PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Prediction of survival benefits from progression-free survival benefits in advanced non small cell lung cancer: evidence from a pooled analysis of 2,334 patients from 5 randomized trials
AUTHORS	Laporte, Silvy; Squifflet, Pierre; Baroux, Noémie; Fossella, Frank; Georgoulias, Vassilis; Pujol, Jean-Louis; Douillard, Jean-Yves; Kudoh, Shinzohy; Pignon, Jean-Pierre; Quineaux, Emmanuel; Buyse, Marc

VERSION 1 - REVIEW

REVIEWER	Avinesh Pillai Biostatistician Department of Statistics University of Auckland New Zealand No competing interests.
REVIEW RETURNED	17-Dec-2012

THE STUDY	This paper describes a meta-analysis to determine if progression free survival can be used as a surrogate for overall survival, in patients with metastatic non-small cell lung cancer treated in first line with docetaxel- based or vinorelbine-based chemotherapies. This question of using a surrogate is clinically important and it is reassuring to see that the authors have defined clearly the statistical methods used, for readers to follow if they are thinking of conducting similar analyses. The questions of interest have been clearly defined, as well as what the units of analysis were (centres or strata). The conclusion that, 'These analyses provide only modest support for considering PFS an acceptable surrogate for OS in patients with advanced NSCLC' is justified considering the limitations of the trial (page 2).
GENERAL COMMENTS	There is an extra dot in the p value on page 8 line 7, "p=.0.01". I summed up the 'N' in Table 1 and got 2,355, instead of 2,334. I see from page 5 that 2 trials were excluded but am not sure the reason for the difference of 21.

REVIEWER	Virginie Westeel, M.D. Ph. D. Chest Disease department

	jean Minjoz University Hospital Besançon, France
REVIEW RETURNED	No competing interests

GENERAL COMMENTS	Main comments: The question of the adequacy of PFS as a surrogate efficacy criteria is crucial, and the paper is well written. However, the authors, by concluding that their results should not be extrapolated to today's environment, reduce the interest of their work. It is important that they highlight why their analysis is worth being published. In this context, the choice of the trials is debatable: - as platinum-based doublets are standard treatment for EGFR wild-type advanced NSCLC, would it not have been more useful for future trials to keep only studies in which there is at least one arm receiving this type of chemotherapy regimen? - There is one study conducted in elderly patients. The relationship between PFS and OS might be different between elderly patients and their younger counterparts. This trial selection could be discussed. The correlation between PFS and OS can be influenced by subsequent treatments. Could the author give some information on second-line treatments in the included trials? Details: How do the authors explain that the difference between docetaxel-based regimens and vinorelbine-based regimens is more pronounced for OS than for PFS? Can the authors develop why different chemotherapy schedules could obscure the relationship between PFS and OS?

VERSION 1 – AUTHOR RESPONSE

Reviewer: Avinesh Pillai
Biostatistician
Department of Statistics
University of Auckland
New Zealand

This paper describes a meta-analysis to determine if progression free survival can be used as a surrogate for overall survival, in patients with metastatic non-small cell lung cancer treated in first line with docetaxel- based or vinorelbine-based chemotherapies.
This question of using a surrogate is clinically important and it is reassuring to see that the authors have defined clearly the statistical methods used, for readers to follow if they are thinking of conducting similar analyses.
The questions of interest have been clearly defined, as well as what the units of analysis were (centres or strata).

The conclusion that, 'These analyses provide only modest support for considering PFS an acceptable surrogate for OS in patients with advanced NSCLC' is justified considering the limitations of the trial (page 2).
Thank you for your positive comments.

There is an extra dot in the p value on page 8 line 7, "p=.0.01".
Corrected. Also, we changed some P-values so that they all have a leading zero before the decimal point.

I summed up the 'N' in Table 1 and got 2,355, instead of 2,334. I see from page 5 that 2 trials were excluded but am not sure the reason for the difference of 21.
The number of patients randomized was equal to 2,355, the number of patients with a valid OS was equal to 2,334 and the number of patients with a valid PFS was equal to 2,331. Hence the surrogacy analyses included 2,331 patients (99% of randomized patients). We have added a comment to this effect in the Results section on page 7. We do not think the small number of patients for whom valid OS or PFS data was unavailable may have had any sizeable impact on the analyses. All data from the 2 trials that could not be included were ignored in the tables and figures of the present paper. Details on these trials can be found in the previously published meta-analysis (Douillard et al 2007).

Reviewer: Virginie Westeel, M.D. Ph. D.
Chest Disease department
jean Minjoz University Hospital
EA 3181 Université de Franche Comté
Besançon, France

No competing interests

Main comments:
The question of the adequacy of PFS as a surrogate efficacy criteria is crucial, and the paper is well written. However, the authors, by concluding that their results should not be extrapolated to today's environment, reduce the interest of their work. It is important that they highlight why their analysis is worth being published.
Our analyses are worth being published because (1) they use a sound methodological approach, and (2) they lead to conclusions that are of interest in spite of the narrow trial selection. It would admittedly have been preferable to have a large number of trials testing a broader range of treatments, but the data currently analysed were available and had been carefully checked prior to being used in a separate meta-analysis (Douillard et al 2007). Obtaining data from more trials to extend the results presented here might be a valuable future research project. However we agree that our conclusion is probably too pessimistic compared to the conclusion we wrote in the abstract. For now, we believe the data stand on their own, and are worthy of publication. We corrected the discussion in this way (p 10).

In this context, the choice of the trials is debatable:

- as platinum-based doublets are standard treatment for EGFR wild-type advanced NSCLC, would it not have been more useful for future trials to keep only studies in which there is at least one arm receiving this type of chemotherapy regimen?
See replies above. The treatment comparisons in this set of trials were of interest at the time the trials were conducted, even though standard therapies for NSCLC have evolved since. Treatment progress, however, does not invalidate surrogacy analyses which may or may not be treatment-specific. Hence the data analysed here are useful to inform claims of surrogacy, admittedly in a historical setting (it is an often noted problem of meta-analyses that they do not generally include current data, if only because of the long follow-up required for definite analyses of overall survival).

- There is one study conducted in elderly patients. The relationship between PFS and OS might be different between elderly patients and their younger counterparts.
This question is interesting, but cannot be reliably answered with the limited dataset available here. Units of analysis only consist of strata, and randomisation was not stratified on age categories in all trials. Analyzing the data within subsets defined by age (or any other prognostic factor) might lead to results that vary simply by chance, just as subset analyses may point to misleading variability in treatment effects in clinical trials. An additional problem arises as a result of differences in selection criteria between the trials. Analyses within subsets could therefore be confounded by differences between trials. In the case of age subsets, the trial that was limited to elderly patients (WJTOG 9904) was rather small (N=181 patients, 8% of available patients) and tended to have much more extreme results than other trials. Hence the subset of elderly patients could be unduly influenced by the atypical results of trial WJOTG 9904. All in all, although the question merits further analyses, we believe our dataset is too limited in size to explore subsets reliably.

This trial selection could be discussed.
See replies above.

The correlation between PFS and OS can be influenced by subsequent treatments. Could the author give some information on second-line treatments in the included trials?
Data on post-progression treatments were not collected. This problem has been encountered by others in similar meta-analyses in advanced disease settings.

Details:
How do the authors explain that the difference between docetaxel-based regimens and vinorelbine-based regimens is more pronounced for OS than for PFS?
We have no explanation for the smaller differences on PFS than on OS, except a possible variability in the measurement techniques or schedules of PFS. Then PFS is maybe more frequently and earlier observed than OS but would be less discriminent.

Can the authors develop why different chemotherapy schedules could obscure the relationship between PFS and OS?
Since the chemotherapies used in this set of trials were relatively homogeneous (it is at once an advantage and a limitation of our meta-analysis), we have no evidence to claim that different chemotherapy schedules could lead to different relationships between PFS and OS. The question is

interesting, but can only be addresses in a much larger meta-analysis including a wide range of treatment regimens.