

## RESEARCH ARTICLE

# Verification of the Correlation between Progression-free Survival and Overall Survival Considering Magnitudes of Survival Post-progression in the Treatment of Four Types of Cancer

Li-Ya Liu, Hao Yu, Jian-Ling Bai, Ping Zeng, Dan-Dan Miao, Feng Chen\*

### Abstract

**Background:** With development and application of new and effective anti-cancer drugs, the median survival post-progression (SPP) is often prolonged, and the role of the median SPP on surrogacy performance should be considered. To evaluate the impact of the median SPP on the correlation between progression-free survival (PFS) and overall survival (OS), we performed simulations for treatment of four types of cancer, advanced gastric cancer (AGC), metastatic colorectal cancer (MCC), glioblastoma (GBM), and advanced non-small-cell lung cancer (ANSCLC). **Materials and Methods:** The effects of the median SPP on the statistical properties of OS and the correlation between PFS and OS were assessed. Further, comparisons were made between the surrogacy performance based on real data from meta-analyses and simulation results with similar scenarios. **Results:** The probability of a significant gain in OS and HR for OS was decreased by an increase of the SPP/OS ratio or by a decrease of observed treatment benefit for PFS. Similarly, for each of the four types of cancer, the correlation between PFS and OS was reduced as the median SPP increased from 2 to 12 months. Except for ANSCLC, for which the median SPP was equal to the true value, the simulated correlation between PFS and OS was consistent with the values derived from meta-analyses for the other three kinds of cancer. Further, for these three types of cancer, when the median SPP was controlled at a designated level (i.e., < 4 months for AGC, < 12 months for MCC, and < 6 months for GBM), the correlation between PFS and OS was strong; and the power of OS reached 34.9% at the minimum. **Conclusions:** PFS is an acceptable surrogate endpoint for OS under the condition of controlling SPPs for AGC, MCC, and GBM at their limit levels; a similar conclusion cannot be made for ANSCLC.

**Keywords:** Progression-free survival - overall survival - survival post-progression - surrogate endpoint - meta-analysis

*Asian Pac J Cancer Prev*, 16 (3), 1001-1006

### Introduction

In relation to evaluations of new anti-cancer drugs, overall survival (OS) is the standard for demonstrating a clinical benefit. However, measurement of OS requires a long follow-up period after disease progression, leading to extended drug development cycles and increases in research costs. Therefore, the evaluation of progression-free survival (PFS) for use in clinical trials has become an important consideration, and health authorities now recognize PFS as a useful endpoint. Between 2005 and 2007, 17% (9 of 53) of approvals for anti-cancer drugs by the U.S. Food and Drug Administration were based on trials with PFS endpoints (Sridhara et al., 2010).

Since additional choices for treatment are available after progression of cancer following prior treatments, the effect of subsequent chemotherapy on OS can result in underestimation of the efficacy of an experimental treatment (Lima et al., 2009; Soria et al., 2010). Indeed,

in recent randomized trials, improvements in PFS did not necessarily lead to an improved OS (Buyse et al., 2010; Saad et al., 2010b; Alimujiang et al., 2013). In clinical trials of first-line treatments, a moderate to large improvement in survival post-progression (SPP) resulted in a reduction of an OS benefit, even though the PFS benefit had been established. Therefore, it is likely that PFS is an appropriate endpoint for clinical trials, particularly those with a crossover design (Mok, 2011; Booth and Eisenhauer, 2012). In this setting, SPP has become prominent due to the availability of effective second- and third-line therapies.

For advanced gastric cancer (AGC), PFS was not an appropriate surrogate endpoint for OS when the median SPP was 4.54 months (Paoletti et al., 2013). For metastatic colorectal cancer (MCC), however, PFS strongly correlated with OS when the median SPP was 10.5 months (Giessen et al., 2013). This conclusion was confirmed in a study of surrogate endpoints for glioblastoma (GBM), in

which the median SPP was 5.75 months (Han et al., 2014). In contrast, in clinical trials for advanced non-small-cell lung cancer (ANSCLC), OS was selected as a primary endpoint when the median SPP was 9.94 months (Cheema and Burkes, 2013). Thus, for different types of cancer, meta-analyses can lead to various conclusions under the same or different median SPPs. Changes in the correlation between PFS and OS and the effect of treatment upon OS with increasing SPP based on the hazard ratio (HR) for PFS have been explored by simulation studies (Broglia and Berry, 2009). The probability of observing a statistically significant difference in OS, however, is dependent on the length of the median SPP and on the magnitude of the HR for PFS.

For the present effort, the primary objectives were: (i) to describe, for four types of cancer, changes of the power of treatment effects on OS and the correlation between PFS and OS with increasing SPP; and (ii) to find, for four types of cancer, the longest median SPP and to determine conditions for which the surrogate endpoint performance is appropriate. For each trial included in meta-analyses for four kinds of cancer, the median PFS or/and time to progression (TTP) from the control as well as HR for PFS or/and TTP were extracted, and the average median PFS or/and TTP and the overall HR for PFS were estimated and employed as simulation parameters. With these parameters, simulations were implemented to assess the probability of a statistically significant benefit in OS and possible surrogacy of PFS for OS through the association between these endpoints.

First described are the methods for generating the multi-trial datasets based on the parameters abstracted from meta-analyses of four kinds of cancer. Results from numerical studies are in Section 3, and a discussion is in Section 4.

## Materials and Methods

### Data generation

Correlations between time-to-event endpoints have been previously considered for cancer treatments (Michael and Schucany, 2002; Broglia and Berry, 2009; Fleischer et al., 2009; Fu et al., 2013). Two of the methods (Michael and Schucany, 2002; Fu et al., 2013) did not control for changes in the median SPP, and another method (Broglia and Berry, 2009) did not provide an assumption of dependency between TTP and OS. Thus, for the more general case to be considered, we used the method proposed by Fleischer et al. (2009). These investigators regarded the TTP to be exponentially distributed with parameter  $\lambda_1$ ,  $TTP \sim \text{Exp}(\lambda_1)$  and a second variable  $X$ , denoted as the time to death without tumor progression, was assumed to be exponentially distributed with parameter  $\lambda_2$ ,  $X \sim \text{Exp}(\lambda_2)$ . Based on the assumption that TTP and  $X$  are independent, PFS was given by the minimum of TTP and  $X$ ,  $PFS = \min(X, TTP)$ . SPP was also exponentially distributed with parameter  $\lambda_3$ ,  $SPP \sim \text{Exp}(\lambda_3)$ . Thus, OS was calculated as follows:

$$OS = \begin{cases} PFS & \text{if } PFS \neq TTP \\ TTP + SPP & \text{otherwise} \end{cases}$$

In this model, the hazard rate for OS was not constant or dependent on whether or not progression had occurred. In general, after progression, the hazard rate for OS was higher than before ( $\lambda_2 > \lambda_3$ ). Based on these assumptions, the correlation between OS and PFS was

$$\text{Corr}(PFS, OS) = \frac{\lambda_3}{\sqrt{\lambda_1^2 + 2\lambda_1\lambda_2 + \lambda_3^2}}$$

and the survival function of OS was given by

$$S_{OS}(t) = \frac{\lambda_1}{\lambda_1 + \lambda_2 - \lambda_3} \exp^{-\lambda_1 t} - \frac{\lambda_3 - \lambda_2}{\lambda_1 + \lambda_2 - \lambda_3} \exp^{-(\lambda_1 + \lambda_2)t}$$

All analyses were accomplished using the R version 3.0.2 statistical package (Team, 2012), except for graphs, which were prepared by use of STATA 12.0 (Stata Corp LP, College Station, Texas, USA). Prior to performing the primary simulations, we conducted pilot simulations to confirm that the desired data characteristics were precisely and consistently represented in our generated datasets.

### Simulation scenarios

The parameters of interest, abstracted and summarized from studies of surrogate endpoints for four kinds of cancer, are listed in Table 1. For each simulation, 2,000 trials were performed for assessment. In all simulations, arrival times were assumed to follow a uniform distribution, and patients were assumed to be accrued at the rate of 30 per month. For each type of cancer, a corresponding additional follow-up time after accrual was assumed to be complete. Based on the concrete clinical context of the trial (accrual time, follow-up time, the median PFS of control, and the HR for PFS) and a two-sided statistical significance level of 0.05, the sample size needed to achieve 80% power could be estimated; e.g., for AGC, the median PFS of control was assumed to be 2 months, the HR for PFS to be 0.4, the accrued time to be 2 months, and follow-up time to be 24 months. The estimated sample size was 23 for each arm. The simulation data were produced based on the same HR for PFS in each trial.

## Results

### Description of the statistical properties of OS

The simulation data were produced based on the same HR for PFS in each trial. In fixing the simulation parameter of the overall HR for PFS from four types of cancer, the impact of the median SPP on HR for OS and the power to detect the effect of treatment upon OS were considered. The log-rank test was applied to compare the survival distributions of two groups.

Both the trial-specific treatment effect on true endpoints and the power available for detecting a benefit in OS were reduced with increasing SPP (Table 2). In other words, these indicators were diluted by increases of SPP/OS. Based on the assumption of exponential distribution, SPP/OS could be calculated as follows:

$$SPP/OS = \frac{1}{2} \times \left( \frac{SPP}{SPP + PFS_{\text{control}}} + \frac{SPP}{SPP + PFS_{\text{control}}/HR} \right)$$

For the four cases, as the median SPP increased from 2 to 12 months, the changes of the power of OS and the HR for OS were relatively major, shifting from 86.05% (for MCC) to 15.16% (for GBM) and from 0.67 (for ANSCLC) to 0.95 (for GBM), respectively. For ANSCLC, the power of OS was 34% when the median SPP was 8 months. For AGC and GBM, however, the power of OS was 34% when the median SPP was 6 months. For MCC, the power of OS reached 41% when the median SPP was 12 months. For GBM and MCC, the HR for OS was 0.9 when the median SPP was 6 months. When the median SPP was 8 months in AGC, the HR for OS was also 0.9. However, for ANSCLC, the HR for OS was only 0.83 when the median SPP was 12 months.

Under the same proportion of SPP/OS, higher HRs of PFS corresponded to higher values of OS and higher HRs for OS. Based on the same median SPP, the differences of SPP/OS between three tumors (ANSCLC, AGC, and GBM) were small (<5%); thus, the values of OS and the HR for OS between these tumors were comparable. The same conclusion, however, was not applicable to MCCs because the differences of SPP/OS between MCCs and the other three tumors were large (>10%). When the median SPP was 2 months, the probability of a statistically significant difference in OS decreased as the HR for PFS approached 1.0. However, the results for MCC were not consistent, because the corresponding value of SPP/OS was lower than those of the others. In addition, for ANSCLC, AGC, and GBM, when the median SPP was 6 months, the treatment effect on OS decreased from 0.77 to 0.91 as the HR for PFS changed from 0.65 to 0.85. With the same median SPP, the HR of OS for MCC was higher than that for the other three tumors.

#### Correlation between PFS and OS

For these determinations, the simulation framework was different. First, the parameters of each trial, including the HR of PFS, the median PFS of the control group, and the total sample size were abstracted from the meta-analyses for each type of cancer. Second, the simulation

data of each trial were generated based on the above fixed parameters and the unfixed parameter of the median SPP. Third, to estimate trial-level surrogacy, the median times on the true endpoint and the surrogate endpoint within each trial were estimated, and the Spearman's rank correlation coefficient ( $r_s$ ) between two endpoints was calculated. The treatment effects on both endpoints within each trial were also calculated, and weighted linear regressions (WLRs) (weights equal to the sample size of the trial) were performed to evaluate relationships between effects.  $R^2$ , estimated from a weighted linear regression, was used to evaluate the model fitting accuracy. The 95% confidence intervals (CI) for  $r_s$  and  $R^2$  were obtained by use of the percentile bootstrap (Hall and Martin, 1989).

#### Correlation between the median PFS and OS

For AGC, MCC, GBM, and ANSCLC, the numbers of reviews were 20, 50, 91 and 21, respectively. From these reviews, the publications that reported both the HR for PFS and OS were 20, 28, 11, and 17, and the averages of the median SPP were 4.54, 10.5, 5.75, and 9.94 months, respectively. The impact of the median SPP on the correlation between the median PFS and OS was determined, and the true surrogacy performance based on the real data for meta-analyses of four kinds of cancer was compared with the simulation results.

The Spearman's rank correlation coefficients between the median PFS and OS for four scenarios decreased as the median SPP varied from 2 to 12 months (Figure 1). When the median SPP was  $\leq 12$  months, the median PFS of three types of cancer strongly correlated with the median OS. For AGC, however, the correlation between the median PFS and OS was strong if the median SPP was < 6 months.

When the median SPP was 4.54 months, the Spearman's rank correlation coefficient between the median PFS and OS for AGC ( $r_s=0.85$ , 95%CI, 0.852-0.854), measured by the individual-level association, was higher than the simulated results for  $r_s$  ( $r_s=0.76$ , 95%CI, 0.41-0.92) (Table 3). For MCC, when the median SPP was 10.5 months, the true value  $r_s$  between the median PFS

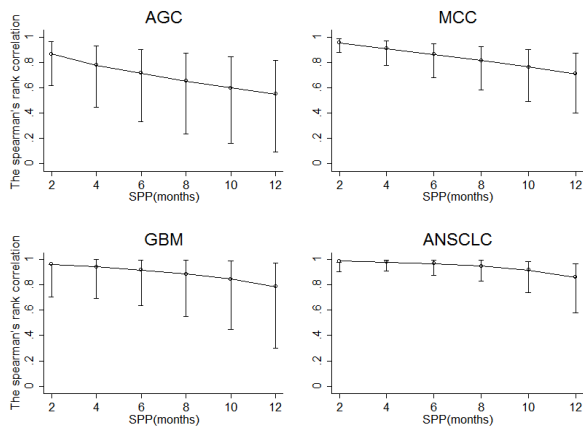
**Table 1. Simulation Scenarios**

Cancer	Fixed parameters				Unfixed parameters
	$\lambda_{1c}$	$\lambda_{1c}+\lambda_{2c}$	HR <sub>PFS</sub>	Follow-up time(months)	Median SPP(months)
AGC	Log(2)/3.75	Log(2)/3.5	0.79	24	2,4,6,8,10,12
MCC	Log(2)/9.2	Log(2)/7.65	0.88	40	
GBM	Log(2)/4.7	Log(2)/3.9	0.85	46	
ANSCLC	Log(2)/4.8	Log(2)/3.75	0.65	30	

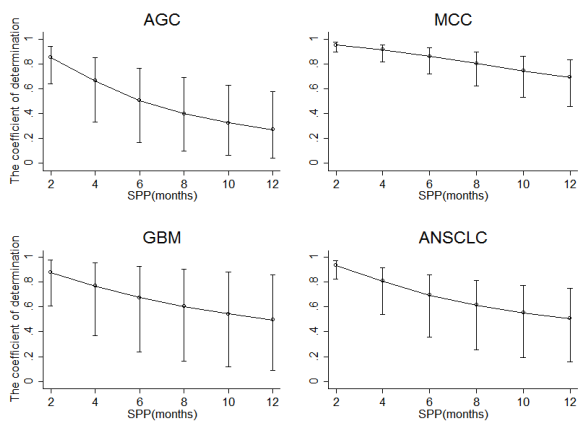
$\lambda_{1c}$ , Parameter of exponentially distributed of TTP;  $\lambda_{1c}+\lambda_{2c}$ , Parameter of exponentially distributed of PFS.

**Table 2. Statistical Properties for OS**

Median SPP	ANSCLC(HR <sub>PFS</sub> =0.65)			AGC(HR <sub>PFS</sub> =0.79)			GBM(HR <sub>PFS</sub> =0.85)			MCC(HR <sub>PFS</sub> =0.88)		
	SPP/OS(%)	HR <sub>OS</sub>	Power <sub>OS</sub>	SPP/OS(%)	HR <sub>OS</sub>	Power <sub>OS</sub>	SPP/OS(%)	HR <sub>OS</sub>	Power <sub>OS</sub>	SPP/OS(%)	HR <sub>OS</sub>	Power <sub>OS</sub>
2	30.28	0.67	81.2	33.73	0.81	74.1	32.12	0.86	71.72	19.72	0.88	86.05
4	46.3	0.72	64.33	50.39	0.85	53.17	48.6	0.89	53.53	32.93	0.89	80.84
6	56.28	0.77	45.38	60.34	0.88	34.3	58.63	0.91	34.88	42.4	0.9	71.3
8	63.11	0.79	34.13	66.96	0.9	25.35	65.38	0.93	24.32	49.53	0.91	59.83
10	68.09	0.82	27	71.69	0.91	20.35	70.24	0.94	19.13	55.08	0.92	49.9
12	71.88	0.83	22.58	75.23	0.91	16.82	73.9	0.95	15.16	59.53	0.93	40.6



**Figure 1. The Spearman's Rank Correlation Coefficient between the Median PFS and OS for Varied Median SPP**



**Figure 2. The Coefficient of Determination of WLR between the Treatment Effects on PFS and OS for Varied Median SPP**

and OS ( $r_s=0.86$ , 95%CI, 0.79-0.91) was higher than the simulated  $r_s$  ( $r_s=0.75$ , 95%CI, 0.47-0.90). For GBM, when the median SPP was 5.75 months, the correlation between the median PFS and OS ( $r_s=0.85$ , 95%CI, 0.68-0.94), which was estimated based on data from meta-analyses, was less than the  $r_s$  of the simulation ( $r_s=0.92$ , 95%CI, 0.64-0.99). For AGC, MCC, and GBM, differences between the simulated  $r_s$  and the true value were ~10%.

In summary, for three kinds of cancer, results of the simulations were consistent with the actual results, and a conclusion was that the median PFS strongly correlated with the median OS. For ANSCLC, however, the difference between the simulated and the true value, ~30%, was greater than that for the other three types.

*Correlation between the HR for PFS and OS*

Summaries of the simulation results for  $r_s$  between the HR for PFS and OS and  $R^2$  of the WLR with increasing SPP are presented in Figure 2. Only when the median SPP was controlled at 4 months for AGC, 12 months for MCC, and 6 months for GBM, the treatment effect on PFS showed a strong correlation with the treatment effect on OS. For these three types of cancer, when the median SPP exceeded these limits, the surrogacy measures performed poorly. The same conclusion, however, was not valid for ANSCLC, since its simulation results ( $R^2=0.55$ , 95%CI,

**Table 3. Comparison of Real data Estimation and the Simulation Results (The Spearman's Rank Correlation between the Median PFS and OS)**

Cancer	Real data estimation (95%CI)	Simulation results (95%CI)
AGC	0.85(0.852,0.854)	0.76(0.41,0.92)
MCC	0.86(0.79,0.91)	0.75(0.47,0.90)
GBM	0.85(0.68,0.94)	0.92(0.64,0.99)
ANSCLC	0.72(0.55,0.88)	0.92(0.73,0.99)

**Table 4. Comparison of Real data Estimation and the Simulation Results (Correlation between the Treatment Effect PFS and OS)**

Cancer	Real data estimation (95%CI)	Simulation results (95%CI)
AGC	0.61 <sup>a</sup> (0.04,1.00)	0.61 <sup>a</sup> (0.27,0.82)
MCC	0.87 <sup>b</sup> (0.67,0.93)	0.83 <sup>b</sup> (0.59,0.94)
GBM	0.92 <sup>a</sup> (0.71-0.99)	0.68 <sup>a</sup> (0.25,0.93)
ANSCLC	0.003 <sup>a</sup> (6.99e-05-0.35)	0.55 <sup>a</sup> (0.19-0.78)

<sup>a</sup>The coefficient of determination of WLS between the treatment effects on PFS and OS; <sup>b</sup>The Spearman's rank correlation between treatment effects on PFS and OS

0.19-0.78) deviated from the true value ( $R^2=0.003$ , 95%CI, 6.99e-05-0.35) (Table 4).

The correlations between the HR for PFS and OS based on the actual data were similar to those simulated for the same median SPP (Table 4). First, because model-based measures proposed by Burzykowski et al., which involved an error-in-variables linear regression that took into account the uncertainty about the estimated effects, were used to evaluate the surrogacy of PFS for AGC, the results of the two different methods (model-based measures based on true data vs WLR based on simulation data) can not be compared directly (Burzykowski et al., 2001).

Nevertheless, Shi et al. showed that the estimated performance of WLR was similar to model-based measures and that the estimated values of the error-in-variables linear regression could be viewed as the estimated values of WLR (Shi et al., 2011). When the median SPP was 4.54 months, the  $R^2$  of the linear regression, adjusted for estimation errors (0.61, 95%CI, 0.04-1.00), was equal to the  $R^2$  based on the simulation data (0.61, 95%CI, 0.27-0.82). Second, for MCC, the Spearman's rank correlation between treatment effects on PFS and OS was high across all studies ( $r_s=0.87$ , 95%CI, 0.67-0.93), which was consistent with the simulated  $r_s$  ( $r_s=0.83$ , 95%CI, 0.59-0.94) when the median SPP was 10.5 months.

Third, for GBM, the surrogacy measures performed well relative to the real data ( $R^2=0.92$ , 95%CI, 0.71-0.99). The same conclusion was reached in the simulation study ( $R^2=0.68$ , 95%CI, 0.25-0.93) when the median SPP was 5.75 months.

**Discussion**

Several investigators have assessed the association between SPP and OS (Saad et al., 2010a; Hayashi et al., 2012; Hayashi et al., 2013; Kawakami et al., 2013;



Petrelli and Barni, 2013; Shitara et al., 2013). Hayashi et al. concluded that the median OS was highly associated with the median SPP but not with the median PFS ( $r_s=0.94$  and  $0.51$ , respectively) and that there was only a weak association between the treatment benefits for PFS and OS ( $r_s=0.29$ ) for patients with ANSCLC who received second- or third-line chemotherapy (Hayashi et al., 2012; Hayashi et al., 2013). A similar result was reported for patients with AGC (Kawakami et al., 2013). The explanation for these conclusions was that the average median SPP was longer than the average median PFS. In other words, the ratio of SPP/OS was higher than that for PFS/OS.

For the present report, simulations were accomplished to assess the performance of the correlation between PFS and OS with increasing SPP based on different parameters from four types of cancer. For simplicity, we assumed that an increase in median PFS led to no change in SPP. Although, in real applications, this assumption may not be true, in a review of advanced ovarian cancer, an increase in median PFS generally resulted in little change in SPP (Sundar et al., 2012). Others drew the same conclusion for four types of metastatic cancer (Bowater et al., 2008). Another limitation for our study is that our simulations assumed that PFS and SPP follow exponential distributions. In general, however, our overall conclusions, based on other distributions in these simulations, were similar to those based on the experimental data.

The process of data generation developed here facilitates systematic simulations, in which a key factor, SPP, was varied in a controlled manner; the other factors, abstracted from meta-analyses for four types of cancer, were fixed. Our simulations demonstrated that the probability of a significant increase in OS depended on the size of SPP/OS and the magnitude of the observed treatment benefits for PFS. Such a statistically significant difference in OS is more frequently reported when there are significant gains in PFS, and the size of SPP/OS is small. The correlation between PFS and OS was reduced as the median SPP varied from 2 months to 12 months.

In addition, we can draw useful conclusions about the four types of cancer. For our simulations of three types, when the medians of SPP were controlled at selected levels (<4 months for AGC, <12 months for MCC, and <6 months for GBM), the correlations between the PFS and OS were strong, and the power of OS, which was reduced by SPP, reached 34.88% at the minimum. Thus, there is evidence that, for three types of cancer, PFS is an acceptable surrogate for the OS endpoint if the SPPs are control at their limit levels. However, Cheema et al. showed that, for advanced NSCLC, there are examples of improvement in PFS without an OS benefit, and an OS benefit without a PFS benefit, suggesting that factors other than preventing disease progression may be important in improving OS (Cheema and Burkes, 2013). Due to the simulation assumption of no treatment difference in SPP, the situation of an OS benefit without a PFS benefit cannot be achieved. This may be the main reason that the value used in the simulation for ANSCLC was larger than the value derived from meta-analyses. Our simulation results are generally consistent with those of published

meta-analyses, except that the clinical trials of an OS benefit without a PFS benefit were not included in the meta-analyses. Unlike AGC, GBM and MCC, similar conclusions could not be drawn for ANSCLC.

## Acknowledgements

The work was supported by the National Natural Science Foundation of China (81273184; 81473070), the National Natural Science Foundation of China Grant for Young Scientists (81302512), and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD). The authors thank Dr. Hill for the discussion of grammar, which improved the content of the article.

## References

- Alimujiang S, Zhang T, Han ZG, et al (2013). Epidermal growth factor receptor tyrosine kinase inhibitor versus placebo as maintenance therapy for advanced non- small-cell lung cancer: a meta-analysis of randomized controlled trials. *Asian Pac J Cancer Prev*, **14**, 2413-9.
- Booth CM, Eisenhauer EA (2012). Progression-Free Survival: Meaningful or Simply Measurable? *J Clin Oncol*, **30**, 1030-3.
- Bowater RJ, Bridge LJ, Lilford RJ (2008). The relationship between progression-free and post-progression survival in treating four types of metastatic cancer. *Cancer Lett*, **262**, 48-53.
- Broglio KR, Berry DA (2009). Detecting an overall survival benefit that is derived from progression-free survival. *J Natl Cancer Inst*, **101**, 1642-9.
- Burzykowski T, Molenberghs G, Buyse M, et al (2001). Validation of surrogate end points in multiple randomized clinical trials with failure time end points. *J R Stat Soc Ser C Appl Stat*, **50**, 405-22.
- Buyse M, Sargent DJ, Grothey A, et al (2010). Biomarkers and surrogate end points—the challenge of statistical validation. *Nat Rev Clin Oncol*, **7**, 309-17.
- Cheema PK, Burkes RL (2013). Overall survival should be the primary endpoint in clinical trials for advanced non-small-cell lung cancer. *Curr Oncol*, **20**, 150-60.
- Fleischer F, Gaschler-Markefski B, Bluhmki E (2009). A statistical model for the dependence between progression-free survival and overall survival. *Stat Med*, **28**, 2669-86.
- Fu H, Wang Y, Liu J, et al (2013). Joint modeling of progression-free survival and overall survival by a Bayesian normal induced copula estimation model. *Stat Med*, **32**, 240-54.
- Giessen C, Laubender RP, Ankerst DP, et al (2013). Progression-free survival as a surrogate endpoint for median overall survival in metastatic colorectal cancer: literature-based analysis from 50 randomized first-line trials. *Clin Cancer Res*, **19**, 225-35.
- Hall P, Martin MA (1989). A note on the accuracy of bootstrap percentile method confidence intervals for a quantile. *Stat Probabil Lett*, **8**, 197-200.
- Han K, Ren M, Wick W, et al (2014). Progression-free survival as a surrogate endpoint for overall survival in glioblastoma: a literature-based meta-analysis from 91 trials. *Neuro Oncol*, **16**, 696-706.
- Hayashi H, Okamoto I, Morita S, et al (2012). Postprogression survival for first-line chemotherapy of patients with advanced non-small-cell lung cancer. *Ann Oncol*, **23**, 1537-41.
- Hayashi H, Okamoto I, Taguri M, et al (2013). Postprogression survival in patients with advanced non-small-cell lung cancer who receive second-line or third-line chemotherapy. *Clin*

- Kawakami H, Okamoto I, Hayashi H, et al (2013). Postprogression survival for first-line chemotherapy in patients with advanced gastric cancer. *Eur J Cancer*, **49**, 3003-9.
- Lima JPD, Santos LVd, Sasse EC, et al (2009). Optimal duration of first-line chemotherapy for advanced non-small cell lung cancer: A systematic review with meta-analysis. *Eur J Cancer*, **45**, 601-7.
- Michael JR, Schucany WR (2002). The mixture approach for simulating bivariate distributions with specified correlations. *Am Stat*, **56**, 48-54.
- Mok TSK (2011). Personalized medicine in lung cancer: what we need to know. *Nat Rev Clin Oncol*, **8**, 661-8.
- Paoletti X, Oba K, Bang YJ, et al (2013). Progression-free survival as a surrogate for overall survival in advanced/recurrent gastric cancer trials: a meta-analysis. *J Natl Cancer Inst*, **105**, 1667-70.
- Petrelli F, Barni S (2013). Surrogate end points and postprogression survival in renal cell carcinoma: an analysis of first-line trials with targeted therapies. *Clin Genitourin Cancer*, **11**, 385-9.
- Saad ED, Katz A, Buyse M (2010a). Overall survival and post-progression survival in advanced breast cancer: a review of recent randomized clinical trials. *J Clin Oncol*, **28**, 1958-62.
- Saad ED, Katz A, Hoff PM, et al (2010b). Progression-free survival as surrogate and as true end point: insights from the breast and colorectal cancer literature. *Ann Oncol*, **21**, 7-12.
- Shi Q, Renfro LA, Bot BM, et al (2011). Comparative assessment of trial-level surrogacy measures for candidate time-to-event surrogate endpoints in clinical trials. *Comput Stat Data An*, **55**, 2748-57.
- Shitara K, Matsuo K, Muro K, et al (2013). Progression-free survival and post-progression survival in patients with advanced gastric cancer treated with first-line chemotherapy. *J Cancer Res Clin Oncol*, **139**, 1383-9.
- Soria JC, Massard C, Le Chevalier T (2010). Should progression-free survival be the primary measure of efficacy for advanced NSCLC therapy? *Ann Oncol*, **21**, 2324-32.
- Sridhara R, Johnson JR, Justice R, et al (2010). Review of oncology and hematology drug product approvals at the US Food and Drug Administration between July 2005 and December 2007. *J Natl Cancer Inst*, **102**, 230-43.
- Sundar S, Wu J, Hillaby K, et al (2012). A systematic review evaluating the relationship between progression free survival and post progression survival in advanced ovarian cancer. *Gynecol Oncol*, **125**, 493-9.
- Team RC (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.