EUROPEAN JOURNAL OF CANCER 48 (2012) 385–388

Available at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.ejconline.com

Current perspective

# Poor correlation between progression-free and overall survival in modern clinical trials: Are composite endpoints the answer?

*Eitan Amir* [a,*], *Bostjan Seruga* [b], *Ryan Kwong* [a], *Ian F. Tannock* [a], *Alberto Ocaña* [a,c]

[a] *Division of Medical Oncology and Hematology, Princess Margaret Hospital, Toronto, Canada*
[b] *Sector of Medical Oncology, Institute of Oncology Ljubljana, Ljubljana, Slovenia*
[c] *Albacete University Hospital and AECC Unit, Albacete, Spain*

ARTICLE INFO

*Article history:*
Available online 22 November 2011

*Keywords:*
Cancer Trials
Endpoints
Surrogacy
Patient reported outcomes
Survival

ABSTRACT

It can be difficult to identify endpoints that accurately reflect patient benefit in metastatic solid tumors. Overall survival (OS) is the gold standard although progression-free survival (PFS) is sometimes used as a surrogate for OS. Statistical modelling has suggested that the association between OS and PFS becomes weaker in diseases with longer survival post-progression (SPP). To evaluate these statistical hypotheses we determined the relationship between PFS and OS in control and experimental arms of randomised trials conducted in the last 10 years, which have led to drug approval. Our data confirm that PFS is a poor surrogate for OS when SPP is long, but it is a better surrogate where SPP is short. In cancers with short SPP designing trials to show OS benefit is feasible and, therefore, remains the preferred approach. In tumours with long SPP, PFS is not clinically meaningful unless it is also associated with improvement in patient reported outcomes such as quality of life. The oncology community should consider the further development and validation of composite endpoints including patient reported outcomes and PFS across different disease sites. Such endpoints have been successfully used in cancer trials in the past. With improvements in therapy and prolonged survival of patients with many cancers, and with increasing pressure from healthcare payers to prove that treatment leads to patient benefit, the choice of optimal endpoints for clinical trials is increasingly important. Composite measures comprising patient reported outcomes and intermediate endpoints such as PFS may be the solution and should be investigated further.



## 1. Introduction

For patients with metastatic solid tumours, designing clinical trials to show a difference in overall survival (OS) between the experimental and control arms can be difficult. Cross-over of patients from one therapy to another can decrease the ability to detect an improvement in OS due to the experimental therapy, while heterogeneity in the natural history of disease post-progression, and differences in post-progression management, can dilute any effect due to earlier treatment.

* *Corresponding author:* Address: Division of Medical Oncology and Hematology, Princess Margaret Hospital, 610 University Avenue, Toronto, ON, Canada M5G 2M9. Tel.: +1 416 946 4501x5181; fax: +1 416 946 4563.
E-mail address: eitan.amir@uhn.on.ca (E. Amir).
0959-8049/$ - see front matter 
doi:10.1016/j.ejca.2011.10.028

Therefore, it has been suggested that intermediate endpoints such as progression-free survival (PFS) can be used as a surrogate for OS.

For an outcome measure to be valid as a surrogate endpoint it needs to fulfil certain criteria. First, there should be strong and consistent correlation between the surrogate and definitive endpoints. Once correlation is established, a surrogate endpoint should also predict the net effect of treatment on the clinical outcome.[1,2] While it is simple to assess correlation between a surrogate and definitive endpoint, assessment of net effect requires complex regression analyses.[3] To date, no intermediate endpoint has been shown to be a universal measure of direct patient benefit or a consistent surrogate for OS across different disease sites.[4] Evidence for the validity of surrogate endpoints in specific malignancies is also variable. In the setting of first-line treatment of advanced colorectal cancer with fluorouracil-based chemotherapy, PFS has been shown to correlate strongly with OS.[5] The correlation coefficient between treatment effects on PFS and on OS is also high[6] and therefore, PFS has been validated as a surrogate endpoint for OS. PFS has also been validated as a surrogate for OS in ovarian cancer,[7] but in advanced breast cancer criteria for surrogacy of PFS were not met.[8]

It is unclear why PFS can be validated as a surrogate for OS in some cancers but not others. Statistical modelling assessing the correlation of PFS and OS for different hypothetical durations of survival post-progression (SPP) in a single disease site model suggested that the association between OS and PFS is likely to become weaker with increasing SPP.[9] However, assumptions made during these statistical simulations could compromise external validity. For example, it was postulated that treatment only influenced PFS and had no effect on SPP. This may not always be true as treatments given in early stages of disease might alter tumour biology and influence outcomes at later stages.[10] Preclinical studies have suggested accelerated tumour growth, invasion and metastasis after withdrawal of certain anti-neoplastic drugs.[11,12] Clinical data do not support these findings,[13] however, this remains an area of uncertainty.[14]

## 2. Methods

We therefore explored the correlation between PFS and OS in real randomised controlled trials (RCTs) supporting registration of new anti-cancer drugs approved by the US Food and Drug Administration in the last 10 years. Included agents were approved for the treatment of metastatic malignancies in adults. Agents used in supportive cancer care were not included. Eligible studies included those reporting HR for PFS (or time to progression, TTP) and for OS as well as data allowing for the calculation of SPP. The association between OS and PFS was assessed using linear regression weighted by the trial sample size. For the primary analysis, identification of a cut-off for median SPP with the greatest discrimination was conducted using a non-parametric (spline) smooth function applied to the correlation between the ratio of OS to PFS and to SPP. Secondary analyses were conducted for other cut-offs for median SPP. These were chosen by either multiplying or dividing the primary cut-off by 1.5 and 2.0 times. The magnitude of correlation was assessed as described by Burnand et al.[15]

## 3. Results

A total of 25 drugs were identified and 48 RCTs were referenced by the FDA in support of the approval process. Fourteen studies did not report HR for both OS and PFS (or TTP). A further eight studies did not report data allowing for the calculation of SPP. Therefore, 26 studies met the inclusion criteria. These studies included the evaluation of both chemotherapy and targeted therapy. Characteristics of included studies are shown in the Supplementary Table.

Patients with pancreatic and non-small cell lung cancers had the shortest SPP, those with colorectal, renal and head and neck cancer had intermediate SPP while those with breast and ovarian cancers had the longest SPP. Non-parametric smoothing analysis identified that the correlation between the ratio of OS to PFS and median SPP became less

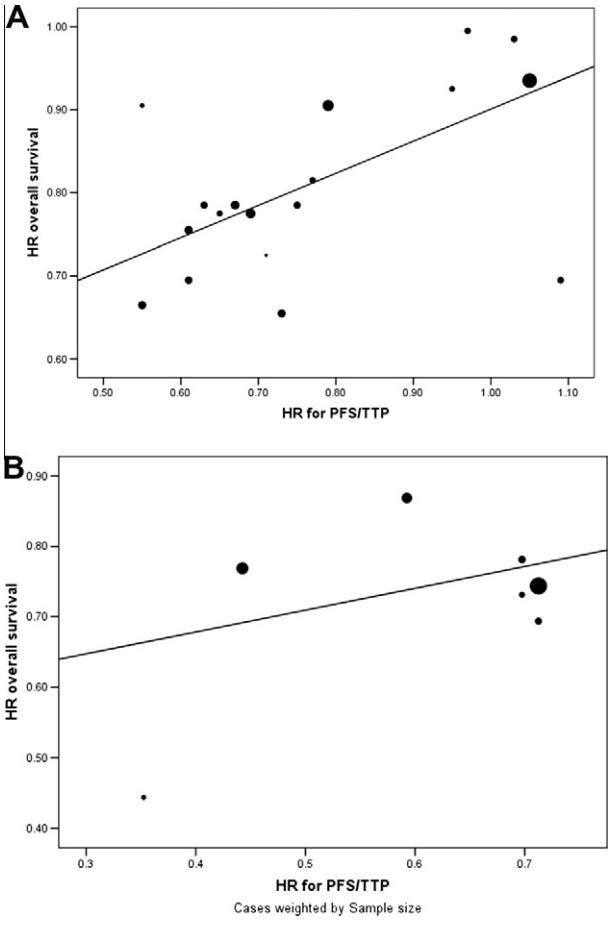


**Fig. 1 – Correlation of overall survival and progression-free survival. (A) for survival post-progression less than 12 months (Pearson Coefficient R = 0.64). (B) for survival post-progression of 12 months or more (Pearson Coefficient R = 0.38). The diagonal line is the fitted weighted linear regression line. The size of the circle is relative to the sample size of the study.**

**Table 1 – Correlation of overall survival and progression-free survival for different cut-offs of survival post progression (SPP). Primary analysis is shaded.**

| Cut-off (months) | Below cut-off | | | Above cut-off | | |
|---|---|---|---|---|---|---|
| | Number of studies | Median SPP | Pearson Coefficient *R* | Number of studies | Median SPP | Pearson Coefficient *R* |
| 6 | 8 | 5.0 | 0.50 | 18 | 10.0 | 0.68 |
| 8 | 15 | 5.6 | 0.62 | 11 | 11.6 | 0.50 |
| 12 | 19 | 6.4 | 0.64 | 7 | 18.5 | 0.38 |
| 18 | 22 | 7.8 | 0.57 | 4 | 28.5 | 0.99 |
| 24 | 23 | 7.8 | 0.63 | 3 | 31.9 | −0.28 |

strong when SPP was approximately 12 months or longer. The primary analysis was therefore carried out for studies with SPP <12 months and those with SPP ⩾12 months. This cut-off is also pragmatic and clinically relevant. Secondary analyses for SPP cut-offs of 6, 8, 18 and 24 months were also conducted. For the group with SPP <12 months, there was reasonably good correlation between the HRs for OS and for PFS ($R = 0.64$, see Fig. 1A). For the group with SPP ⩾12 months, the correlation between the HRs for OS and for PFS was poor ($R = 0.38$ see Fig. 1B). The correlation coefficients for different cut-offs for SPP are shown in Table 1. In general, correlations between PFS and OS were poor to moderate regardless of the cut-off used.

## 4. Discussion

Our data support those derived from statistical modelling and confirm that when SPP is short, there is better correlation between OS and PFS than when SPP is longer. However, even where SPP was <12 months the correlation coefficient between PFS and OS in clinical trials was only 0.64. The use of PFS as a primary endpoint therefore appears of limited clinical significance in most solid tumours and the oncology community may be over-interpreting the value of improvement in this endpoint.[16]

This analysis has limitations. First, as described above, included studies were conducted in a variety of malignancies and at different points in the course of disease. This introduces heterogeneity to the analysis as outcomes and treatment options differ between studies. However, it would be expected that these differences would be accounted for by differences in SPP and grouping of studies was matched for this variable. Secondly, only studies reporting HR for OS and PFS as well as data from which SPP could be assessed were included, and our selection may be influenced by publication bias. Finally, the FDA website was utilised to identify RCTs which were cited in support of the drug approval process. Since these studies were 'positive', this likely led to higher correlation between OS and PFS. Our data therefore represent a best-case scenario. It is expected that in unselected studies the correlation between PFS and OS would be lower than that observed in the present analysis.

The choice of optimal primary endpoint for clinical trials therefore remains contentious. In cancers where SPP is long relative to PFS (such as in first-line therapy of hormone receptor-positive breast cancers), PFS does not predict for OS and therefore is a weak endpoint for patient benefit. Unfortunately, it is also difficult to design and execute a trial to show benefit in OS in this setting. The prevention of progression and symptoms associated with progression can be as important as improving long-term survival in such diseases. Furthermore, avoidance of equipoise where improvements in PFS may be countered by treatment-related toxicity is important especially as newly approved cancer therapies have been shown to have increased toxicity.[17] In such settings, improved PFS may be a clinically important outcome, but only if it leads to improvement in symptoms or quality of life. If PFS is long relative to SPP (such as first-line therapy in pancreas or non-small cell lung cancers or later line treatments of other cancers), then OS values are dominated by PFS, so any effect on PFS will likely also show a benefit in OS. However, even though PFS may be a valid surrogate, OS is a more definitive endpoint[18] and should therefore be preferred. Maintenance or improvement of quality of life is important in this setting and rigorous assessments of patient reported outcomes are imperative in this setting as well.

The oncology community should consider the development and validation of a composite endpoint including PFS and patient reported outcomes for cancers with long SPP. Such composite endpoints are in frequent use in other branches of medicine such as cardiology[19,20] as well as in decision analysis[21] pharmaco-economics.[22] They have also been used successfully in oncology.[23,24] In the AXIS trial of axitinib versus sorafenib as second-line therapy for metastatic renal cell carcinoma, the primary endpoint was PFS and reported results showed superiority of axitinib.[24] Investigators then undertook a secondary analysis of time to deterioration which was defined as a composite endpoint of death, progression or worsening in patient reported outcomes, whichever occurred first. Results showed that axitinib delayed time to this composite endpoint compared with sorafenib.[25] The use of such composite endpoints appears desirable, but would require validation and testing under a variety of conditions before being applied to other settings.

Important questions regarding the use of such composite endpoint remains, especially the relative weighting of PFS and patient reported outcomes in the final measurement. It is our view that in order to ensure clinically relevant results, the composite endpoint should be weighted in favour of patient reported outcomes and effect sizes for the different components of the endpoint should be reported to ensure transparency.[26]

With improvements in therapy and prolonged survival of patients with many cancers, and with increasing pressure from healthcare payers to prove that treatment leads to patient benefit,[27] the choice of optimal endpoints for clinical trials is increasingly important. Composite measures comprising patient reported outcomes and intermediate endpoints such as PFS may be the solution and should be investigated further.

## Conflict of interest statement

None declared.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejca.2011.10.028.

REFERENCES


1. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med* 1989;**8**(4):431–40.
2. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;**125**(7):605–13.
3. Lin DY, Fleming TR, De GV. Estimating the proportion of treatment effect explained by a surrogate marker. *Stat Med* 1997;**16**(13):1515–27.
4. Korn EL, Freidlin B, Abrams JS. Overall survival as the outcome for randomized clinical trials with effective subsequent therapies. *J Clin Oncol* 2011;**29**(17):2439–42.
5. Tang PA, Bentzen SM, Chen EX, Siu LL. Surrogate end points for median overall survival in metastatic colorectal cancer: literature-based analysis from 39 randomized controlled trials of first-line chemotherapy. *J Clin Oncol* 2007;**25**(29):4562–8.
6. Buyse M, Burzykowski T, Carroll K, et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. *J Clin Oncol* 2007;**25**(33):5218–24.
7. Bast RC, Thigpen JT, Arbuck SG, et al. Clinical trial endpoints in ovarian cancer: report of an FDA/ASCO/AACR Public Workshop. *Gynecol Oncol* 2007;**107**(2):173–6.
8. 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* 2008;**26**(12):1987–92.
9. Broglio KR, Berry DA. Detecting an overall survival benefit that is derived from progression-free survival. *J Natl Cancer Inst* 2009;**101**(23):1642–9.
10. Oliver TG, Mercer KL, Sayles LC, et al. Chronic cisplatin treatment promotes enhanced damage repair and tumor progression in a mouse model of lung cancer. *Genes Dev* 2010;**24**(8):837–52.
11. Ebos JM, Lee CR, Cruz-Munoz W, et al. Accelerated metastasis after short-term treatment with a potent inhibitor of tumor angiogenesis. *Cancer Cell* 2009;**15**(3):232–9.
12. Paez-Ribes M, Allen E, Hudock J, et al. Antiangiogenic therapy elicits malignant progression of tumors to increased local invasion and distant metastasis. *Cancer Cell* 2009;**15**(3):220–31.
13. Miles D, Harbeck N, Escudier B, et al. Disease course patterns after discontinuation of bevacizumab: pooled analysis of randomized phase III trials. *J Clin Oncol* 2011;**29**(1):83–8.
14. Potemski P. Is the postprogression survival time really not shortened in the bevacizumab-containing arms of phase III clinical trials? *J Clin Oncol* 2011;**29**(13):e384–5.
15. Burnand B, Kernan WN, Feinstein AR. Indexes and boundaries for "quantitative significance" in statistical decisions. *J Clin Epidemiol* 1990;**43**(12):1273–84.
16. Ocana A, Amir E, Vera F, Eisenhauer EA, Tannock IF. Addition of bevacizumab to chemotherapy for treatment of solid tumors: similar results but different conclusions. *J Clin Oncol* 2011;**29**(3):254–6.
17. Niraula S, Amir E, Ocana A, Seruga B, Tannock I. The balance between benefits and harms of molecular targeted agents. *J Clin Oncol* 2011 [abstr. 6030].
18. Korn EL, Freidlin B, Abrams JS. Overall survival as the outcome for randomized clinical trials with effective subsequent therapies. *J Clin Oncol* 2011;**29**(17):2439–42.
19. Taylor AL, Ziesche S, Yancy C, et al. Combination of isosorbide dinitrate and hydralazine in blacks with heart failure. *N Engl J Med* 2004;**351**(20):2049–57.
20. Lim E, Brown A, Helmy A, Mussa S, Altman DG. Composite outcomes in cardiovascular research: a survey of randomized trials. *Ann Intern Med* 2008;**149**(9):612–7.
21. Pauker SG, Kassirer JP. Decision analysis. *N Engl J Med* 1987;**316**(5):250–8.
22. Torrance GW. Preferences for health outcomes and cost-utility analysis. *Am J Manag Care* 1997;**3**(Suppl):S8–S20.
23. Burris III HA, Moore MJ, Andersen J, et al. Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: a randomized trial. *J Clin Oncol* 1997;**15**(6):2403–13.
24. Rini BI, Escudier B, Tomczak P, et al. Axitinib versus sorafenib as second-line therapy for metastatic renal cell carcinoma (mRCC): results of phase III AXIS trial. *J Clin Oncol* 2011 [abstr. 4503].
25. Cella D, Escudier B, Rini BI, et al. Patient-reported outcomes (PROs) in a phase III AXIS trial of axitinib versus sorafenib as second-line therapy for metastatic renal cell carcinoma (mRCC). *J Clin Oncol* 2011 [abstr. 4504].
26. Tomlinson G, Detsky AS. Composite end points in randomized trials: there is no free lunch. *JAMA* 2010;**303**(3):267–8.
27. Sullivan R, Peppercorn J, Sikora K, et al. Delivering affordable cancer care in high-income countries. *Lancet Oncol* 2011 Sep;**12**(10):933–80.