would avoid calling anything statistically significant or not. I would just give the positives and the pertinent negatives. Having listed the variables you considered above (at least I listed them), you could just comment on the stage.

Comment [A6]: This is not at all significant if you take multiple comparisons into account. Instead, this is an interesting group to explore further

Comment [A7]: I think it highly important that you give the results in the other subgroup as well—we need to know whether the drug is harmful or merely equal to placebo in the advanced stage group

Comment [A8]: I would soften your wording here. Your last sentence is sufficiently cautious, this one not as much.

(LDH) (*3, 5, 6*) and alkaline phosphatase levels *(6)* have been identified as important prognostic factors for NSCLC.

Comment [A9]: This sentence seems a little out of place. Perhaps move it back to where you talk about prognosis.

Many patients with NSCLC experience relapse, necessitating second-line chemotherapy. Docetaxel and paclitaxel are typically used as second-line chemotherapeutic agents, but there is a growing need for additional and improved second-line chemotherapeutic regimens for NSCLC *(1)*. New therapies combining docetaxel with molecular-targeted agents like tyrosine kinase inhibitors have shown promise *(7)*.

The current study investigates the use of TFD725, an experimental tyrosine kinase inhibitor, as a second-line chemotherapeutic agent in combination with docetaxel.

Questions of Interest

1. Does TFD725 increase or decrease the instantaneous risk of death for patients with NSCLC when used as second-line chemotherapy with docetaxel, relative to docetaxel alone?
2. Does TFD725 increase or decrease survival probability at twelve months for patients with NSCLC when used as a second-line chemotherapy with docetaxel, relative to docetaxel alone?
3. Do known predictors of poor patient outcomes affect the primary and secondary outcomes of interest? Specifically, do abnormal LDH levels, abnormal alkaline phosphatase levels, or stage of disease at initial diagnosis, increase or decrease either the instaneous risk of death or the survival probability at twelve months for patients with NSCLC by treatment group?

Comment [A10]: I do not think these are at all interesting questions. We already know that these are important risk factors, and the only value in reporting these would be to show that your sample was typical of other NSCLC studies. Instead, we would be interested in looking at analyses of the treatment effect adjusted for these variables and looking at effect of treatment within subgroups defined by these variables. (I think you meant to say: "increase or decrease the treatment's effect on either the instantaneous risk of death of the survival probability". You never even presented the answer to the question as you wrote it (and as noted, we do not want to see the answer to the question as you wrote it._

Source of the Data

This trial was a multicenter, Phase IIB, double-blind, randomized trial of TFD725 and docetaxel versus docetaxel alone. All patients (n=188) had progressed on platinum-based chemotherapy as a first-line treatment for NSCLC. To be eligible, patients must have had either stage IIIB or stage IV disease at diagnosis, must not have received docetaxel as part of their first-line chemotherapy regimen, must have had ECOG performance status 2 or better at randomization, must have been ≤80 years old at randomization, and must have agreed to use effective contraception for the duration of the study.

Patients were randomized in a stratified manner, by clinical site and disease stage at diagnosis (stage IIIB without malignant pleural effusion vs. more advanced disease). Patients in the treatment arm (n=98) received docetaxel (50 mg/m^2 per 3 weeks) and TFD725 (50 mg/m^2 per day), whereas patients in the control arm (n=90) received docetaxel alone (75 mg/m^2 per 3 weeks). Patients were to discontinue study drug if unacceptable toxicities occurred that did not resolve after making pre-specified dose modifications to their regimens.

At baseline, each patient's medical history, physical exam findings, and ambulatory status were recorded. These data included demographics (study site, age, sex), the presence/absence of advanced disease at diagnosis, tumor response to first-line treatment, laboratory measures of disease at randomization (LDH and alkaline phosphatase), and performance status on the ECOG scale. Patients were followed for survival until death or the end of the study, whichever occurred first.

There were no missing values on any variable for the 188 patients enrolled in the study. Patients who were not observed to have died during the study were labeled "censored" for the purposes of study analysis.

Comment [A11]: You could strengthen this statement to say that "patients still alive at the time of data analysis were censored". This makes more clear the fact that there was no censoring due to loss of follow-up.

Methods

The study was designed with an intent-to-treat analysis; all patients were followed for the duration of the study regardless of whether they complied with the treatment regimen. The primary outcome of the study was the hazard ratio comparing instantaneous risk of death between the treatments arms, and secondary endpoint of 12-month survival.

Descriptive statistics for each treatment group were calculated using standard formulas for the arithmetic mean, standard deviation, median, and percentiles. In order to identify any statistically significant differences in demographic characteristics at baseline, we performed a two-sample t-test with unequal variances on the continuous variables (age and time from initial diagnosis to randomization), and a test of binomial proportions to compare the

Comment [A12]: It is silly to perform such tests. We do not care. Any "significant" differences are certainly type I errors. Lack of statistical significance does not mean there is no meaningful confounding, and presence of statistical significance does not mean there is meaningful confounding.

Comment [A13]: We would generally call this the "chi squared test" in order to facilitate its recognition by the reader.

mean or proportion between the treatment arms on the binary variables (European site, male, advanced stage, response to first line therapy, abnormal LDH, and abnormal alkaline phosphatase). A chi-squared test was performed to assess whether the proportion of subjects having ECOG levels 0, 1, and 2 were different between the treatment arms. For these and all other statistical tests, the level of significance was pre-specified to be $\alpha=0.05$.

A Cox proportional hazards model was used to compare the instantaneous risk of death between treatment groups, which was the primary endpoint, with a hazard ratio of less than one suggesting better survival in the treatment group. This method provides the average estimated hazard ratio over time, for each group, where the "hazard" is the risk of the event occurring (in this study, death) at any given time point, given that a subject survived to at least that time point. Confidence intervals were calculated under the assumption that hazards ratio are constant over time (the "proportional hazards" assumption). There was nothing to indicate that this was not a valid assumption in this dataset.

Comment [A14]: Under the strong null hypothesis, this assumption must hold. So this is not such a big deal to worry about. I do note that we never have much power to detect nonproportional hazards.

We used the Kaplan-Meier method to estimate the proportion of patients who survived to 9, 12, 15, and 18 months in the control and treatment arms of the study; of these, we chose 12-month survival as the secondary endpoint because it is a clinically-accepted comparison measure. The 95% confidence intervals (CI) were derived using standard errors calculated with Greenwood's Formula; the intervals indicate the range of the true population proportion surviving for which it would not be unusual to obtain the observed results. The Kaplan-Meier method assumes that censoring is non-informative; that is, patients who were censored are no more or less likely to experience an event (in this study, death) than patients who remained under observation. Although nothing in a dataset can fully validate this assumption, we analyzed the censoring distribution with a Cox regression, redefining censoring as the event of interest, to see if significant differences existed between treatment groups. *A priori*, non-informative censoring was a reasonable assumption.

Comment [A15]: The PH model also has this assumption.

Comment [A16]: Owing to the "administrative censoring" being the sole cause of censoring.

We used a Cox regression model to assess whether pre-specified variables were effect modifiers or precision variables. (We did not expect to find confounders, due to the randomized nature of the study; we were most interested in identifying precision variables and effect modifiers.) Previous clinical knowledge indicates that abnormal LDH, abnormal alkaline phosphatase, and advanced stage of disease at diagnosis are strong predictors of poor outcomes. Therefore, we planned to begin our exploratory analyses by testing each variable individually as an interaction term in the Cox regression model. For the covariates that were not effect modifiers, we then checked whether they were associated with survival time in each treatment group; that is, we checked to see if the variable was a precision variable. We fit six simple Cox regression models for survival time. In each model we restricted our analysis to either the placebo or treatment group and tested one covariate of interest alone in the model (stage of disease, abnormal LDH, and abnormal alkaline phosphatase). If the results were significant in either treatment group, we regarded it as an indication of association between survival time and the covariate, given treatment group. We explored the association between treatment and survival time among each subgroup defined by the effect modifier or precision variable to see whether the treatment effect was significant in certain subgroups. Because these analyses were exploratory, we planned to interpret their results and p-values with caution. There were no adjustments made for multiple comparisons.

Comment [A17]: Testing for intereractions is less of interest than is assessing whether the treatment works in subgroups. It requires a larger sample size to detect an interaction than it does to establish an effect of treatment within a subgroup.

Comment [A18]: This is the reverse order that we would usually do. We would perform an adjusted analysis before looking within subgroups. Furthermore, we do not care to "prove" something is a precision variable. We generally go on prior results.

Comment [A19]: This *ad hoc* method is not particularly well positioned to answer your questions. Why not examine the effect in both groups together with double the sample size?

All analyses were conducted using Stata (version 11, copyright StataCorp) and R (version 2.91).

Results

Baseline characteristics of the study population are presented in Table 1. The patients in each treatment arm had similar demographic characteristics and severity of disease at baseline, as expected in a randomized trial. Only the proportion of patients with abnormal alkaline phosphatase was significantly different between the treatment arms ($p=0.04$).

Comment [A20]: Again, statistical significance is irrelevant. The question is: Is the 9% difference in proportion of patients with abnormal LDH indicative of an important difference in survival? The fastest way to address this is just to do an adjusted analysis with all important predictors.

The duration of follow-up varied from 56 days to 615 days (20.2 months). In the treatment arm, 30 (30.6%) patients were censored over the duration of follow-up, versus 18 (20.0%) in the control arm. The median duration of follow-up was 551 days in the control group and 546 days in the treatment group. From the Cox regression, the censoring distribution did not differ significantly between treatment groups (two-sided p=0.856).

The proportion of patients surviving in each treatment group at 9, 12, 15, and 18 months is presented in Table 2, as estimated using the Kaplan Meier method. Median survival was 374 days for the control group, and 414 days for the treatment group. Among the TFD725+docetaxel group, the survival probability at 12 months was 6.8% higher than that of control group (95% CI, 7.3% lower to 20.9% higher; p=0.34). The survival proportions became more divergent between the treatments groups as time from randomization increased. At 9 months the survival probability in the TFD725+docetaxel group was 2.7% greater than that of the docetaxel group (95% CI, 8.7% lesser to 14.1% greater; p=0.64); whereas at 18 months, the survival probability in the TFD725+docetaxel group was 12.5% greater than that of the docetaxel group (95% CI, 0.65% lesser to 25.7% greater; p=0.08). The wide confidence intervals suggest a lack of precision in these estimates, however.

The hazard ratio was 0.75, comparing the instantaneous risk of death between the TFD725+docetaxel group and the control group (95% CI, 0.54 to 1.04; p=0.08), as shown in Table 3.

The survival curves of the proportion of patients surviving versus days from randomization by treatment group are presented in Figure 1A. Survival between treatment groups was nearly identical soon after randomization and the survival between groups diverge after 12 months. A statistically significant difference in survival probability between the groups was observed at 15 months (14.0% higher among the treatment group; 95% CI, 0.4% higher to 27.5% higher; p=0.04).

Among the three variables individually tested in the Cox regression model – abnormal LDH, abnormal alkaline phosphatase, and advanced disease at diagnosis – only advanced disease status was a potential effect modifier, although it was not statistically significant (two-sided p = 0.075). The survival probability estimated at 12 months in each group defined by treatment group and stage is shown in Table 3. We found that among patients with Stage IIIb disease without malignant pleural effusion at diagnosis, the treatment group had a 20.2% higher (95% CI, 2.4% lower to 42.7% higher; p=0.08) probability of survival at 12 months. According to the Cox regression, the risk of death in the treatment group was 0.53 (95% CI, 0.29 to 0.99) times that of the control group. This difference was significant (p = 0.05) and allows us to reject the null hypothesis that there was no difference in survival between treatment groups among patients with Stage IIIB disease at diagnosis. However, among patients with advanced disease at diagnosis (Stage IV disease or malignant pleural effusion), the treatment group had a 1.7% lower 95% CI, 19.6% lower to 16.3% higher; p=0.85) survival probability at 12 months and those in the TFD725+docetaxel group had 0.99 (95% CI, 0.67 to 1.46, p=0.96) times the instantaneous risk of death as compared to the patients treated with docetaxel alone.

Abnormal LDH and abnormal alkaline phosphatase were also significantly associated with survival time. Since the treatment groups did not have equal proportions of patients with abnormal LDH and abnormal alkaline phosphatase at randomization, we performed subgroup analyses with these two variables. The survival curves defined by treatment group and each of these variables are shown in Figure 1C and Figure 1D, respectively. Table 3 shows the estimated difference in survival probability at 12 months and the hazard ratios for each group. None of the results were significant, although among patients with abnormal alkaline phosphatase at baseline, there was a 6.2% lower (95% CI , 22.5% higher to 34.9% lower; p=0.67) survival probability in the treatment group relative to placebo at 12 months.

Comment [A21]: NO. The minimum follow-up was much longer than 56 days. The earliest death was 56 days, but we followed that patient for much longer. (He/she is still dead.) It is very difficult to get descriptive minima and maxima from censored observations, though you could report the minimum observed censoring time if that were really important.

Comment [A22]: Mention your figure here, because your discussion about diverging survival curves (which cannot possibly be PH if they are truly divergent in this manner) is readily apparent in the graphs.

Comment [A23]: HUGE multiple comparison problem fishing through your data to find a point of maximal statistical significance. Therefore, i would not play up the statistical significance here. Let the results of Table 2 and your discussion above suffice.

Comment [A24]: I would be a little more descriptive here. Talk about interesting trends toward differences in survival that were seen in these exploratory analyses. Reference your figures and table. Then comment that these differences cannot be judged statistically significant due to the multiple comparison issues.

Comment [A25]: Nope, not significant due to the multiple comparisons. (Intriguing for future study, and if this were our primary question in a later study and was stat signif in that study, then I would be talking like you are here.)

Comment [A26]: In any subgroup analysis, it is EXTREMELY important to talk about both subgroups like you do here. I would have included this verbage in the Summary as well. There is a huge difference between no effect overall with a benefit in one subgroup and harm in the other and a case where slight effect in the overall sample appears due to strong effect in one subgroup and neutral effect in the other. We are far more willing to continue to explore the latter, than the former due to safety.

Comment [A27]: You continue to make too much of statistical significance. This is an exploratory analysis in which we do not even begin to know how to interpret your “p value” as a true p value.

Discussion

In this multicenter, Phase IIB, double-blind, randomized trial, TFD725 did not significantly improve survival when added to docetaxel as second-line therapy for NSCLC, compared docetaxel alone. Because the censoring distributions did not differ between treatment groups, we can infer that the addition of TFD725 did not compromise a patient's capacity to participate in the study. That is, neither side effects nor lack of treatment effect influenced the likelihood that a patient would return for observation and data collection.

Although the difference between treatment groups was not statistically significant, patients receiving TFD725 in addition to docetaxel had a 6.8% higher survival probability at 12 months than patients receiving docetaxel alone. The hazard ratio indicated that instantaneous risk of death was lower for patients receiving TFD725 than those receiving docetaxel alone; however, this difference was also not statistically significant (HR = 0.75; 95% CI, 0.54 to 1.04). Even though no statistically significant difference was observed between the treatment groups, the data do not preclude the existence of such a difference. Patients receiving TFD725 had a 14% higher survival proportion at 15 months than those who received docetaxel alone, and this difference was significant (two-sided p=0.04). These data may suggest that any purported benefit of TFD725 is manifest only over the long-term; patients who die early during the course of second-line therapy may be at such increased risk of death that treatment will not affect their survival outcomes. Since NSCLC survival is low over time intervals longer than 12 months, TFD725 may have limited utility as a treatment for this type of cancer. The width of the 95% confidence interval for the difference in 15-month survival (0.412% to 27.5% greater survival in the treatment group) highlights the lack of statistical precision in the survival estimates. Further investigation is warranted to determine whether the true survival difference is closer to 0.412% or 27.5%, as the magnitude of the difference will guide clinical decision-making and affect how side effects or adverse events are weighed against possible benefits of TFD725. Also, further investigation with larger sample sizes may be able to demonstrate a significant association between TFD725 and survival at 9- and 12-month time intervals.

Alternatively, the lack of statistically significant differences in 12-month survival may be due to effect modification by stage of disease at diagnosis. Our exploratory analyses suggest that TFD725 treatment may improve survival only for patients diagnosed with Stage IIIB without malignant pleural effusion (Table 3; Figure 1B). Among patients with Stage IV disease or malignant pleural effusion at diagnosis, treatment did not significantly alter (or worsen) survival (HR=0.99; 95% confidence interval, 0.67 to 1.46). However, because this inference relies on subgroup analysis, it should be interpreted with caution. A further study, including only patients with Stage IIIB disease without malignant pleural effusion at diagnosis, would be required to demonstrate the benefit of TFD725 in this patient population.

Our exploratory analyses also indicate that TFD725 may not provide any benefit to patients who have an abnormal alkaline phosphatase at the time second-line treatment is initiated. Among patients with abnormal alkaline phosphatase levels, treatment was associated with a lower survival at 12 months and a higher risk of death. However, these results should also be interpreted with extreme caution, because the confidence intervals were wide and p>>0.05 for both statistics. There were relatively few people with abnormal alkaline phosphatase in this study (29 in the control group, 19 in the treatment group); to evaluate whether TFD725 toxicity differs for patients with abnormal alkaline phosphatase, another trial with a larger sample size of these patients would be necessary.

Taken as a whole, this multicenter, Phase IIB, double-blind, randomized trial does not demonstrate that the combination of TFD725 and docetaxel improves survival of late-stage NSCLC patients as compared to docetaxel alone.

Comment [A28]: We never conclude equality. That takes an infinite sample size.

Comment [A29]: You don't know this. Maybe side effects on the treatment arm balanced out lack of effect on the placebo arm.

Comment [A30]: And these might be the stage IV patients that you discuss below

Comment [A31]: This statement makes no sense.

Comment [A32]: Or closer to -10%. You have a multiple comparison problem here. Mention it. You are focusing on one of four different time periods you looked at. This might (in worse case) be a 80% CI, not a 95% CI.

Comment [A33]: Or it might demonstrate that TFD725 is harmful.

Comment [A34]: I would phrase this more as "our results suggest that it ight be most interesting to restrict further study to the stage III group.

Comment [A35]: Are these the stage IV patients?

Comment [A36]: We never (well almost never) expect a Phase II trial to demonstrate efficacy or effectiveness. Instead we hope the results of such a trial to guide us in performing a Phase III study

References

1. Bonomi PD. Therapeutic advances in second-line treatment of advanced non-small-cell lung cancer. *Clin Lung Cancer*. 2004;6(3):154-161.

2. Silvestri GA, Tanoue LT, Margolis ML, Barker J, Detterbeck F. The noninvasive staging of non-small cell lung cancer: the guidelines. *Chest*. 2003;123;147S-156S.
3. Osterlind K. Factors confounding evaluation of treatment effect in lung cancer. *Lung Cancer*. 1994;10 Suppl 1:S97-103.

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

5. Cağlayan B, Fidan A, Salepçi B, et al. Effects of prognostic factors and treatment on survival in advanced non-small cell lung cancer. *Tuberk Toraks*. 2004;52(4):323-332.

6. Ray P, Quantin X, Grenìer J, Pujol JL. Predictive factors of tumor response and prognostic factors of survival during lung cancer chemotherapy. *Cancer Detect. Prev*. 1998;22(4):293-304.

7. Fossella FV, Lynch T, Shepherd FA. Second line chemotherapy for NSCLC: establishing a gold standard. *Lung Cancer*. 2002;38 Suppl 4:5-12.

Table 1. Descriptive statistics for patient characteristics by treatment group.

	Docetaxel Alone (n=90)				*Docetaxel and TFD725 (n=98)*			
	mean (sd)	**min**	**median**	**max**	**mean (sd)**	**min**	**median**	**max**
Age at randomization (years)	60.51 (4.79)	50	61	75	60.38 (5.41)	46	60	71
Time from diagnosis to randomization (months)	10.23 (4.35)	3	10	27	10.39 (4.78)	3	10	31
	proportion				**proportion**			
Proportion from a European site	18.1%				17.3%			
Proportion male	52.2%				58.2%			
Proportion with malignant pleural effusion (stage IV)	65.6%				60.2%			
Proportion with tumor response to first therapy	56.7%				57.1%			
Proportion with abnormal LDH level at randomization	17.8%				9.2%			
Proportion with abnormal alkaline phosphatase level at randomization	32.2%				19.4%			
ECOG Scale Score of 0	25.6%				34.7%			
ECOG Scale Score of 1	68.9%				61.2%			
ECOG Scale Score of 2	5.6%				4.1%			

* Measurement taken at randomization

**Patients with malignant pleural effusion were Stage IV, and remaining patients were Stage IIIB

Table 2. Kaplan-Meier estimates of survival probability by treatment group.

	% Survival (95% CI) by Treatment Group	
Months	***Docetaxel***	***Docetaxel + TFD725***
9	78.9 (68.9, 86.0)	81.6 (72.4, 88.0)
12	54.4 (43.6, 64.1)	61.2 (50.8, 70.1)
15	28.9 (12.0, 38.4)	42.9 (33.0, 52.4)
18	19.5 (11.4, 29.1)	32.0 (22.7, 41.6)

Comment [A37]: You gave the differences and inference in the text. That could suffice, but I would likely have included it as a third column here.

Table 3: Estimated differences in survival probabilities and hazard ratios between treatment groups for all patients and for pre-specified subgroups.

Covariate[1]	Stratum	% Difference in 12-Month Survival Probability[2], Treatment – Control (95% CI)	Two-sided p-value	Hazards Ratio, Treatment:Control (95% CI)	Two sided p-value
All Patients	n/a	6.8 (-7.3, 20.9)	0.34	0.75 (0.54, 1.04)	0.08
Advanced disease at diagnosis[3]	Not advanced	20.2 (-2.4, 42.7)	0.08	0.53 (0.29, 0.99)	0.05
	Advanced	-1.7 (-19.6, 16.3)	0.85	0.99 (0.67, 1.46)	0.96
LDH at diagnosis	Normal	1.9 (-13.0, 16.8)	0.80	0.79 (0.55, 1.14)	0.21
	Abnormal	14.6 (-21.7, 50.8)	0.43	0.72 (0.32, 1.62)	0.43
Alkaline Phosphatase at diagnosis	Normal	8.4 (-7.8, 24.7)	0.31	0.74 (0.49, 1.11)	0.14
	Abnormal	-6.2 (-34.9, 22.5)	0.67	1.04 (0.56, 1.95)	0.90

1. Each covariate was entered individually into a simple Cox proportional hazards regression model.
2. Difference between treatment arms in absolute percentages (not relative).
3. Not advanced = Stage IIIB without malignant pleural effusion at diagnosis; Advanced = Stage IIIB with malignant pleural effusion or Stage IV at diagnosis.

Figure 1: Kaplan- Meier survival estimates by treatment group and subgroups of interest.
(A) Proportion of all patients surviving versus days from randomization by treatment group (B) Proportion of patient. (A) Proportion of patients surviving versus days from randomization by treatment group (placebo=docetaxel alone; study drug= docetaxel + TFD725) (B) Proportion of patient surviving versus days from randomization by treatment group and stage of disease at diagnosis (advdis=0: Stage IIIB; advis=1: stage IV or malignant pleural effusion) (C) Proportion of patients surviving versus days from randomization by treatment group and abnormal LDH at randomization (abnLDH=0: normal LDH; abnLDH=1: abnormal LDH) (D) Proportion of patients surviving versus days from randomization by treatment group and abnormal alkaline phosphatase level at randomization (Alkphos=0: normal alkaline phosphatase level; Alkphos=1: abnormal alkaline phosphatase level).

Comment [A38]: You talk about months elsewhere. I would have used months everywhere. 30.4 and 365 are such difficult numbers to work with.