BMJ open

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

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2012-001802
Article Type:	Research
Date Submitted by the Author:	12-Nov-2012
Complete List of Authors:	Laporte, Silvy; University Hospital of saint-Etienne, Dept. of Clinical Pharmacology Squifflet, Pierre Baroux, Noémie; ICL, Fossella, Frank Georgoulias, Vassilis Pujol, Jean-Louis Douillard, Jean-Yves Kudoh, Shinzohy Pignon, Jean-Pierre Quineaux, Emmanuel Buyse, Marc
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Complementary medicine, Pharmacology and therapeutics
Keywords:	CLINICAL PHARMACOLOGY, Adult oncology < ONCOLOGY, STATISTICS & RESEARCH METHODS, CHEMOTHERAPY, Clinical trials < THERAPEUTICS

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

Laporte S [1,2], Squifflet P [3], Baroux N [4], Fossella F [5], Georgoulias V [6], Pujol JL [7], Douillard JY [8], Kudoh S [9], Pignon JP [10], Quinaux E [3], Buyse M [3,11]

(1) Université Jean Monnet, Saint-Etienne, France
(2) INSERM, CIE3, Saint-Etienne, France
(3) IDDI, Louvain-la-Neuve, Belgium
(4) Institut Cancérologie de la Loire, France
(5) University of Texas MD Anderson Cancer Center, Houston, TX, USA
(6) University General Hospital of Heraklion, Heraklion, Greece
(7) University Hospital of Montpellier, France
(8) ICO Centre René Gauducheau, St Herblain, France
(9) Osaka City General Hospital, Osaka, Japan
(10) Meta-Analysis Unit, Institut Gustave Roussy, Villejuif, France
(11) Center for Statistics, Hasselt University, Diepenbeek, Belgium

Address for correspondence and reprints:
Marc Buyse, ScD, IDDI (International Drug Development Institute), 30 avenue Provinciale, 1340 Ottignies Louvain-la-Neuve, Belgium,
Tel: + 32 10614444, Fax: +32 10618888, marc.buyse@iddi.com

Running head: prediction of survival benefits in lung cancer

ARTICLE SUMMARY

1) Article Focus

To investigate whether progression-free survival (PFS) can be considered a surrogate endpoint for overall survival (OS) in advanced non small cell lung cancer (NSCLC)

2) Key Messages

- Our analyses provide only modest support for considering PFS an acceptable surrogate for OS in patients with advanced NSCLC
- Only treatments that have a major impact on PFS (risk reduction of at least 50%) would be expected to also have a significant effect on OS

3) Strengths and Limitations

- Strengths : (1) analyses based on individual patient data, (2) widely accepted statistical methodology for surrogate endpoint validation
- Limitations : (1) data available on a limited number of trials, (2) results may not apply to targeted therapies

ABSTRACT

Objectives. To investigate whether progression-free survival (PFS) can be considered a surrogate endpoint for overall survival (OS) in advanced non small cell lung cancer (NSCLC).

Design. Meta-analysis of individual patient data from randomized trials

Setting/Participants. Randomized trials comparing docetaxel-based to vinorelbine-based chemotherapy for first-line treatment of NSCLC.

Primary and secondary outcome measures: Surrogacy of PFS for OS was assessed through the association between these endpoints and between the treatment effects on these endpoints. The surrogate threshold effect was the minimum treatment effect on PFS required to predict a non-zero treatment effect on OS.

Results. The median follow-up of patients still alive was 23.4 months. Median OS was 10.0 months and median PFS was 5.5 months. The treatment effects on PFS and on OS were correlated, whether using centers (R^2 = 0.62, 95% C.I. = 0.52–0.72) or prognostic strata (R^2 = 0.72, 95% C.I. = 0.60–0.84) as units of analysis. The surrogate threshold effect was a PFS hazard ratio of 0.49 using centers or 0.53 using prognostic strata.

Conclusions. These analyses provide only modest support for considering PFS an acceptable surrogate for OS in patients with advanced NSCLC. Only treatments that have a major impact on PFS (risk reduction of at least 50%) would be expected to also have a significant effect on OS. Whether these results also apply to targeted therapies is an open question that requires independent evaluation.

INTRODUCTION

A surrogate endpoint is a measure that can substitute for a "final" or "true" clinical endpoint to predict patient outcomes earlier or more conveniently than with the true endpoint. The conditions required for an endpoint to be considered a valid surrogate have been intensely studied in the recent statistical literature, and whether a surrogate can ever be "validated" is still a matter of debate today. (1- 8) Two independent conditions have proven useful to explore potential surrogate endpoints in clinical settings: one stipulates that the surrogate endpoint should predict the clinical endpoint, the other that the effect of a treatment on the surrogate endpoint should predict the effect of that treatment on the true endpoint. (9)

Overall survival (OS) remains one of the most important clinical outcomes for assessing the efficacy of cancer treatments in randomized clinical trials. However, in most cases deaths occur only after prolonged follow-up, and with the increasing number of active cancer treatments, the effect of a first-line agent on OS may be confounded by subsequent therapies. Progression-free survival (PFS), measured from randomization until objective tumor progression or death, can be assessed earlier than OS, but whether it can be considered a valid surrogate for OS depends on the malignancy and the treatment under investigation. For example, OS differences can be reliably predicted from progression-free survival (PFS) differences in advanced colorectal cancer treated with fluoropyrimidines, but not in advanced breast cancer treated with anthracyclines or taxanes. (10-11)

We investigated whether PFS is an acceptable surrogate for OS in patients with advanced non-small cell lung cancer (NSCLC) using individual data from 2,334 patients enrolled in five randomized controlled trials comparing docetaxel-based with vinorelbine-based chemotherapy as first-line treatment for advanced NSCLC.

PATIENTS AND METHODS

Trials

We analyzed data from 7 randomized controlled trials (12-18) included in a published meta-analysis of OS comparing docetaxel-based with vinca-alkaloids-based chemotherapy as first-line treatment for NSCLC. (19) Eligible trials included at least one treatment arm with either docetaxel alone or in combination with either a platinum agent (cisplatin or carboplatin) or gemcitabine and at least one vinca alkaloid-based treatment arm. Two of the seven trials included in the meta-analysis of OS could not be included in our analysis of surrogacy because the definition of PFS could not be ascertained reliably in spite of in-depth review of the case report forms. (13-14)

Table 1 provides details on the remaining five trials. The experimental arm consisted of docetaxel plus a platinum (cisplatin or carboplatin) in two trials, docetaxel plus gemcitabine in two trials, and docetaxel alone in one trial. The control arm consisted of vinorelbine plus cisplatin in four trials, and vinorelbine alone in one trial. Standard chemotherapy doses and schedules were used in the experimental and control arms.

Table 1 here

Table 2 shows the distribution of baseline patient characteristics in the 5 trials analyzed. The WGTOG trial only included patients ≥ 70 years and performance status ≥ 1. (18) The Taxobel 303 trial only included stage IV patients. (15) Approximately three-quarters of the patients were male and one-third of the tumors were squamous cell carcinomas.

Table 2 here

Data

A first meta-analysis, based on summary data extracted from the papers describing the results of the 7 trials, suggested that docetaxel-based regimens were slightly superior to vinca-alkaloid-based regimens in terms of OS for first-line therapy of advanced NSCLC.

(19) A subsequent meta-analysis confirmed these results using individual patient data from a total of 322 centers participating in the same set of 7 trials. (20) The following data were requested for the subsequent meta-analysis: patient identifier, center identifier, randomization date, treatment assigned by randomization, age, gender, body mass index, performance status, stage, overall tumor response to the first assigned treatment, date of response, date of progression with the first allocated treatment, date of death or last visit, survival status, and cause of death if applicable.

Time to Event Analyses

Progression-free survival was defined as the time from random assignment to disease progression (as assessed in each individual trial) or death from any cause. Overall survival was defined as the time from random assignment to death from any cause. The distributions of PFS and OS were estimated using the Kaplan-Meier method. Treatment groups were compared using a Cox regression model. The median follow-up time was estimated using the Kaplan-Meier method with censoring for death.

Surrogacy Analyses

A two-level modeling approach was adopted to estimate the association between PFS and OS, and between the treatment effects on these endpoints. Treatment effects were estimated as logarithms of the hazard ratio (logHR). The logHR has intuitive appeal as a measure of treatment effect: it is equal to zero in the absence of a treatment effect, and is approximately equal to the risk reduction for small treatment effects (hence a logHR of -0.10 corresponds to a risk reduction of about 10%, and a logHR of 0.10 corresponds to a risk increase of about 10%). The logHRs were estimated within units of analysis consisting of 135 centers or 64 strata. Centers were either individual centers if they had more than 3 patients per treatment arm, or groups of small centers with an average size equal to the average size of the big centers of the same trial (Tax 326: average size = 15, Taxobel 303: average size = 17, HORG: average size = 25, WJTOG 9904: average size = 14, French: average size = 20). Strata were defined within each trial by the cross-classification of the following prognostic factors: age (<60 *vs.* >= 60 years), gender (male *vs.* female), performance status (ECOG 0 or 1 *vs.* 2 or 3), BMI (<18.5 *vs.* >= 18.5 kg/m^2),

histology (squamous *vs.* non-squamous) and stage (IIIb *vs.* IV *vs.* unknown). Prognostic strata were formed as follows: a Cox model for overall survival was fit within each trial with treatment, each of the prognostic factors listed above, and the treatment-prognostic factor interaction. The prognostic factors were ordered by increasing level of significance for the treatment-prognostic factor interaction. The first prognostic factor selected in this way was the factor most predictive of the effect of treatment on overall survival, and was used to split the patients of a trial into two (or three) strata. Each of these strata was then split by the second prognostic factor; and so on. The splitting was stopped when it produced strata with less than 3 patients per treatment arm.

The association between PFS and OS was quantified through a bivariate copula model fitted on individual patient data. Kendall's τ was used to quantify the correlation between the endpoints. (21) A linear regression model was fitted on the estimated treatment effects on PFS and OS (logHRs for PFS and OS). Coefficients of determination (equal to squared correlation coefficients) were estimated using weighted linear regression. (21) Coefficients of determination (R^2) quantify the proportion of variance explained by the regression. The surrogate threshold effect was defined as the minimum treatment effect on PFS required to predict a non-zero treatment effect on OS in a future trial. (22)

RESULTS

Treatment Effects on PFS and OS

A total of 2,334 patients were included in the analysis. The median follow-up of patients still alive was 23.4 months. For the entire cohort, the median OS was 10.0 months and the median PFS was 5.5 months (Figure 1), with little difference between the curves until about 12 months.

Figure 1 here

The hazard ratios were 0.97 for PFS (95% C.I., 0.89 – 1.05, P = .44) and 0.92 for OS (95% C.I., 0.84 – 1.01, P = .089) (Figure 2). There was significant heterogeneity between the 5 trials in terms of PFS (P = .0.01) but not in terms of OS (P = .72)

Figure 2 here

Correlation between PFS and OS

PFS showed some correlation with OS (τ = 0.59; 95% C.I., 0.58 – 0.61).

Correlation Between Treatment Effects

The coefficient of determination between treatment effects estimated within centers was R^2 = 0.62 (95% C.I. = 0.52 – 0.72). The linear regression equation was logHR(OS) = -0.048 + 0.76 × logHR(PFS) (Figure 3). Using centers as the unit of analysis, the surrogate threshold effect was a PFS hazard ratio of 0.49, indicating that a risk reduction of 51% in terms of PFS would predict a non-zero effect on OS.

Figure 3 here

The coefficient of determination between treatment effects estimated within strata was R^2 = 0.88 (95% C.I. = 0.60 – 0.84). The linear regression equation was logHR(OS) = -0.071 + 0.87 × logHR(OS) (Figure 4). Using strata as the unit of analysis, the surrogate threshold effect was a PFS hazard ratio of 0.53, indicating that a risk reduction of 47% in terms of PFS would predict a non-zero effect on OS.

Figure 4 here

DISCUSSION

Our analyses suggest that PFS is not a statistically acceptable surrogate endpoint for OS in patients with metastatic non-small cell lung cancer treated in first line with docetaxel-based or vinorelbine-based chemotherapies. Indeed, although about two-thirds of the treatment effects on OS are explained by the treatment effects on PFS, the surrogate threshold effect ranges from 0.49 to 0.53 depending on whether strata or centers are used as the unit of analysis, which implies that only a major benefit of some new drug on PFS (hazard reduction of about one half or greater) would be expected to also produce a non-zero benefit on OS. These analyses are quite similar whether treatment effects are estimated in the centers participating to the trials, or in the strata defined by the characteristics of the patients most predictive of survival benefits, despite the fact that the latter analysis could have overestimated the association between the treatment effects through deliberate confounding by the prognostic factors used to define the strata.

In this set of trials, docetaxel-based regimens showed a trend towards better results than vinorelbine-based regimens, but the difference was not significant for either OS (HR = 0.92, $P = .089$) or PFS (HR = 0.97, $P = .44$). If anything, the difference was more pronounced for OS than for PFS, an unusual finding with advanced solid tumors, for which a benefit on PFS is generally diluted to yield a smaller benefit on OS. (10-11) Other meta-analyses did not support this finding. A meta-analysis comparing gemcitabine-platinum with other platinum-containing regimens found about the same benefit of gemcitabine-platinum on PFS (HR = 0.88, information available on 14 of 17 trials) as on OS (HR = 0.91 for the same 14 trials). (23) A meta-analysis of trials comparing longer with shorter durations of chemotherapy found a much more pronounced benefit of longer chemotherapy duration on PFS (HR = 0.75 on 9 of 13 trials) than on OS (HR = 0.93 for the same 9 trials). (24)

The difference between the median PFS and the median OS was only 4.5 months in this and other meta-analyses (Le Chevalier 2005), and therefore the gain in time from using PFS instead of OS in future trials of chemotherapy for advanced NSCLC would not be as

large as in other tumor types. (10-11) The short survival time post progression implies that differences in overall survival are likely to be observed for truly effective new treatments. (25) All in all, these findings suggest that even if PFS could be proposed as a plausible surrogate for OS from a statistical point of view, it would not be a very attractive one to evaluate the worth of conventional chemotherapies for advanced non-small cell lung cancer. The exclusion of two trials with unreliable PFS from our meta-analysis casts further doubts on the usefulness of this endpoint in advanced NSCLC, at least as measured a decade ago. Such exclusions also cast some doubts on meta-analyses that rely solely on published papers rather than on carefully reviewed individual patient data. (26)

Our analyses have several limitations. We used a pragmatic approach, using PFS as measured by the investigators in each trial, ignoring any possible differences in measurement techniques or schedules. While such differences may have an impact on PFS duration, they are unlikely to have much impact on the PFS hazard ratio. It is also unlikely that the results would have been much different, had a blinded central review of PFS been available in all trials. (27) The fact that treatment doses and schedules differed from trial to trial does not raise any particular concern. Indeed, such differences could have obscured (rather than enhanced) the relationship between treatment effects on PFS and on OS, hence the observed relationship is probably an underestimate of what would have been observed in a more homogeneous setting. More importantly, the randomized comparisons in this set of trials were between two standard combinations of cytotoxic drugs. The relationship between PFS and OS, and between treatment effects on PFS and OS, might not be the same for different cytostatic agents, or for targeted agents. Likewise, our results should not be extrapolated to today's environment, since more drugs with demonstrated activity in lung cancer are currently available than was the case in the trials analyzed here. Given the obvious advantages of using PFS as the primary endpoint in randomized trials, these issues deserve further investigation through further meta-analyses of contemporary randomized trials. (28)

SOURCES OF FUNDING

This work was supported by an unrestricted research grant from sanofi-aventis, Paris, France, and by the Ligue Nationale Contre le Cancer, France.

CONTRIBUTORSHIP

SL. conception and design, collection of data, data extraction, interpretation of results, revision and final approval of manuscript

PS. Data analysis, drafting the article

NB. Data extraction and monitoring

FF. collection of data, revision and final approval of manuscript

VG. collection of data, revision and final approval of manuscript

JLP. collection of data, revision and final approval of manuscript

JYD. conception, revision and final approval of manuscript

SK. collection of data, revision and final approval of manuscript

JPP. conception and design, revision and final approval of manuscript

EQ. Data analysis, drafting the article

MB. conception and design, interpretation of results, drafting and revising the article, final approval of manuscript

DATA SHARING

There is no additional data available.

REFERENCES

1. Alonso A, Molenberghs G. Evaluating time to cancer recurrence as a surrogate marker for survival from an information theory perspective. Statistical Methods in Medical Research 17: 497-504, 2008
2. Baker SG. Two simple approaches for validating a binary surrogate endpoint using data from multiple trials. Statistical Methods in Medical Research 17: 505-14, 2008
3. Burzykowski T. Surrogate endpoints: wishful thinking or reality? Statistical Methods in Medical Research 17: 463-6, 2008
4. ChakravartyA, Sridhara R. Use of progression-free survival as a surrogate marker in oncology trials: some regulatory issues. Statistical Methods in Medical Research 17: 515-18, 2008
5. Green EM, Yothers G, Sargent DJ. Surrogate endpoint validation: statistical elegance versus clinical relevance. Statistical Methods in Medical Research 17: 477-86, 2008
6. Grothey A. Surrogate endpoints for overall survival in early colorectal cancer from the clinician's perspective. Statistical Methods in Medical Research 17: 529-35, 2008
7. Hughes MD. Practical issues arising in an exploratory analysis evaluating progression-free survival as a surrogate endpoint for overall survival in advanced colorectal cancer. Statistical Methods in Medical Research 17: 487-95, 2008
8. Piedbois P, Croswell JM. Surrogate endpoints for overall survival in advanced colorectal cancer: a clinician's perspective. Statistical Methods in Medical Research 17: 519-27, 2008
9. Buyse M, Burzykowski T, Michiels S, Carroll K. Individual- and trial-level surrogacy in colorectal cancer. Statistical Methods in Medical Research 17: 467-75, 2008
10. Buyse M, Burzykowski T, Carroll K, et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. J Clin Oncol 25: 5218-24, 2007
11. Burzykowski T, Buyse M, Piccart-Gebhart MJ, et al. Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. J Clin Oncol 26: 1987-92, 2008
12. Fossella F, Pereira JR, von Pawel J, 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 21: 3016–3024, 2003

13. Monnier A, Zatloukal P, Siffnerova H, et al. Docetaxel (D) versus vinorelbine–cisplatin (VC) as front-line treatment for patients with advanced non-small cell lung cancer (NSCLC) and with a normal level of serum alpha 1 acid glycoprotein (AAG): Additional data on quality of life (QOL). Preliminary results. Proc Am Soc Clin Oncol 2003 [Abstract 2680].
14. Kubota K, Watanabe K, Kunitoh H, et al. Phase III randomized trial of docetaxel plus cisplatin versus vindesine plus cisplatin in patients with stage IV non-small-cell lung cancer: The Japanese Taxotere Lung Cancer Study Group. J Clin Oncol 22: 254–261, 2004
15. Douillard JY, Gervais R, Dabouis G, et al. Sequential two-line strategy for stage IV non-small-cell lung cancer: docetaxel–cisplatin versus vinorelbine–cisplatin followed by crossover to single-agent docetaxel or vinorelbine at progression: final results of a randomized phase II study. Ann Oncol 16: 81–89, 2005
16. Georgoulias V, Ardavanis A, Tsiafaki X, et al. Vinorelbine plus cisplatin versus docetaxel plus gemcitabine in advanced non-small-cell lung cancer: A phase III randomized trial. J Clin Oncol 23: 2937–2945, 2005
17. Pujol JL, Breton JL, Gervais R, et al. Gemcitabine–docetaxel versus cisplatin–vinorelbine in advanced or metastatic non-small-cell lung cancer: A phase III study addressing the case for cisplatin. Ann Oncol 16: 602–610, 2005
18. Kudoh S, Takeda K, Nakagawa K, et al. Phase III study of docetaxel compared with vinorelbine in elderly patients with advanced non-small-cell lung cancer: Results of the West Japan Thoracic Oncology Group Trial (WJTOG 9904). J Clin Oncol 24: 3657–3663, 2006
19. Douillard JY, Laporte S, Fossella F, et al. Comparison of docetaxel- and vinca alkaloid-based chemotherapy in the first-line treatment of advanced non-small cell lung cancer: a meta-analysis of seven randomized clinical trials. J Thorac Oncol 2: 939-46, 2007
20. Laporte S, Buyse M, Baroux N, Fossella F, et al. Meta-analysis comparing docetaxel and vinca-alkaloids in the first-line treatment of NSCLC. Comparison of results based

on individual patient data, study report data, and published data. Eur J Cancer Suppl. 1: Abstract 4, 2007

21. Burzykowski T, Molenberghs G, Buyse M, et al: Validation of surrogate endpoints in multiple randomized clinical trials with failure-time end points. J R Stat Soc C 50: 405-422, 2001
22. Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharmaceutical Stat 5: 173-186, 2006
23. Le Chevalier T, Scagliotti G, Natale R, et al. Efficacy of gemcitabine plus platinum chemotherapy compared with other platinum containing regimens in advanced non-small-cell lung cancer: a meta-analysis of survival outcomes. Lung Cancer 47: 69–80, 2005
24. Soon YY, Stockler MR, Askie LM, Boyer MJ. Duration of chemotherapy for advanced non-small-cell lung cancer: a systematic review and meta-analysis of randomized trials. J Clin Oncol 27: 3277–3283, 2009
25. Broglio KR, Berry DA. Detecting and overall survival benefit that is derived from progression-free survival. J Natl Cancer Inst 101: 1642-1649, 2009
26. Johnson KR, Ringland C, Stodkes BJ, et al. Response rate or time to progression as predictors of survival in trials of metastatic colorectal cancer or non-small-cell lung cancer: a meta-analysis. Lancet Oncology 7: 741-6, 2006
27. Amit O, Mannino F, Stone AM, et al. Blinded independent central review of progression in cancer clinical trials: Results from a meta-analysis. Eur J Cancer 47:1772-8; 2011
28. Soria JC, Massard C, Le Chevalier T. Should progression-free survival be the primary measure of efficacy for advanced NSCLC therapy? Annals Oncol 21: 2324-32, 2010

Table 1. Trials included in the surrogacy analysis (N: number of patients randomized; D, docetaxel; C, cisplatin; Cb, carboplatin; G, gemcitabine, V, vinorelbine; AUC, Area Under the Curve; HORG, Hellenic Oncology Research Group; WJTOG, West Japan Thoracic Oncology Group; q3/4 wks = every 3/4 weeks). Drug doses are indicated in mg/m² except for Cb, which was dosed to obtain an AUC of 6 mg/mL.

Trial name	First Author	Accrual period	N	Follow-up (months)	Docetaxel arm	Vinorelbine arm
Tax 326	Fossella	1998-2000	1218	21	D 75 C 75 q3 wks × 6 or D 75 Cb AUC 6 q3 wks × 6	V 25 C 100 q4 wks × 6
HORG	Georgoulias	1999-2002	413	20	D 100 G 1000 q3 wks × 6*	V 30 C 80 q3 wks × 6*
French	Pujol	1999-2001	311	25	D 85 G 1000 q3 wks × 8	V 30 C 100 q4 wks × 6
Taxobel 303	Douillard	1998-2000	233	43	D 75 C 100 q3 wks × 6	V 30 C 100 q3 wks × 6
WJTOG 9904	Kudoh	2000-2003	180	26	D 60 q3 wks × 4	V 25 q4 wks × 3

* Three additional cycles were given to patients in complete or partial response after the sixth cycle

Table 2. Baseline patient characteristics for the trials listed in Table 1 (BMI: body mass index; ECOG PS: Eastern Cooperative Oncology Group Performance Status).

Trial name	Age in years (median, range)	BMI (median, range)	Gender (Male / Female)
Tax 326	60 (23 - 87)	24.3 (15.4 – 49.6)	73% / 27%
HORG	64 (36 – 78)	Not available	89% / 11%
French	58 (37 – 75)	23.9 (15.0 – 44.1)	80% / 20%
Taxobel 303	58 (27 – 77)	23.6 (15.6 – 40.0)	82% / 18%
WJTOG 9904	76 (70 – 86)	20.5 (14.4 – 28.8)	75% / 25%

Trial name	ECOG PS (% 0 / % 1 / % 2+)	Stage (% IIIb / % IV)	Squamous cell carcinoma (% Yes / % No)
Tax 326	16% / 80% / 4%	33% / 67%	33% / 67%
HORG	44% / 46% / 10%	38% / 62%	Not available
French	21% / 71% / 8%	16% / 84%	28% / 72%
Taxobel 303	31% / 54% / 15%	0% / 100%	33% / 67%
WJTOG 9904	0% / 96% / 4%	36% / 64%	33% / 67%

LEGEND TO FIGURES

Figure 1. Kaplan-Meier estimates of overall survival and progression-free survival by treatment arm.

Figure 2. Forest plot of hazard ratios for overall survival and progression-free survival.

Figure 3. Correlation between treatment effects (log hazard ratios) on progression-free survival (horizontal axis) and overall survival (vertical axis) in participating centers. The size of each circle is proportional to the number of patients in the corresponding center. The reference circle in black corresponds to 10 patients.

Figure 4. Correlation between treatment effects (log hazard ratios) on progression-free survival (horizontal axis) and overall survival (vertical axis) in prognostic strata (see text). The size of each circle is proportional to the number of patients in the corresponding stratum. The reference circle in black corresponds to 10 patients.

Figure 1

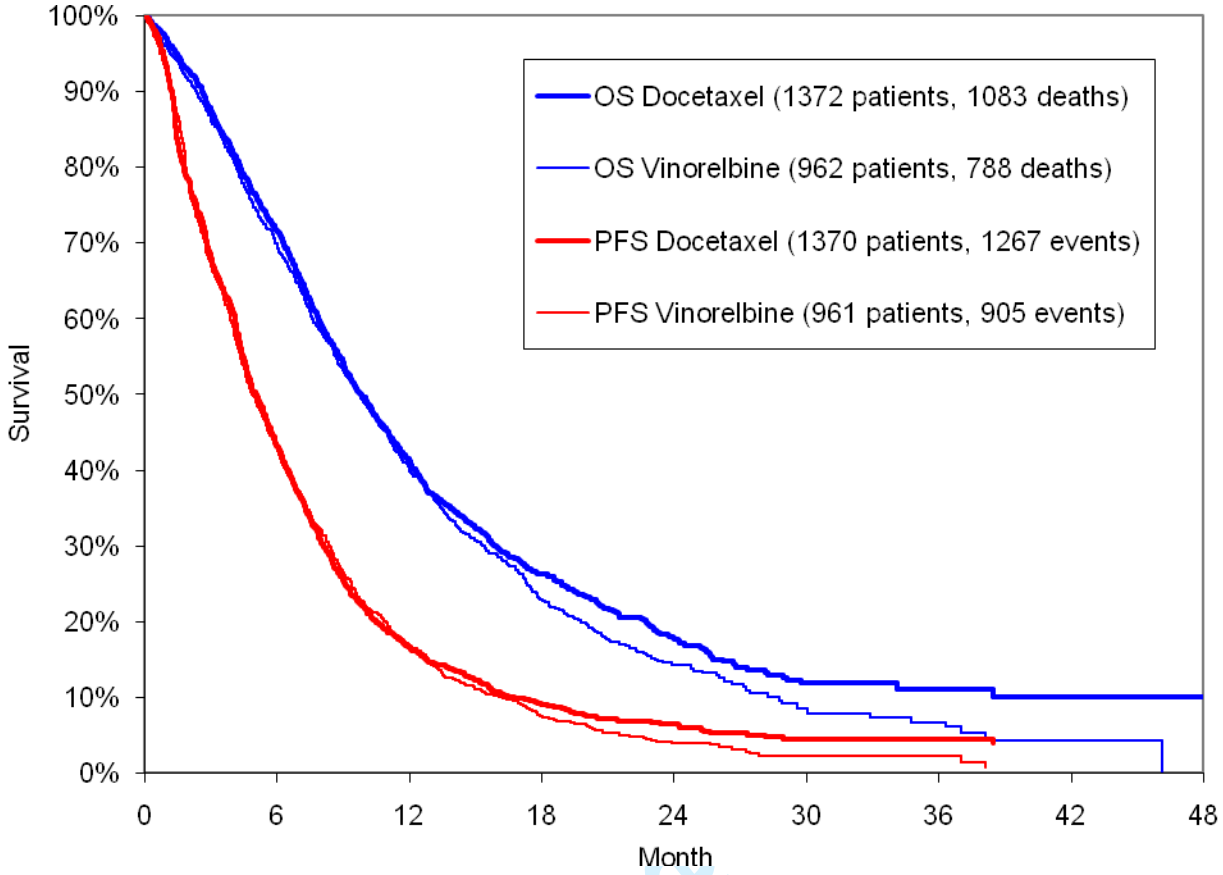


Figure 2

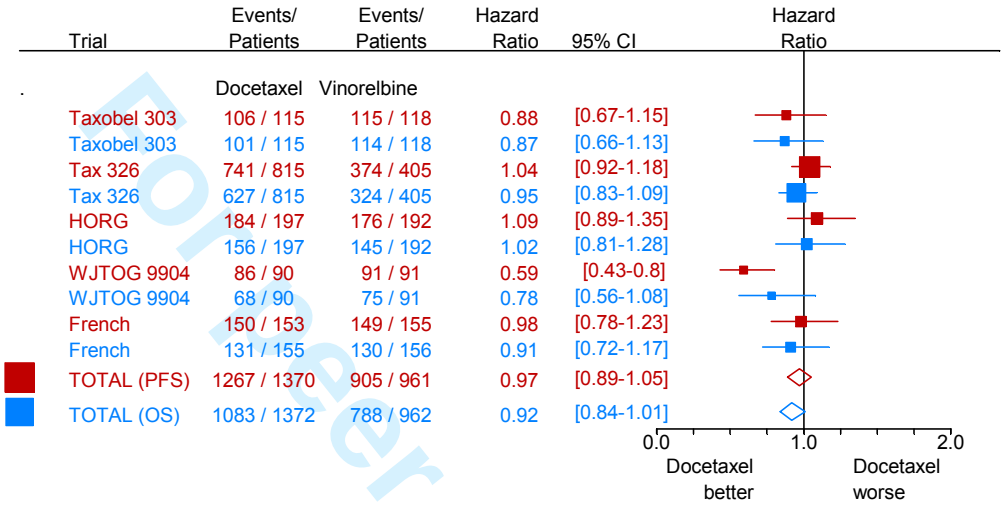


Figure 3

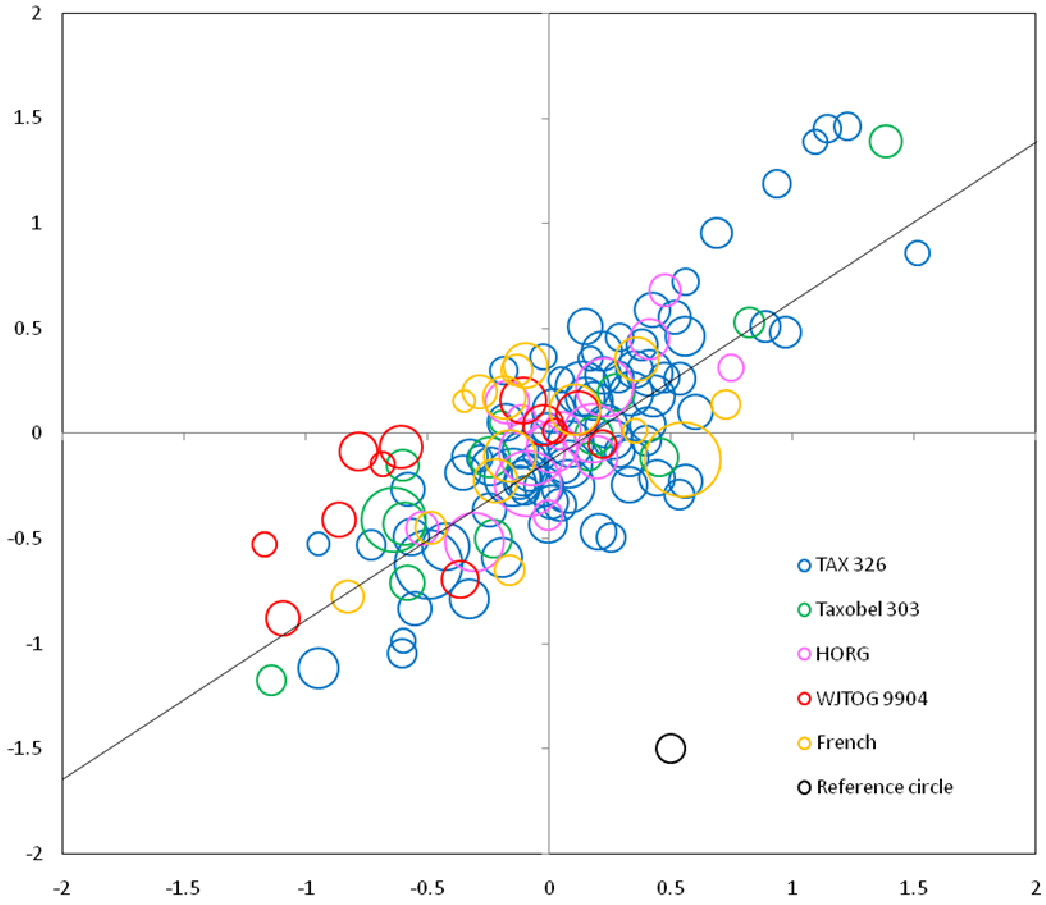
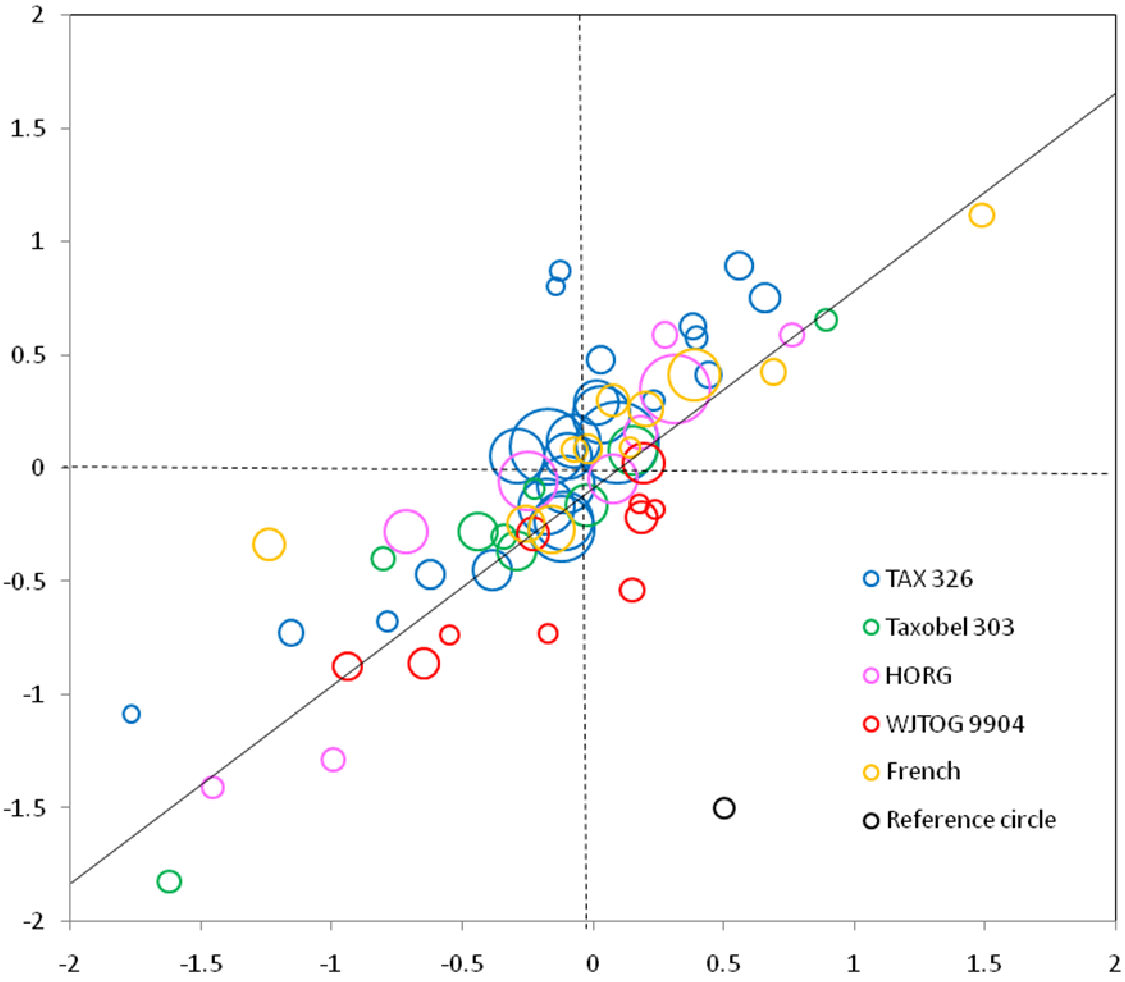


Figure 4



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

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2012-001802.R1
Article Type:	Research
Date Submitted by the Author:	30-Jan-2013
Complete List of Authors:	Laporte, Silvy; University Hospital of saint-Etienne, Dept. of Clinical Pharmacology Squifflet, Pierre; IDDI, Baroux, Noémie; ICL, Fossella, Frank; University of texas, MD Anderson Cancer Center Georgoulias, Vassilis; University General Hospital of Heraklion, Pujol, Jean-Louis; University Hospital of Montpellier, Douillard, Jean-Yves; ICO Centre René Gauducheau, Kudoh, Shinzohy; Osaka City General Hospital, Pignon, Jean-Pierre; Institut Gustave Roussy, Meta-Analysis Unit Quineaux, Emmanuel; IDDI, Buyse, Marc
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Complementary medicine, Pharmacology and therapeutics
Keywords:	CLINICAL PHARMACOLOGY, Breast tumours < ONCOLOGY, Clinical trials < THERAPEUTICS

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

Laporte S [1,2], Squifflet P [3], Baroux N [4], Fossella F [5], Georgoulias V [6], Pujol JL [7], Douillard JY [8], Kudoh S [9], Pignon JP [10], Quinaux E [3], Buyse M [3,11]

(1) Université Jean Monnet, Saint-Etienne, France
(2) INSERM, CIE3, Saint-Etienne, France
(3) IDDI, Louvain-la-Neuve, Belgium
(4) Institut Cancérologie de la Loire, France
(5) University of Texas MD Anderson Cancer Center, Houston, TX, USA
(6) University General Hospital of Heraklion, Heraklion, Greece
(7) University Hospital of Montpellier, France
(8) ICO Centre René Gauducheau, St Herblain, France
(9) Osaka City General Hospital, Osaka, Japan
(10) Meta-Analysis Unit, Institut Gustave Roussy, Villejuif, France
(11) Center for Statistics, Hasselt University, Diepenbeek, Belgium

Address for correspondence and reprints:
Marc Buyse, ScD, IDDI (International Drug Development Institute), 30 avenue Provinciale, 1340 Ottignies Louvain-la-Neuve, Belgium,
Tel: + 32 10614444, Fax: +32 10618888, marc.buyse@iddi.com

Running head: prediction of survival benefits in lung cancer

ARTICLE SUMMARY

1) Article Focus

To investigate whether progression-free survival (PFS) can be considered a surrogate endpoint for overall survival (OS) in advanced non small cell lung cancer (NSCLC)

2) Key Messages

- Our analyses provide only modest support for considering PFS an acceptable surrogate for OS in patients with advanced NSCLC
- Only treatments that have a major impact on PFS (risk reduction of at least 50%) would be expected to also have a significant effect on OS

3) Strengths and Limitations

- Strengths : (1) analyses based on individual patient data, (2) widely accepted statistical methodology for surrogate endpoint validation
- Limitations : (1) data available on a limited number of trials, (2) results may not apply to targeted therapies

ABSTRACT

Objectives. To investigate whether progression-free survival (PFS) can be considered a surrogate endpoint for overall survival (OS) in advanced non small cell lung cancer (NSCLC).

Design. Meta-analysis of individual patient data from randomized trials

Setting/Participants. Randomized trials comparing docetaxel-based to vinorelbine-based chemotherapy for first-line treatment of NSCLC.

Primary and secondary outcome measures: Surrogacy of PFS for OS was assessed through the association between these endpoints and between the treatment effects on these endpoints. The surrogate threshold effect was the minimum treatment effect on PFS required to predict a non-zero treatment effect on OS.

Results. The median follow-up of patients still alive was 23.4 months. Median OS was 10.0 months and median PFS was 5.5 months. The treatment effects on PFS and on OS were correlated, whether using centers (R^2 = 0.62, 95% C.I. = 0.52–0.72) or prognostic strata (R^2 = 0.72, 95% C.I. = 0.60–0.84) as units of analysis. The surrogate threshold effect was a PFS hazard ratio of 0.49 using centers or 0.53 using prognostic strata.

Conclusions. These analyses provide only modest support for considering PFS an acceptable surrogate for OS in patients with advanced NSCLC. Only treatments that have a major impact on PFS (risk reduction of at least 50%) would be expected to also have a significant effect on OS. Whether these results also apply to targeted therapies is an open question that requires independent evaluation.

INTRODUCTION

A surrogate endpoint is a measure that can substitute for a "final" or "true" clinical endpoint to predict patient outcomes earlier or more conveniently than with the true endpoint. The conditions required for an endpoint to be considered a valid surrogate have been intensely studied in the recent statistical literature, and whether a surrogate can ever be "validated" is still a matter of debate today. (1- 8) Two independent conditions have proven useful to explore potential surrogate endpoints in clinical settings: one stipulates that the surrogate endpoint should predict the clinical endpoint, the other that the effect of a treatment on the surrogate endpoint should predict the effect of that treatment on the true endpoint. (9)

Overall survival (OS) remains one of the most important clinical outcomes for assessing the efficacy of cancer treatments in randomized clinical trials. However, in most cases deaths occur only after prolonged follow-up, and with the increasing number of active cancer treatments, the effect of a first-line agent on OS may be confounded by subsequent therapies. Progression-free survival (PFS), measured from randomization until objective tumor progression or death, can be assessed earlier than OS, but whether it can be considered a valid surrogate for OS depends on the malignancy and the treatment under investigation. For example, OS differences can be reliably predicted from progression-free survival (PFS) differences in advanced colorectal cancer treated with fluoropyrimidines, but not in advanced breast cancer treated with anthracyclines or taxanes. (10-11)

We investigated whether PFS is an acceptable surrogate for OS in patients with advanced non-small cell lung cancer (NSCLC) using individual data from 2,334 patients enrolled in five randomized controlled trials comparing docetaxel-based with vinorelbine-based chemotherapy as first-line treatment for advanced NSCLC.

PATIENTS AND METHODS

Trials

We analyzed data from 7 randomized controlled trials (12-18) included in a published meta-analysis of OS comparing docetaxel-based with vinca-alkaloids-based chemotherapy as first-line treatment for NSCLC. (19) Eligible trials included at least one treatment arm with either docetaxel alone or in combination with either a platinum agent (cisplatin or carboplatin) or gemcitabine and at least one vinca alkaloid-based treatment arm. Two of the seven trials included in the meta-analysis of OS could not be included in our analysis of surrogacy because the definition of PFS could not be ascertained reliably in spite of in-depth review of the case report forms. (13-14)

Table 1 provides details on the remaining five trials. The experimental arm consisted of docetaxel plus a platinum (cisplatin or carboplatin) in two trials, docetaxel plus gemcitabine in two trials, and docetaxel alone in one trial. The control arm consisted of vinorelbine plus cisplatin in four trials, and vinorelbine alone in one trial. Standard chemotherapy doses and schedules were used in the experimental and control arms.

Table 1 here

Table 2 shows the distribution of baseline patient characteristics in the 5 trials analyzed. The WGTOG trial only included patients ≥ 70 years and performance status ≥ 1. (18) The Taxobel 303 trial only included stage IV patients. (15) Approximately three-quarters of the patients were male and one-third of the tumors were squamous cell carcinomas.

Table 2 here

Data

A first meta-analysis, based on summary data extracted from the papers describing the results of the 7 trials, suggested that docetaxel-based regimens were slightly superior to vinca-alkaloid-based regimens in terms of OS for first-line therapy of advanced NSCLC.

(19) A subsequent meta-analysis confirmed these results using individual patient data from a total of 322 centers participating in the same set of 7 trials. (20) The following data were requested for the subsequent meta-analysis: patient identifier, center identifier, randomization date, treatment assigned by randomization, age, gender, body mass index, performance status, stage, overall tumor response to the first assigned treatment, date of response, date of progression with the first allocated treatment, date of death or last visit, survival status, and cause of death if applicable.

Time to Event Analyses

Progression-free survival was defined as the time from random assignment to disease progression (as assessed in each individual trial) or death from any cause. Overall survival was defined as the time from random assignment to death from any cause. The distributions of PFS and OS were estimated using the Kaplan-Meier method. Treatment groups were compared using a Cox regression model. The median follow-up time was estimated using the Kaplan-Meier method with censoring for death.

Surrogacy Analyses

A two-level modeling approach was adopted to estimate the association between PFS and OS, and between the treatment effects on these endpoints. Treatment effects were estimated as logarithms of the hazard ratio (logHR). The logHR has intuitive appeal as a measure of treatment effect: it is equal to zero in the absence of a treatment effect, and is approximately equal to the risk reduction for small treatment effects (hence a logHR of -0.10 corresponds to a risk reduction of about 10%, and a logHR of 0.10 corresponds to a risk increase of about 10%). The logHRs were estimated within units of analysis consisting of 135 centers or 64 strata. Centers were either individual centers if they had more than 3 patients per treatment arm, or groups of small centers with an average size at least equal to the average size of the big centers of the same trial (Tax 326: average size = 15, Taxobel 303: average size = 17, HORG: average size = 25, WJTOG 9904: average size = 14, French: average size = 20). Strata were defined within each trial by the cross-classification of the following prognostic factors: age (<60 *vs.* >= 60 years), gender (male *vs.* female), performance status (ECOG 0 or 1 *vs.* 2 or 3), BMI (<18.5 *vs.* >= 18.5 kg/m^2),

histology (squamous *vs.* non-squamous) and stage (IIIb *vs.* IV *vs.* unknown). Prognostic strata were formed as follows: a Cox model for overall survival was fit within each trial with treatment, each of the prognostic factors listed above, and the treatment-prognostic factor interaction. The prognostic factors were ordered by increasing level of significance for the treatment-prognostic factor interaction. The first prognostic factor selected in this way was the factor most predictive of the effect of treatment on overall survival, and was used to split the patients of a trial into two (or three) strata. Each of these strata was then split by the second prognostic factor; and so on. The splitting was stopped when it produced strata with less than 3 patients per treatment arm.

The association between PFS and OS was quantified through a bivariate copula model fitted on individual patient data. Kendall's τ was used to quantify the correlation between the endpoints. (21) A linear regression model was fitted on the estimated treatment effects on PFS and OS (logHRs for PFS and OS). Coefficients of determination (equal to squared correlation coefficients) were estimated using weighted linear regression. (21) Coefficients of determination (R^2) quantify the proportion of variance explained by the regression. The surrogate threshold effect was defined as the minimum treatment effect on PFS required to predict a non-zero treatment effect on OS in a future trial. (22)

RESULTS

Treatment Effects on PFS and OS

A total of 2,331 patients were included in the analysis, since the PFS and/or OS was missing for 24 (1%) of all patients. The median follow-up of patients still alive was 23.4 months. For the entire cohort, the median OS was 10.0 months and the median PFS was 5.5 months (Figure 1), with little difference between the curves until about 12 months.

Figure 1 here

The hazard ratios were 0.97 for PFS (95% C.I., 0.89 – 1.05, $P = 0.44$) and 0.92 for OS (95% C.I., 0.84 – 1.01, $P = 0.089$) (Figure 2). There was significant heterogeneity between the 5 trials in terms of PFS ($P = 0.01$) but not in terms of OS ($P = 0.72$)

Figure 2 here

Correlation between PFS and OS

PFS showed some correlation with OS ($\tau = 0.59$; 95% C.I., 0.58 – 0.61).

Correlation Between Treatment Effects

The coefficient of determination between treatment effects estimated within centers was $R^2 = 0.62$ (95% C.I. = 0.52 – 0.72). The linear regression equation was logHR(OS) = -0.048 + 0.76 × logHR(PFS) (Figure 3). Using centers as the unit of analysis, the surrogate threshold effect was a PFS hazard ratio of 0.49, indicating that a risk reduction of 51% in terms of PFS would predict a non-zero effect on OS.

Figure 3 here

The coefficient of determination between treatment effects estimated within strata was $R^2 = 0.88$ (95% C.I. = 0.60 – 0.84). The linear regression equation was logHR(OS) = -0.071 + 0.87 × logHR(OS) (Figure 4). Using strata as the unit of analysis, the surrogate threshold effect was a PFS hazard ratio of 0.53, indicating that a risk reduction of 47% in terms of PFS would predict a non-zero effect on OS.

Figure 4 here

DISCUSSION

Our analyses suggest that PFS is not a statistically acceptable surrogate endpoint for OS in patients with metastatic non-small cell lung cancer treated in first line with docetaxel-based or vinorelbine-based chemotherapies. Indeed, although about two-thirds of the treatment effects on OS are explained by the treatment effects on PFS, the surrogate threshold effect ranges from 0.49 to 0.53 depending on whether strata or centers are used as the unit of analysis, which implies that only a major benefit of some new drug on PFS (hazard reduction of about one half or greater) would be expected to also produce a non-zero benefit on OS. These analyses are quite similar whether treatment effects are estimated in the centers participating to the trials, or in the strata defined by the characteristics of the patients most predictive of survival benefits, despite the fact that the latter analysis could have overestimated the association between the treatment effects through deliberate confounding by the prognostic factors used to define the strata.

In this set of trials, docetaxel-based regimens showed a trend towards better results than vinorelbine-based regimens, but the difference was not significant for either OS (HR = 0.92, $P = 0.089$) or PFS (HR = 0.97, $P = 0.44$). If anything, the difference was more pronounced for OS than for PFS, an unusual finding with advanced solid tumors, for which a benefit on PFS is generally diluted to yield a smaller benefit on OS. (10-11) Other meta-analyses did not support this finding. A meta-analysis comparing gemcitabine-platinum with other platinum-containing regimens found about the same benefit of gemcitabine-platinum on PFS (HR = 0.88, information available on 14 of 17 trials) as on OS (HR = 0.91 for the same 14 trials). (23) A meta-analysis of trials comparing longer with shorter durations of chemotherapy found a much more pronounced benefit of longer chemotherapy duration on PFS (HR = 0.75 on 9 of 13 trials) than on OS (HR = 0.93 for the same 9 trials). (24)

The difference between the median PFS and the median OS was only 4.5 months in this and other meta-analyses (Le Chevalier 2005), and therefore the gain in time from using PFS instead of OS in future trials of chemotherapy for advanced NSCLC would not be as

large as in other tumor types. (10-11) The short survival time post progression implies that differences in overall survival are likely to be observed for truly effective new treatments. (25) All in all, these findings suggest that even if PFS could be proposed as a plausible surrogate for OS from a statistical point of view, it would not be a very attractive one to evaluate the worth of conventional chemotherapies for advanced non-small cell lung cancer. The exclusion of two trials with unreliable PFS from our meta-analysis casts further doubts on the usefulness of this endpoint in advanced NSCLC, at least as measured a decade ago. Such exclusions also cast some doubts on meta-analyses that rely solely on published papers rather than on carefully reviewed individual patient data. (26)

Our analyses have several limitations. We used a pragmatic approach, using PFS as measured by the investigators in each trial, ignoring any possible differences in measurement techniques or schedules. While such differences may have an impact on PFS duration, they are unlikely to have much impact on the PFS hazard ratio. It is also unlikely that the results would have been much different, had a blinded central review of PFS been available in all trials. (27) The fact that treatment doses and schedules differed from trial to trial does not raise any particular concern. Indeed, such differences could have obscured (rather than enhanced) the relationship between treatment effects on PFS and on OS, hence the observed relationship is probably an underestimate of what would have been observed in a more homogeneous setting. More importantly, the randomized comparisons in this set of trials were between two standard combinations of cytotoxic drugs. Although these analyses provide modest support for considering PFS an acceptable surrogate for OS in patients with advanced NSCLC, treatments that have a major impact on PFS would be expected to also have a significant effect on OS. The relationship between PFS and OS, and between treatment effects on PFS and OS, which might not be the same for different cytostatic agents or for targeted agents, and given the obvious advantages of using PFS as the primary endpoint in randomized trials, these issues deserve further investigation through further meta-analyses of contemporary randomized trials. (28)

SOURCES OF FUNDING

This work was supported by an unrestricted research grant from sanofi-aventis, Paris, France, and by the Ligue Nationale Contre le Cancer, France.

DATA SHARING

No additional data available.

COMPETING INTERESTS

I will upload an ICMJE conflicts of interest form for each author of this manuscript.

CONTRIBUTORSHIP

SL. conception and design, collection of data, data extraction, interpretation of results, revision and final approval of manuscript

PS. Data analysis, drafting the article

NB. Data extraction and monitoring

FF. collection of data, revision and final approval of manuscript

VG. collection of data, revision and final approval of manuscript

JLP. collection of data, revision and final approval of manuscript

JYD. conception, revision and final approval of manuscript

SK. collection of data, revision and final approval of manuscript

JPP. conception and design, revision and final approval of manuscript

EQ. Data analysis, drafting the article

MB. conception and design, interpretation of results, drafting and revising the article, final approval of manuscript

REFERENCES

1. Alonso A, Molenberghs G. Evaluating time to cancer recurrence as a surrogate marker for survival from an information theory perspective. Statistical Methods in Medical Research 17: 497-504, 2008
2. Baker SG. Two simple approaches for validating a binary surrogate endpoint using data from multiple trials. Statistical Methods in Medical Research 17: 505-14, 2008
3. Burzykowski T. Surrogate endpoints: wishful thinking or reality? Statistical Methods in Medical Research 17: 463-6, 2008
4. ChakravartyA, Sridhara R. Use of progression-free survival as a surrogate marker in oncology trials: some regulatory issues. Statistical Methods in Medical Research 17: 515-18, 2008
5. Green EM, Yothers G, Sargent DJ. Surrogate endpoint validation: statistical elegance versus clinical relevance. Statistical Methods in Medical Research 17: 477-86, 2008
6. Grothey A. Surrogate endpoints for overall survival in early colorectal cancer from the clinician's perspective. Statistical Methods in Medical Research 17: 529-35, 2008
7. Hughes MD. Practical issues arising in an exploratory analysis evaluating progression-free survival as a surrogate endpoint for overall survival in advanced colorectal cancer. Statistical Methods in Medical Research 17: 487-95, 2008
8. Piedbois P, Croswell JM. Surrogate endpoints for overall survival in advanced colorectal cancer: a clinician's perspective. Statistical Methods in Medical Research 17: 519-27, 2008
9. Buyse M, Burzykowski T, Michiels S, et al. Individual- and trial-level surrogacy in colorectal cancer. Statistical Methods in Medical Research 17: 467-75, 2008
10. Buyse M, Burzykowski T, Carroll K, et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. J Clin Oncol 25: 5218-24, 2007
11. Burzykowski T, Buyse M, Piccart-Gebhart MJ, et al. Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. J Clin Oncol 26: 1987-92, 2008
12. Fossella F, Pereira JR, von Pawel J, 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 21: 3016–3024, 2003

13. Monnier A, Zatloukal P, Siffnerova H, et al. Docetaxel (D) versus vinorelbine–cisplatin (VC) as front-line treatment for patients with advanced non-small cell lung cancer (NSCLC) and with a normal level of serum alpha 1 acid glycoprotein (AAG): Additional data on quality of life (QOL). Preliminary results. Proc Am Soc Clin Oncol 2003 [Abstract 2680].
14. Kubota K, Watanabe K, Kunitoh H, et al. Phase III randomized trial of docetaxel plus cisplatin versus vindesine plus cisplatin in patients with stage IV non-small-cell lung cancer: The Japanese Taxotere Lung Cancer Study Group. J Clin Oncol 22: 254–261, 2004
15. Douillard JY, Gervais R, Dabouis G, et al. Sequential two-line strategy for stage IV non-small-cell lung cancer: docetaxel–cisplatin versus vinorelbine–cisplatin followed by crossover to single-agent docetaxel or vinorelbine at progression: final results of a randomized phase II study. Ann Oncol 16: 81–89, 2005
16. Georgoulias V, Ardavanis A, Tsiafaki X, et al. Vinorelbine plus cisplatin versus docetaxel plus gemcitabine in advanced non-small-cell lung cancer: A phase III randomized trial. J Clin Oncol 23: 2937–2945, 2005
17. Pujol JL, Breton JL, Gervais R, et al. Gemcitabine–docetaxel versus cisplatin–vinorelbine in advanced or metastatic non-small-cell lung cancer: A phase III study addressing the case for cisplatin. Ann Oncol 16: 602–610, 2005
18. Kudoh S, Takeda K, Nakagawa K, et al. Phase III study of docetaxel compared with vinorelbine in elderly patients with advanced non-small-cell lung cancer: Results of the West Japan Thoracic Oncology Group Trial (WJTOG 9904). J Clin Oncol 24: 3657–3663, 2006
19. Douillard JY, Laporte S, Fossella F, et al. Comparison of docetaxel- and vinca alkaloid-based chemotherapy in the first-line treatment of advanced non-small cell lung cancer: a meta-analysis of seven randomized clinical trials. J Thorac Oncol 2: 939-46, 2007
20. Laporte S, Buyse M, Baroux N, et al. Meta-analysis comparing docetaxel and vinca-alkaloids in the first-line treatment of NSCLC. Comparison of results based on

individual patient data, study report data, and published data. Eur J Cancer Suppl. 1: Abstract 4, 2007

21. Burzykowski T, Molenberghs G, Buyse M, et al: Validation of surrogate endpoints in multiple randomized clinical trials with failure-time end points. J R Stat Soc C 50: 405-422, 2001
22. Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharmaceutical Stat 5: 173-186, 2006
23. Le Chevalier T, Scagliotti G, Natale R, et al. Efficacy of gemcitabine plus platinum chemotherapy compared with other platinum containing regimens in advanced non-small-cell lung cancer: a meta-analysis of survival outcomes. Lung Cancer 47: 69–80, 2005
24. Soon YY, Stockler MR, Askie LM, et al. Duration of chemotherapy for advanced non-small-cell lung cancer: a systematic review and meta-analysis of randomized trials. J Clin Oncol 27: 3277–3283, 2009
25. Broglio KR, Berry DA. Detecting and overall survival benefit that is derived from progression-free survival. J Natl Cancer Inst 101: 1642-1649, 2009
26. Johnson KR, Ringland C, Stodkes BJ, et al. Response rate or time to progression as predictors of survival in trials of metastatic colorectal cancer or non-small-cell lung cancer: a meta-analysis. Lancet Oncology 7: 741-6, 2006
27. Amit O, Mannino F, Stone AM, et al. Blinded independent central review of progression in cancer clinical trials: Results from a meta-analysis. Eur J Cancer 47:1772-8; 2011
28. Soria JC, Massard C, Le Chevalier T. Should progression-free survival be the primary measure of efficacy for advanced NSCLC therapy? Annals Oncol 21: 2324-32, 2010

Table 1. Trials included in the surrogacy analysis (N: number of patients randomized; D, docetaxel; C, cisplatin; Cb, carboplatin; G, gemcitabine, V, vinorelbine; AUC, Area Under the Curve; HORG, Hellenic Oncology Research Group; WJTOG, West Japan Thoracic Oncology Group; q3/4 wks = every 3/4 weeks). Drug doses are indicated in mg/m² except for Cb, which was dosed to obtain an AUC of 6 mg/mL.

Trial name	First Author	Accrual period	N	Follow-up (months)	Docetaxel arm	Vinorelbine arm
Tax 326	Fossella	1998-2000	1218	21	D 75 C 75 q3 wks × 6 or D 75 Cb AUC 6 q3 wks × 6	V 25 C 100 q4 wks × 6
HORG	Georgoulias	1999-2002	413	20	D 100 G 1000 q3 wks × 6*	V 30 C 80 q3 wks × 6*
French	Pujol	1999-2001	311	25	D 85 G 1000 q3 wks × 8	V 30 C 100 q4 wks × 6
Taxobel 303	Douillard	1998-2000	233	43	D 75 C 100 q3 wks × 6	V 30 C 100 q3 wks × 6
WJTOG 9904	Kudoh	2000-2003	180	26	D 60 q3 wks × 4	V 25 q4 wks × 3

* Three additional cycles were given to patients in complete or partial response after the sixth cycle

Table 2. Baseline patient characteristics for the trials listed in Table 1 (BMI: body mass index; ECOG PS: Eastern Cooperative Oncology Group Performance Status).

Trial name	Age in years (median, range)	BMI (median, range)	Gender (Male / Female)
Tax 326	60 (23 - 87)	24.3 (15.4 – 49.6)	73% / 27%
HORG	64 (36 – 78)	Not available	89% / 11%
French	58 (37 – 75)	23.9 (15.0 – 44.1)	80% / 20%
Taxobel 303	58 (27 – 77)	23.6 (15.6 – 40.0)	82% / 18%
WJTOG 9904	76 (70 – 86)	20.5 (14.4 – 28.8)	75% / 25%

Trial name	ECOG PS (% 0 / % 1 / % 2+)	Stage (% IIIb / % IV)	Squamous cell carcinoma (% Yes / % No)
Tax 326	16% / 80% / 4%	33% / 67%	33% / 67%
HORG	44% / 46% / 10%	38% / 62%	Not available
French	21% / 71% / 8%	16% / 84%	28% / 72%
Taxobel 303	31% / 54% / 15%	0% / 100%	33% / 67%
WJTOG 9904	0% / 96% / 4%	36% / 64%	33% / 67%

LEGEND TO FIGURES

Figure 1. Kaplan-Meier estimates of overall survival and progression-free survival by treatment arm.

Figure 2. Forest plot of hazard ratios for overall survival and progression-free survival.

Figure 3. Correlation between treatment effects (log hazard ratios) on progression-free survival (horizontal axis) and overall survival (vertical axis) in participating centers. The size of each circle is proportional to the number of patients in the corresponding center. The reference circle in black corresponds to 10 patients.

Figure 4. Correlation between treatment effects (log hazard ratios) on progression-free survival (horizontal axis) and overall survival (vertical axis) in prognostic strata (see text). The size of each circle is proportional to the number of patients in the corresponding stratum. The reference circle in black corresponds to 10 patients.

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

Laporte S [1,2], Squifflet P [3], Baroux N [4], Fossella F [5], Georgoulias V [6], Pujol JL [7], Douillard JY [8], Kudoh S [9], Pignon JP [10], Quinaux E [3], Buyse M [3,11]

(1) Université Jean Monnet, Saint-Etienne, France
(2) INSERM, CIE3, Saint-Etienne, France
(3) IDDI, Louvain-la-Neuve, Belgium
(4) Institut Cancérologie de la Loire, France
(5) University of Texas MD Anderson Cancer Center, Houston, TX, USA
(6) University General Hospital of Heraklion, Heraklion, Greece
(7) University Hospital of Montpellier, France
(8) ICO Centre René Gauducheau, St Herblain, France
(9) Osaka City General Hospital, Osaka, Japan
(10) Meta-Analysis Unit, Institut Gustave Roussy, Villejuif, France
(11) Center for Statistics, Hasselt University, Diepenbeek, Belgium

Address for correspondence and reprints:
Marc Buyse, ScD, IDDI (International Drug Development Institute), 30 avenue Provinciale, 1340 Ottignies Louvain-la-Neuve, Belgium,
Tel: + 32 10614444, Fax: +32 10618888, marc.buyse@iddi.com

Running head: prediction of survival benefits in lung cancer

ARTICLE SUMMARY

1) Article Focus

To investigate whether progression-free survival (PFS) can be considered a surrogate endpoint for overall survival (OS) in advanced non small cell lung cancer (NSCLC)

2) Key Messages

- Our analyses provide only modest support for considering PFS an acceptable surrogate for OS in patients with advanced NSCLC
- Only treatments that have a major impact on PFS (risk reduction of at least 50%) would be expected to also have a significant effect on OS

3) Strengths and Limitations

- Strengths : (1) analyses based on individual patient data, (2) widely accepted statistical methodology for surrogate endpoint validation
- Limitations : (1) data available on a limited number of trials, (2) results may not apply to targeted therapies

ABSTRACT

Objectives. To investigate whether progression-free survival (PFS) can be considered a surrogate endpoint for overall survival (OS) in advanced non small cell lung cancer (NSCLC).

Design. Meta-analysis of individual patient data from randomized trials

Setting/Participants. Randomized trials comparing docetaxel-based to vinorelbine-based chemotherapy for first-line treatment of NSCLC.

Primary and secondary outcome measures: Surrogacy of PFS for OS was assessed through the association between these endpoints and between the treatment effects on these endpoints. The surrogate threshold effect was the minimum treatment effect on PFS required to predict a non-zero treatment effect on OS.

Results. The median follow-up of patients still alive was 23.4 months. Median OS was 10.0 months and median PFS was 5.5 months. The treatment effects on PFS and on OS were correlated, whether using centers (R^2 = 0.62, 95% C.I. = 0.52–0.72) or prognostic strata (R^2 = 0.72, 95% C.I. = 0.60–0.84) as units of analysis. The surrogate threshold effect was a PFS hazard ratio of 0.49 using centers or 0.53 using prognostic strata.

Conclusions. These analyses provide only modest support for considering PFS an acceptable surrogate for OS in patients with advanced NSCLC. Only treatments that have a major impact on PFS (risk reduction of at least 50%) would be expected to also have a significant effect on OS. Whether these results also apply to targeted therapies is an open question that requires independent evaluation.

There is no additional data available.

INTRODUCTION

A surrogate endpoint is a measure that can substitute for a "final" or "true" clinical endpoint to predict patient outcomes earlier or more conveniently than with the true endpoint. The conditions required for an endpoint to be considered a valid surrogate have been intensely studied in the recent statistical literature, and whether a surrogate can ever be "validated" is still a matter of debate today. (1- 8) Two independent conditions have proven useful to explore potential surrogate endpoints in clinical settings: one stipulates that the surrogate endpoint should predict the clinical endpoint, the other that the effect of a treatment on the surrogate endpoint should predict the effect of that treatment on the true endpoint. (9)

Overall survival (OS) remains one of the most important clinical outcomes for assessing the efficacy of cancer treatments in randomized clinical trials. However, in most cases deaths occur only after prolonged follow-up, and with the increasing number of active cancer treatments, the effect of a first-line agent on OS may be confounded by subsequent therapies. Progression-free survival (PFS), measured from randomization until objective tumor progression or death, can be assessed earlier than OS, but whether it can be considered a valid surrogate for OS depends on the malignancy and the treatment under investigation. For example, OS differences can be reliably predicted from progression-free survival (PFS) differences in advanced colorectal cancer treated with fluoropyrimidines, but not in advanced breast cancer treated with anthracyclines or taxanes. (10-11)

We investigated whether PFS is an acceptable surrogate for OS in patients with advanced non-small cell lung cancer (NSCLC) using individual data from 2,334 patients enrolled in five randomized controlled trials comparing docetaxel-based with vinorelbine-based chemotherapy as first-line treatment for advanced NSCLC.

PATIENTS AND METHODS

Trials

We analyzed data from 7 randomized controlled trials (12-18) included in a published meta-analysis of OS comparing docetaxel-based with vinca-alkaloids-based chemotherapy as first-line treatment for NSCLC. (19) Eligible trials included at least one treatment arm with either docetaxel alone or in combination with either a platinum agent (cisplatin or carboplatin) or gemcitabine and at least one vinca alkaloid-based treatment arm. Two of the seven trials included in the meta-analysis of OS could not be included in our analysis of surrogacy because the definition of PFS could not be ascertained reliably in spite of in-depth review of the case report forms. (13-14)

Table 1 provides details on the remaining five trials. The experimental arm consisted of docetaxel plus a platinum (cisplatin or carboplatin) in two trials, docetaxel plus gemcitabine in two trials, and docetaxel alone in one trial. The control arm consisted of vinorelbine plus cisplatin in four trials, and vinorelbine alone in one trial. Standard chemotherapy doses and schedules were used in the experimental and control arms.

Table 1 here

Table 2 shows the distribution of baseline patient characteristics in the 5 trials analyzed. The WGTOG trial only included patients ≥ 70 years and performance status ≥ 1. (18) The Taxobel 303 trial only included stage IV patients. (15) Approximately three-quarters of the patients were male and one-third of the tumors were squamous cell carcinomas.

Table 2 here

Data

A first meta-analysis, based on summary data extracted from the papers describing the results of the 7 trials, suggested that docetaxel-based regimens were slightly superior to vinca-alkaloid-based regimens in terms of OS for first-line therapy of advanced NSCLC.

(19) A subsequent meta-analysis confirmed these results using individual patient data from a total of 322 centers participating in the same set of 7 trials. (20) The following data were requested for the subsequent meta-analysis: patient identifier, center identifier, randomization date, treatment assigned by randomization, age, gender, body mass index, performance status, stage, overall tumor response to the first assigned treatment, date of response, date of progression with the first allocated treatment, date of death or last visit, survival status, and cause of death if applicable.

Time to Event Analyses

Progression-free survival was defined as the time from random assignment to disease progression (as assessed in each individual trial) or death from any cause. Overall survival was defined as the time from random assignment to death from any cause. The distributions of PFS and OS were estimated using the Kaplan-Meier method. Treatment groups were compared using a Cox regression model. The median follow-up time was estimated using the Kaplan-Meier method with censoring for death.

Surrogacy Analyses

A two-level modeling approach was adopted to estimate the association between PFS and OS, and between the treatment effects on these endpoints. Treatment effects were estimated as logarithms of the hazard ratio (logHR). The logHR has intuitive appeal as a measure of treatment effect: it is equal to zero in the absence of a treatment effect, and is approximately equal to the risk reduction for small treatment effects (hence a logHR of -0.10 corresponds to a risk reduction of about 10%, and a logHR of 0.10 corresponds to a risk increase of about 10%). The logHRs were estimated within units of analysis consisting of 135 centers or 64 strata. Centers were either individual centers if they had more than 3 patients per treatment arm, or groups of small centers with an average size at least equal to the average size of the big centers of the same trial (Tax 326: average size = 15, Taxobel 303: average size = 17, HORG: average size = 25, WJTOG 9904: average size = 14, French: average size = 20). Strata were defined within each trial by the cross-classification of the following prognostic factors: age (<60 *vs.* >= 60 years), gender (male *vs.* female), performance status (ECOG 0 or 1 *vs.* 2 or 3), BMI (<18.5 *vs.* >= 18.5 kg/m²),

histology (squamous *vs.* non-squamous) and stage (IIIb *vs.* IV *vs.* unknown). Prognostic strata were formed as follows: a Cox model for overall survival was fit within each trial with treatment, each of the prognostic factors listed above, and the treatment-prognostic factor interaction. The prognostic factors were ordered by increasing level of significance for the treatment-prognostic factor interaction. The first prognostic factor selected in this way was the factor most predictive of the effect of treatment on overall survival, and was used to split the patients of a trial into two (or three) strata. Each of these strata was then split by the second prognostic factor; and so on. The splitting was stopped when it produced strata with less than 3 patients per treatment arm.

The association between PFS and OS was quantified through a bivariate copula model fitted on individual patient data. Kendall's τ was used to quantify the correlation between the endpoints. (21) A linear regression model was fitted on the estimated treatment effects on PFS and OS (logHRs for PFS and OS). Coefficients of determination (equal to squared correlation coefficients) were estimated using weighted linear regression. (21) Coefficients of determination (R^2) quantify the proportion of variance explained by the regression. The surrogate threshold effect was defined as the minimum treatment effect on PFS required to predict a non-zero treatment effect on OS in a future trial. (22)

RESULTS

Treatment Effects on PFS and OS

A total of 2,~~334~~ 331 patients were included in the analysis, since the PFS and/or OS was missing for 24 (1%) of all patients. The median follow-up of patients still alive was 23.4 months. For the entire cohort, the median OS was 10.0 months and the median PFS was 5.5 months (Figure 1), with little difference between the curves until about 12 months.

Figure 1 here

The hazard ratios were 0.97 for PFS (95% C.I., 0.89 – 1.05, P = 0.44) and 0.92 for OS (95% C.I., 0.84 – 1.01, P = 0.089) (Figure 2). There was significant heterogeneity between the 5 trials in terms of PFS (P = .01) but not in terms of OS (P = 0.72)

Figure 2 here

Correlation between PFS and OS

PFS showed some correlation with OS (τ = 0.59; 95% C.I., 0.58 – 0.61).

Correlation Between Treatment Effects

The coefficient of determination between treatment effects estimated within centers was R^2 = 0.62 (95% C.I. = 0.52 – 0.72). The linear regression equation was logHR(OS) = -0.048 + 0.76 × logHR(PFS) (Figure 3). Using centers as the unit of analysis, the surrogate threshold effect was a PFS hazard ratio of 0.49, indicating that a risk reduction of 51% in terms of PFS would predict a non-zero effect on OS.

Figure 3 here

The coefficient of determination between treatment effects estimated within strata was R^2 = 0.88 (95% C.I. = 0.60 – 0.84). The linear regression equation was logHR(OS) = -0.071 + 0.87 × logHR(OS) (Figure 4). Using strata as the unit of analysis, the surrogate threshold effect was a PFS hazard ratio of 0.53, indicating that a risk reduction of 47% in terms of PFS would predict a non-zero effect on OS.

Figure 4 here

DISCUSSION

Our analyses suggest that PFS is not a statistically acceptable surrogate endpoint for OS in patients with metastatic non-small cell lung cancer treated in first line with docetaxel-based or vinorelbine-based chemotherapies. Indeed, although about two-thirds of the treatment effects on OS are explained by the treatment effects on PFS, the surrogate threshold effect ranges from 0.49 to 0.53 depending on whether strata or centers are used as the unit of analysis, which implies that only a major benefit of some new drug on PFS (hazard reduction of about one half or greater) would be expected to also produce a non-zero benefit on OS. These analyses are quite similar whether treatment effects are estimated in the centers participating to the trials, or in the strata defined by the characteristics of the patients most predictive of survival benefits, despite the fact that the latter analysis could have overestimated the association between the treatment effects through deliberate confounding by the prognostic factors used to define the strata.

In this set of trials, docetaxel-based regimens showed a trend towards better results than vinorelbine-based regimens, but the difference was not significant for either OS (HR = 0.92, P = 0.089) or PFS (HR = 0.97, P = 0.44). If anything, the difference was more pronounced for OS than for PFS, an unusual finding with advanced solid tumors, for which a benefit on PFS is generally diluted to yield a smaller benefit on OS. (10-11) Other meta-analyses did not support this finding. A meta-analysis comparing gemcitabine-platinum with other platinum-containing regimens found about the same benefit of gemcitabine-platinum on PFS (HR = 0.88, information available on 14 of 17 trials) as on OS (HR = 0.91 for the same 14 trials). (23) A meta-analysis of trials comparing longer with shorter durations of chemotherapy found a much more pronounced benefit of longer chemotherapy duration on PFS (HR = 0.75 on 9 of 13 trials) than on OS (HR = 0.93 for the same 9 trials). (24)

The difference between the median PFS and the median OS was only 4.5 months in this and other meta-analyses (Le Chevalier 2005), and therefore the gain in time from using PFS instead of OS in future trials of chemotherapy for advanced NSCLC would not be as

large as in other tumor types. (10-11) The short survival time post progression implies that differences in overall survival are likely to be observed for truly effective new treatments. (25) All in all, these findings suggest that even if PFS could be proposed as a plausible surrogate for OS from a statistical point of view, it would not be a very attractive one to evaluate the worth of conventional chemotherapies for advanced non-small cell lung cancer. The exclusion of two trials with unreliable PFS from our meta-analysis casts further doubts on the usefulness of this endpoint in advanced NSCLC, at least as measured a decade ago. Such exclusions also cast some doubts on meta-analyses that rely solely on published papers rather than on carefully reviewed individual patient data. (26)

Our analyses have several limitations. We used a pragmatic approach, using PFS as measured by the investigators in each trial, ignoring any possible differences in measurement techniques or schedules. While such differences may have an impact on PFS duration, they are unlikely to have much impact on the PFS hazard ratio. It is also unlikely that the results would have been much different, had a blinded central review of PFS been available in all trials. (27) The fact that treatment doses and schedules differed from trial to trial does not raise any particular concern. Indeed, such differences could have obscured (rather than enhanced) the relationship between treatment effects on PFS and on OS, hence the observed relationship is probably an underestimate of what would have been observed in a more homogeneous setting. More importantly, the randomized comparisons in this set of trials were between two standard combinations of cytotoxic drugs. Although these analyses provide modest support for considering PFS an acceptable surrogate for OS in patients with advanced NSCLC, treatments that have a major impact on PFS would be expected to also have a significant effect on OS. The relationship between PFS and OS, and between treatment effects on PFS and OS, which might not be the same for different cytostatic agents or for targeted agents, and given the obvious advantages of using PFS as the primary endpoint in randomized trials, these issues deserve further investigation through further meta-analyses of contemporary randomized trials. (28)~~The relationship between PFS and OS, and between treatment effects on PFS and OS, might not be the same for different cytostatic agents, or for~~

~~targeted agents. Likewise, our results should not be extrapolated to today's environment, since more drugs with demonstrated activity in lung cancer are currently available than was the case in the trials analyzed here. Given the obvious advantages of using PFS as the primary endpoint in randomized trials, these issues deserve further investigation through further meta-analyses of contemporary randomized trials. (28)~~

SOURCES OF FUNDING

This work was supported by an unrestricted research grant from sanofi-aventis, Paris, France, and by the Ligue Nationale Contre le Cancer, France.

REFERENCES

1. Alonso A, Molenberghs G. Evaluating time to cancer recurrence as a surrogate marker for survival from an information theory perspective. Statistical Methods in Medical Research 17: 497-504, 2008
2. Baker SG. Two simple approaches for validating a binary surrogate endpoint using data from multiple trials. Statistical Methods in Medical Research 17: 505-14, 2008
3. Burzykowski T. Surrogate endpoints: wishful thinking or reality? Statistical Methods in Medical Research 17: 463-6, 2008
4. ChakravartyA, Sridhara R. Use of progression-free survival as a surrogate marker in oncology trials: some regulatory issues. Statistical Methods in Medical Research 17: 515-18, 2008
5. Green EM, Yothers G, Sargent DJ. Surrogate endpoint validation: statistical elegance versus clinical relevance. Statistical Methods in Medical Research 17: 477-86, 2008
6. Grothey A. Surrogate endpoints for overall survival in early colorectal cancer from the clinician's perspective. Statistical Methods in Medical Research 17: 529-35, 2008
7. Hughes MD. Practical issues arising in an exploratory analysis evaluating progression-free survival as a surrogate endpoint for overall survival in advanced colorectal cancer. Statistical Methods in Medical Research 17: 487-95, 2008
8. Piedbois P, Croswell JM. Surrogate endpoints for overall survival in advanced colorectal cancer: a clinician's perspective. Statistical Methods in Medical Research 17: 519-27, 2008
9. Buyse M, Burzykowski T, Michiels S, Carroll K. Individual- and trial-level surrogacy in colorectal cancer. Statistical Methods in Medical Research 17: 467-75, 2008
10. Buyse M, Burzykowski T, Carroll K, et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. J Clin Oncol 25: 5218-24, 2007
11. Burzykowski T, Buyse M, Piccart-Gebhart MJ, et al. Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. J Clin Oncol 26: 1987-92, 2008
12. Fossella F, Pereira JR, von Pawel J, 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 21: 3016–3024, 2003

13. Monnier A, Zatloukal P, Siffnerova H, et al. Docetaxel (D) versus vinorelbine–cisplatin (VC) as front-line treatment for patients with advanced non-small cell lung cancer (NSCLC) and with a normal level of serum alpha 1 acid glycoprotein (AAG): Additional data on quality of life (QOL). Preliminary results. Proc Am Soc Clin Oncol 2003 [Abstract 2680].
14. Kubota K, Watanabe K, Kunitoh H, et al. Phase III randomized trial of docetaxel plus cisplatin versus vindesine plus cisplatin in patients with stage IV non-small-cell lung cancer: The Japanese Taxotere Lung Cancer Study Group. J Clin Oncol 22: 254–261, 2004
15. Douillard JY, Gervais R, Dabouis G, et al. Sequential two-line strategy for stage IV non-small-cell lung cancer: docetaxel–cisplatin versus vinorelbine–cisplatin followed by crossover to single-agent docetaxel or vinorelbine at progression: final results of a randomized phase II study. Ann Oncol 16: 81–89, 2005
16. Georgoulias V, Ardavanis A, Tsiafaki X, et al. Vinorelbine plus cisplatin versus docetaxel plus gemcitabine in advanced non-small-cell lung cancer: A phase III randomized trial. J Clin Oncol 23: 2937–2945, 2005
17. Pujol JL, Breton JL, Gervais R, et al. Gemcitabine–docetaxel versus cisplatin–vinorelbine in advanced or metastatic non-small-cell lung cancer: A phase III study addressing the case for cisplatin. Ann Oncol 16: 602–610, 2005
18. Kudoh S, Takeda K, Nakagawa K, et al. Phase III study of docetaxel compared with vinorelbine in elderly patients with advanced non-small-cell lung cancer: Results of the West Japan Thoracic Oncology Group Trial (WJTOG 9904). J Clin Oncol 24: 3657–3663, 2006
19. Douillard JY, Laporte S, Fossella F, et al. Comparison of docetaxel- and vinca alkaloid-based chemotherapy in the first-line treatment of advanced non-small cell lung cancer: a meta-analysis of seven randomized clinical trials. J Thorac Oncol 2: 939-46, 2007
20. Laporte S, Buyse M, Baroux N, Fossella F, et al. Meta-analysis comparing docetaxel and vinca-alkaloids in the first-line treatment of NSCLC. Comparison of results based

on individual patient data, study report data, and published data. Eur J Cancer Suppl. 1: Abstract 4, 2007

21. Burzykowski T, Molenberghs G, Buyse M, et al: Validation of surrogate endpoints in multiple randomized clinical trials with failure-time end points. J R Stat Soc C 50: 405-422, 2001
22. Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharmaceutical Stat 5: 173-186, 2006
23. Le Chevalier T, Scagliotti G, Natale R, et al. Efficacy of gemcitabine plus platinum chemotherapy compared with other platinum containing regimens in advanced non-small-cell lung cancer: a meta-analysis of survival outcomes. Lung Cancer 47: 69–80, 2005
24. Soon YY, Stockler MR, Askie LM, Boyer MJ. Duration of chemotherapy for advanced non-small-cell lung cancer: a systematic review and meta-analysis of randomized trials. J Clin Oncol 27: 3277–3283, 2009
25. Broglio KR, Berry DA. Detecting and overall survival benefit that is derived from progression-free survival. J Natl Cancer Inst 101: 1642-1649, 2009
26. Johnson KR, Ringland C, Stodkes BJ, et al. Response rate or time to progression as predictors of survival in trials of metastatic colorectal cancer or non-small-cell lung cancer: a meta-analysis. Lancet Oncology 7: 741-6, 2006
27. Amit O, Mannino F, Stone AM, et al. Blinded independent central review of progression in cancer clinical trials: Results from a meta-analysis. Eur J Cancer 47:1772-8; 2011
28. Soria JC, Massard C, Le Chevalier T. Should progression-free survival be the primary measure of efficacy for advanced NSCLC therapy? Annals Oncol 21: 2324-32, 2010

Table 1. Trials included in the surrogacy analysis (N: number of patients randomized; D, docetaxel; C, cisplatin; Cb, carboplatin; G, gemcitabine, V, vinorelbine; AUC, Area Under the Curve; HORG, Hellenic Oncology Research Group; WJTOG, West Japan Thoracic Oncology Group; q3/4 wks = every 3/4 weeks). Drug doses are indicated in mg/m² except for Cb, which was dosed to obtain an AUC of 6 mg/mL.

Trial name	First Author	Accrual period	N	Follow-up (months)	Docetaxel arm	Vinorelbine arm
Tax 326	Fossella	1998-2000	1218	21	D 75 C 75 q3 wks × 6 or D 75 Cb AUC 6 q3 wks × 6	V 25 C 100 q4 wks × 6
HORG	Georgoulias	1999-2002	413	20	D 100 G 1000 q3 wks × 6*	V 30 C 80 q3 wks × 6*
French	Pujol	1999-2001	311	25	D 85 G 1000 q3 wks × 8	V 30 C 100 q4 wks × 6
Taxobel 303	Douillard	1998-2000	233	43	D 75 C 100 q3 wks × 6	V 30 C 100 q3 wks × 6
WJTOG 9904	Kudoh	2000-2003	180	26	D 60 q3 wks × 4	V 25 q4 wks × 3

* Three additional cycles were given to patients in complete or partial response after the sixth cycle

Table 2. Baseline patient characteristics for the trials listed in Table 1 (BMI: body mass index; ECOG PS: Eastern Cooperative Oncology Group Performance Status).

Trial name	Age in years (median, range)	BMI (median, range)	Gender (Male / Female)
Tax 326	60 (23 - 87)	24.3 (15.4 – 49.6)	73% / 27%
HORG	64 (36 – 78)	Not available	89% / 11%
French	58 (37 – 75)	23.9 (15.0 – 44.1)	80% / 20%
Taxobel 303	58 (27 – 77)	23.6 (15.6 – 40.0)	82% / 18%
WJTOG 9904	76 (70 – 86)	20.5 (14.4 – 28.8)	75% / 25%

Trial name	ECOG PS (% 0 / % 1 / % 2+)	Stage (% IIIb / % IV)	Squamous cell carcinoma (% Yes / % No)
Tax 326	16% / 80% / 4%	33% / 67%	33% / 67%
HORG	44% / 46% / 10%	38% / 62%	Not available
French	21% / 71% / 8%	16% / 84%	28% / 72%
Taxobel 303	31% / 54% / 15%	0% / 100%	33% / 67%
WJTOG 9904	0% / 96% / 4%	36% / 64%	33% / 67%

LEGEND TO FIGURES

Figure 1. Kaplan-Meier estimates of overall survival and progression-free survival by treatment arm.

Figure 2. Forest plot of hazard ratios for overall survival and progression-free survival.

Figure 3. Correlation between treatment effects (log hazard ratios) on progression-free survival (horizontal axis) and overall survival (vertical axis) in participating centers. The size of each circle is proportional to the number of patients in the corresponding center. The reference circle in black corresponds to 10 patients.

Figure 4. Correlation between treatment effects (log hazard ratios) on progression-free survival (horizontal axis) and overall survival (vertical axis) in prognostic strata (see text). The size of each circle is proportional to the number of patients in the corresponding stratum. The reference circle in black corresponds to 10 patients.

Figure 1

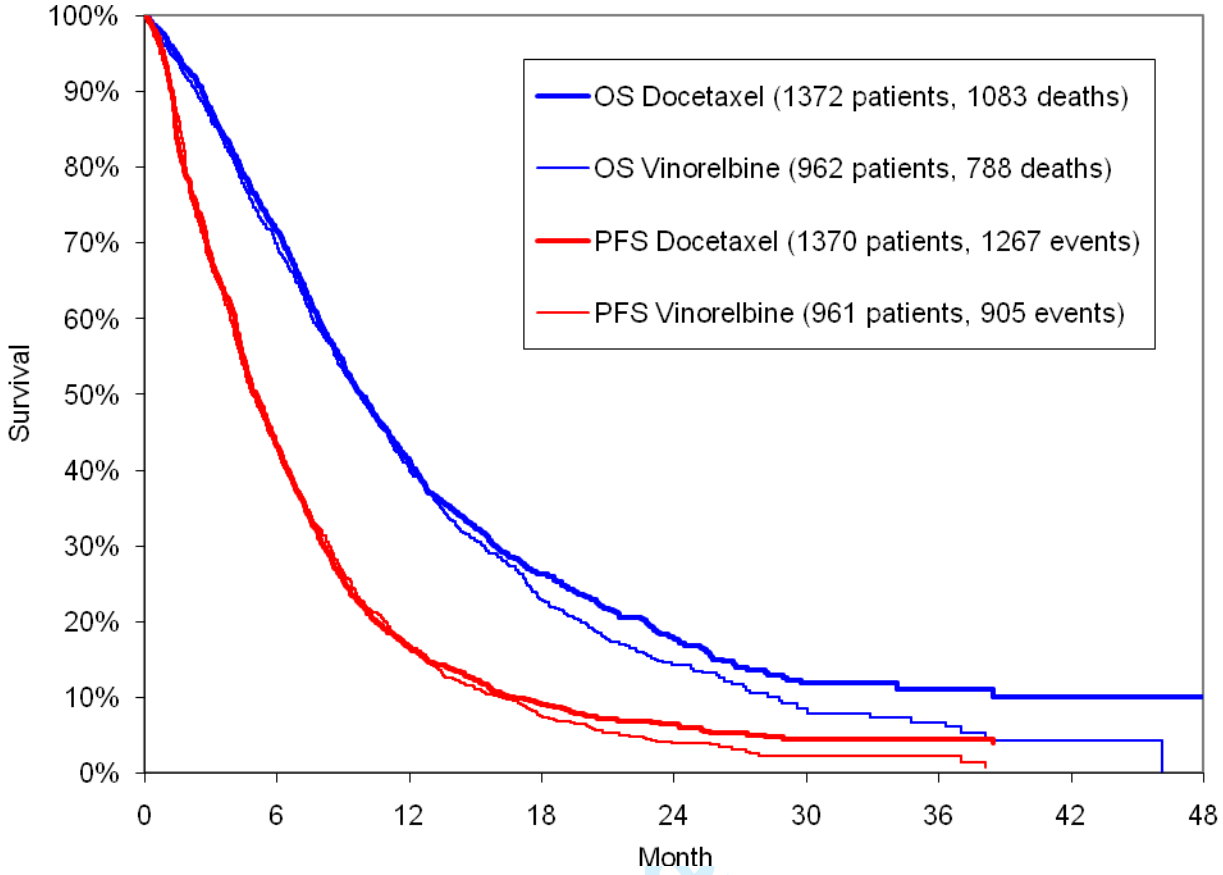


Figure 2

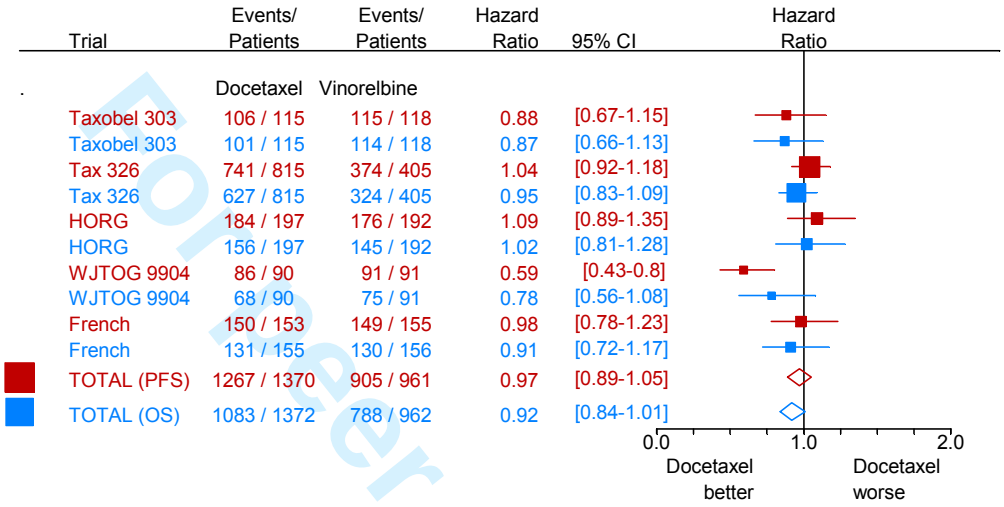


Figure 3

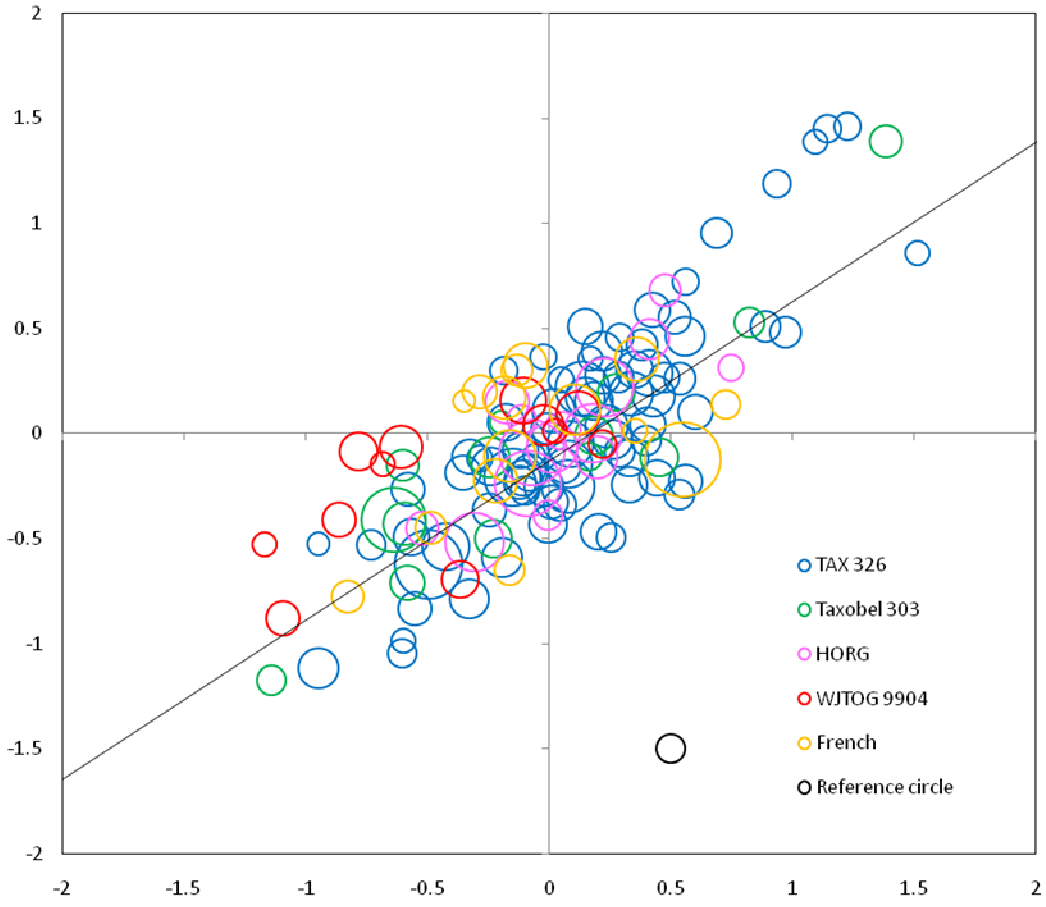


Figure 4

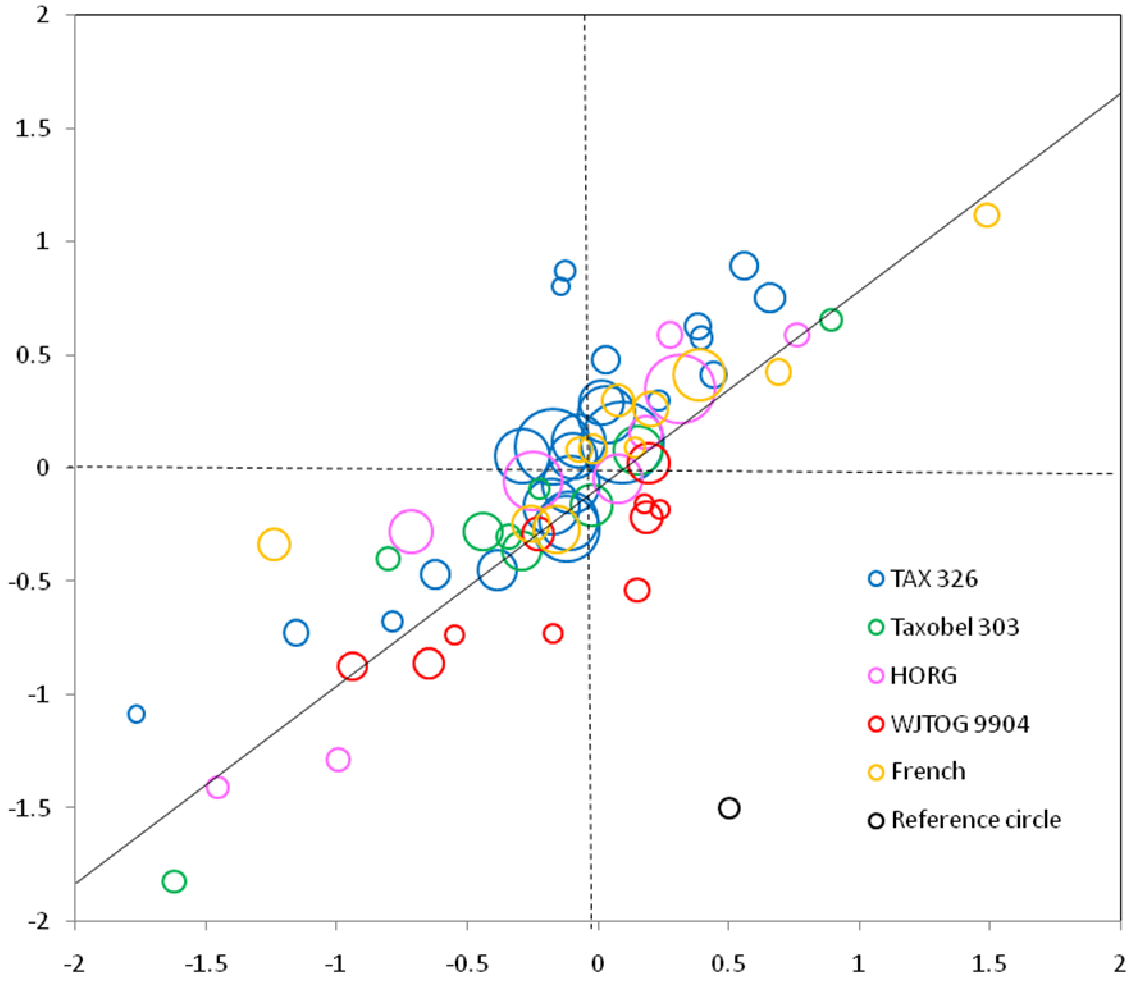
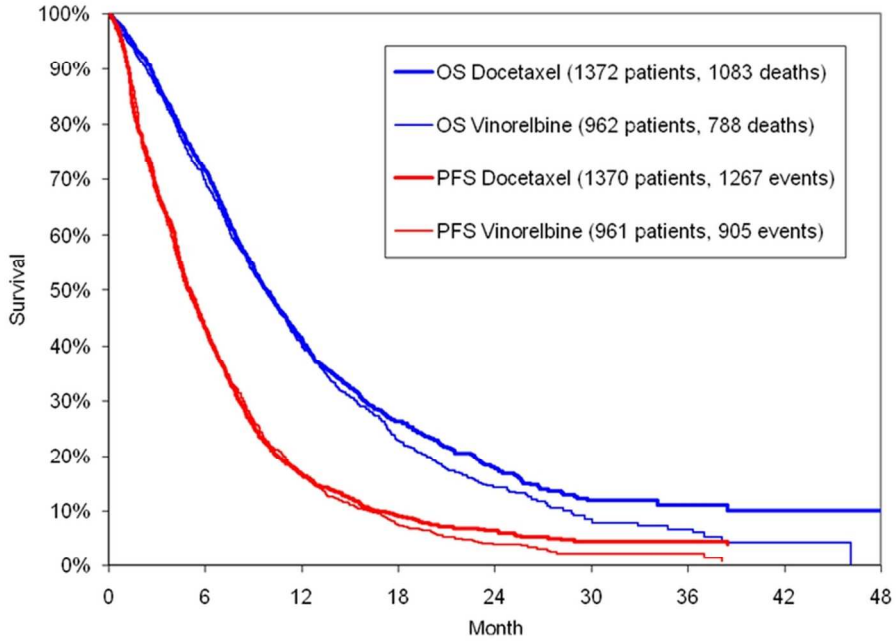
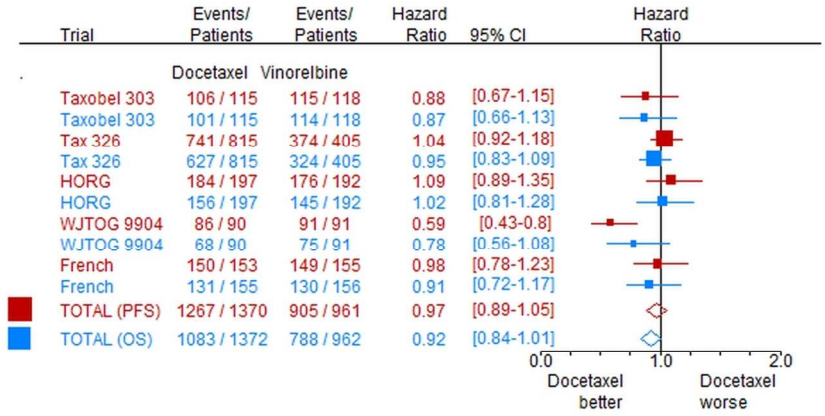


Figure 1



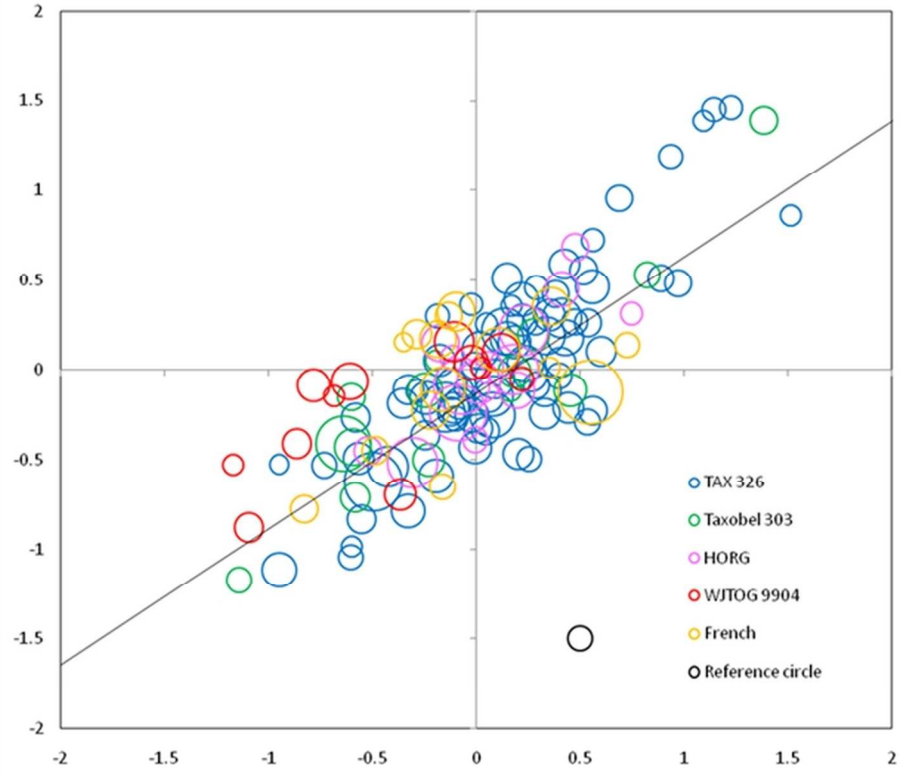
113x90mm (300 x 300 DPI)

Figure 2



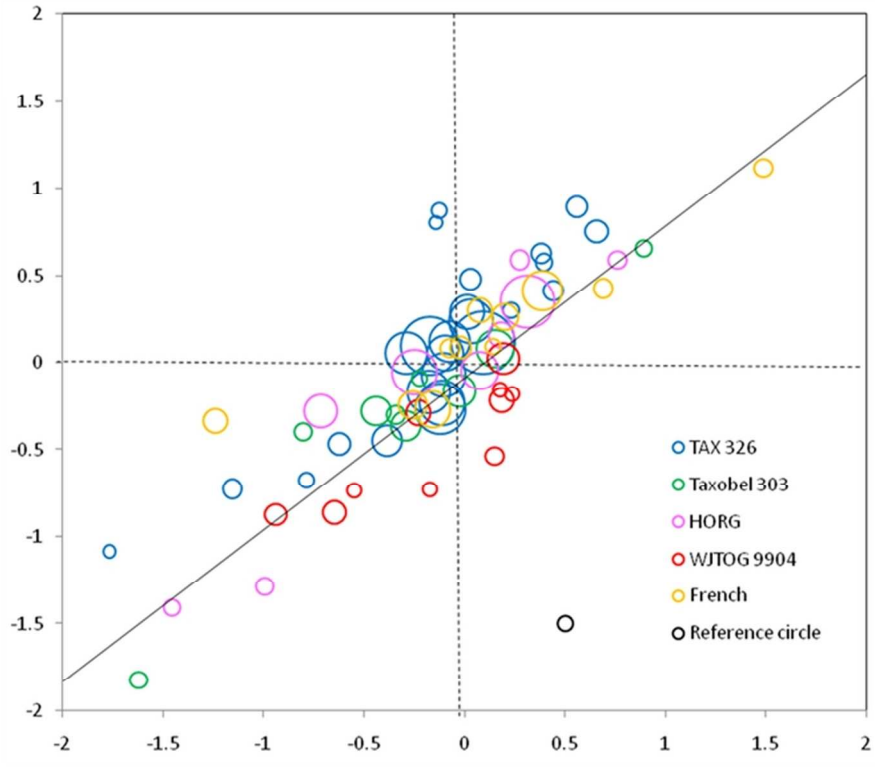
124x90mm (300 x 300 DPI)

Figure 3



97x90mm (300 x 300 DPI)

Figure 4



103x90mm (300 x 300 DPI)