

Association between treatment effects on disease progression end points and overall survival in clinical studies of patients with metastatic renal cell carcinoma

TE Delea^{*1}, A Khuu¹, DYC Heng², T Haas³ and D Soulières⁴

¹PAI (Policy Analysis Inc.), 4 Davis Court, Brookline, MA 02445, USA; ²Tom Baker Cancer Centre, University of Calgary, 1331 29th Street NW Calgary, Alberta T2N 4N2, Canada; ³Novartis Pharma, Basel CH-4002, Switzerland; ⁴Centre Hospitalier de l'Université de Montréal, Montreal, Quebec, Canada

BACKGROUND: The relationship between progression-free survival and time to progression (PFS/TTP) and overall survival (OS) has been demonstrated in a variety of solid tumours but not in metastatic renal cell carcinoma (mRCC).

METHODS: A systematic literature search was conducted to identify controlled trials of cytokine or targeted therapies for mRCC reporting information on treatment effects on PFS/TTP and OS for one or more comparison. The associations between treatment effects on PFS/TTP and OS were analysed using linear regression.

RESULTS: Thirty-one studies representing 10 943 patients, 75 treatment groups, and 41 comparisons were identified. The correlation coefficient between the negative log of the hazard ratio (HR) for PFS/TTP ($-\ln \text{HR}_{\text{PFS/TTP}}$) vs the negative log of the HR for OS ($-\ln \text{HR}_{\text{OS}}$) was 0.80 ($P < 0.0001$). In linear regression, the coefficient on $-\ln \text{HR}_{\text{PFS/TTP}}$ vs $-\ln \text{HR}_{\text{OS}}$ was 0.64 (95% confidence interval (CI): 0.47–0.81; $R^2 = 0.63$), suggesting each 10% relative risk reduction (RRR) for PFS/TTP was associated with a 6% RRR for OS. A 1-month gain in median PFS/TTP was associated with a 1.17-month gain in median OS (95% CI: 0.59, 1.76; $R^2 = 0.28$).

CONCLUSION: In trials of treatments for mRCC, treatment effects on PFS/TTP are strongly associated with treatment effects on OS. *British Journal of Cancer* (2012) **107**, 1059–1068. doi:10.1038/bjc.2012.367 www.bjccancer.com

Published online 30 August 2012

© 2012 Cancer Research UK

Keywords: metastatic; progression; renal cell carcinoma; survival

Overall survival (OS) is the gold standard for the assessment of efficacy in phase III trials of cancer therapies (Sargent and Hayes, 2008). However, use of OS as the primary end point requires that large numbers of patients be followed for an extended period of time to detect statistically significant differences between the treatment groups, thereby increasing study costs and delaying access to potentially beneficial treatments. Also, for ethical or practical reasons, patients randomised to control therapy are often allowed to crossover to study therapy, or receive an off-study investigational or other active treatment upon disease progression, thereby diluting the observed effect of study treatment on OS. These factors make measures of time to disease progression (e.g., progression-free survival (PFS) or time to progression (TTP)) attractive alternatives to OS. Measures of PFS/TTP generally require fewer patients and/or shorter follow-up to detect statistically significant differences between the treatment groups, and are not confounded by use of subsequent therapies upon disease progression. Moreover, PFS/TTP may be important measures *per se*, as disease progression may be associated with reduced patient health-related quality of life and increased healthcare costs.

The use of PFS/TTP as a valid surrogate end point for OS requires that treatment effects on OS can be reliably predicted

from observed treatment effects on PFS/TTP (Fleming and DeMets, 1996; Tang *et al*, 2007; Burzykowski *et al*, 2008). Although the association between treatment effects on PFS/TTP and treatment effects on OS has been examined in a variety of solid tumours (Louvét *et al*, 2001; Johnson *et al*, 2006; Buyse *et al*, 2007; Tang *et al*, 2007; Sherrill *et al*, 2008), it has not been rigorously examined in patients with metastatic renal cell carcinoma (mRCC) (Knox, 2008). The objective of this study was to evaluate the association between treatment effects on PFS/TTP and treatment effects on OS in randomized controlled trials of patients with mRCC.

MATERIALS AND METHODS

Search strategy

Medline was searched to identify clinical trials of interleukin-2, interferon (IFN)- α , axitinib, lapatinib, pazopanib, sunitinib, sorafenib, bevacizumab, everolimus, or temsirolimus in mRCC. The search was limited to studies published in English from January 1997 to January 2010, which reported data on survival and/or mortality in the abstract. Abstracts of identified studies were reviewed by two independent reviewers (AK and TED) to identify studies for which full-text articles would be retrieved and reviewed. This search was supplemented with hand searches of the American Society of Clinical Oncology (ASCO) and European Cancer Organisation (ECCO) Web sites for abstracts, posters, and/or presentations reported between January 2005 and December 2010, as well as reference lists of retrieved articles and prior

*Correspondence: TE Delea; E-mail: tdelea@pai2.com

Revised 31 May 2012; accepted 24 July 2012; published online 30 August 2012

meta-analyses and systematic reviews (Coppin *et al*, 2005, 2008; Coppin, 2008; Thompson Coon *et al*, 2010). Studies were included if they reported median PFS/TTP and median OS for two or more treatment groups or hazard ratios (HR)s for PFS/TTP and HRs for OS for one or more treatment comparison.

Data extraction

For each study selected for inclusion, information was extracted on first author, year of publication, prior treatment (treatment-naive, prior cytokine treatment, prior targeted treatment, mixed prior treatment), treatments evaluated, measures of PFS/TTP used (PFS or TTP), overall response rate (ORR) and whether trial patients were allowed to crossover to other study therapy or other active treatment after progression. For each treatment group, sample sizes for ORR, PFS/TTP and OS, median PFS/TTP, median OS, and corresponding 95% confidence intervals (CIs) for median PFS/TTP and median OS were recorded. Also recorded were HRs (and corresponding 95% CIs) for PFS/TTP and OS. Studies representing duplicate reports of the same trial were excluded, with the report least likely to have been impacted by crossover selected for the analysis (e.g., based on rank-preserving structural failure time (RPSFT) models or inverse probability of censoring weighted (IPCW) analyses, with patients censored at crossover, or at study unblinding before crossover).

Measures of treatment effect

Two measures of treatment effects on PFS/TTP and OS were analysed: (1) the absolute differences between the treatment groups in median PFS/TTP (in months) *vs* the absolute differences between groups in median OS (in months) and (2) the negative of the natural log of the HR for PFS/TTP ($-\ln \text{HR}_{\text{PFS/TTP}}$) *vs* the negative of the natural log of the HR for OS ($-\ln \text{HR}_{\text{OS}}$). For small treatment effects (relative risk reduction (RRR) $\leq \pm 30\%$), the $-\ln$ (HR) is approximately equal to the RRR. The HR is frequently used as the primary measure of treatment effect in controlled clinical trials. However, the median survival for each treatment group is also frequently reported. The advantage of the HR is that it reflects a comparison of hazards for the entirety of the survival distribution, whereas the difference in medians reflects a comparison at a single point on the distribution. On the other hand, if treatment has no effect on post-progression survival, the gain in median PFS will be an unbiased estimate of the gain in median OS regardless of the duration of PPS, whereas the HR for OS will tend to be greater than that for PFS, and the degree of difference will depend on the duration of PPS (Broglia and Berry, 2009). Both of these measures have been used in prior studies of the association between PFS/TTP and OS in other cancers (Louvret *et al*, 2001; Johnson *et al*, 2006; Buyse *et al*, 2007; Tang *et al*, 2007; Sherrill *et al*, 2008).

For studies that reported both PFS and TTP, we recorded PFS. For those that reported TTP only, we combined TTP results with those for PFS from other studies. Although TTP and PFS are different measures, in the setting of mRCC, wherein survival is short and death due to reasons other than mRCC is rare, the HRs and differences in median survival are likely similar for TTP and PFS. TTP and PFS have been combined in prior studies of the association between disease progression end points and OS in other tumours (Tang *et al*, 2007; Sherrill *et al*, 2008). In an evaluation of studies of metastatic breast cancer patients in which TTP and PFS were not combined but were analysed separately, the associations between TTP and PFS on the one hand and OS on the other were similar (Burzykowski *et al*, 2008).

For studies that did not report HRs for PFS/TTP or OS, HRs were estimated using data from Kaplan–Meier curves or numbers of events and log-rank statistics (Tierney *et al*, 2007). For treatment arms for which median OS was not reached but for

which Kaplan–Meier survival curves were reported, median survival was estimated by fitting Weibull survival functions to reported Kaplan–Meier curves (Carroll, 2003). For studies that included more than the two treatment groups, treatment effects on PFS/TTP and OS were calculated for $k-1$ of k potential comparisons (e.g., for a study with treatments A, B, and C, we calculated two comparisons: A *vs* B and A *vs* C). In cases with an obvious control arm, this arm was selected as the reference group for all comparisons.

Statistical analyses

The possibility of publication bias was assessed by examining asymmetry of a funnel plot of estimates of $-\ln \text{HR}_{\text{OS}}$ *vs* its s.e. and using Egger's test (Egger *et al*, 1997). Pearson correlation coefficients between treatment effects on PFS/TTP and treatment effects on OS were calculated. In calculating Pearson correlation coefficients, each treatment comparison was weighted by the sum of the number of patients in the two treatment groups compared (non-parametric Spearman correlation coefficient were also calculated and were virtually the same as Pearson correlations and were not reported). The associations between treatment effects on PFS/TTP and treatment effects on OS also were examined using ordinary least squares regression with each treatment comparison weighted by the sum of the number of patients in the two treatment groups. Ninety-five percent prediction limits were calculated from weighted regressions using the mean number of patients per comparison as a weight.

Analyses were conducted separately by prior treatment for mRCC (none *vs* any), PFS/TTP end point reported (PFS *vs* TTP), whether crossover to active therapy after disease progression was allowed, and year of publication. Analyses also were conducted using all potential comparisons from trials with more than two treatment arms (e.g., for a study with treatments A, B, and C, we calculated three comparisons: A *vs* B, A *vs* C, and B *vs* C), setting the intercept terms in regression models to zero, and using all comparisons and setting intercept terms to zero. An analysis also was conducted to assess the association between ORR and OS for studies that reported ORR. In this analysis, comparisons involving arms with zero or missing response data were excluded. The treatment effect on ORR was measured in terms of the natural log of the relative risk of the response ($\ln \text{RR}_{\text{ORR}}$) and the treatment effect on OS was measured in terms of the $-\ln \text{HR}_{\text{OS}}$. An analysis also was conducted of the association between the $-\ln \text{HR}_{\text{PFS/TTP}}$ and $-\ln \text{HR}_{\text{OS}}$ in which each comparison was weighted by inverse of the variance of the $-\ln \text{HR}_{\text{OS}}$ rather than the number of subjects.

RESULTS

Search results

The search identified 235 potential studies. From these, as well as hand searches of reference lists of retrieved studies, ASCO and ECR web sites, and prior systematic reviews, a total of 31 studies were identified, representing 10 943 patients, 75 treatment groups, and 41 potential treatment comparisons that reported sufficient information for either the analysis of correlation between differences in median PFS/TTP and differences in median OS or between $-\ln \text{HR}_{\text{PFS/TTP}}$ and $-\ln \text{HR}_{\text{OS}}$ (Table 1) (Kruit *et al*, 1997; Negrier *et al*, 1998, 2000, 2007, 2008; Medical Research Council Renal Centre Collaborators, 1999; Pyrhonen *et al*, 1999; Motzer *et al*, 2000, 2007, 2008, 2009, 2010; Atzpodien *et al*, 2001, 2002, 2004, 2006; Dutcher *et al*, 2003; Yang *et al*, 2003; Atkins *et al*, 2004; Aass *et al*, 2005; Donskov *et al*, 2005; McDermott *et al*, 2005; Tannir *et al*, 2006; Bukowski *et al*, 2007; Escudier *et al*, 2007a, b; Hudes *et al*, 2007; Amato *et al*, 2008; Figlin *et al*, 2008; Sternberg

Table 1 Comparison included in analysis

Author (Year)	Prior treatment	PFS/ TTP end point	Cross-over	Treatment 1						Treatment 2						OS	PFS	HR (95% CI)				
				Treatment	N	ORR (%)	PFS	OS	Median months (95% CI)	Treatment	N	ORR (%)	PFS	OS	Median months (95% CI)				N for comparison ^a	RR ORR ^b	PFS	OS
Assi (2005)	None	PFS		IFNA	161	0.0	3.2 (2.8-4.6)	13.2 (11.0-17.8)	IFNA+CRA	159	0.0	5.1 (3.8-6.5)	17.3 (13.1-23.1)	320	1.01	1.9	4.1	0.67 (0.53-0.85)	0.78 (0.61-1.00)			
Arzoo (2008) ^c	Mixed	PFS	y	IFNA+IMA	2	0.0	2.5 (2.4-2.6)	9.0 (2.4-15.6)	IFNA+	2	0.0	9.6 (8.6-10.6)	21.0 (14.1-27.9)	3	1.00	7.1	12.0					
				IFNA+IMA	2	0.0	2.5 (2.4-2.6)	9.0 (2.4-15.6)	IFNA+GEF	12	25.0	4.3 (1.1-16.0)	11.4 (1.1-29.1)	13	1.14	1.8	2.4					
Atkins (2004)	Cytokine	TTP		TEM+25 mg	36	5.6	6.3 (3.6-7.8)	13.8 (9.0-18.7)	TEM+75 mg	38	7.9	6.7 (3.5-8.5)	11.0 (8.6-18.6)	56	1.27	0.4	-2.8	1.02 (0.62-1.67) ^d	1.02 (0.62-1.67) ^d			
				TEM+25 mg	36	5.6	6.3 (3.6-7.8)	13.8 (9.0-18.7)	TEM+250 mg	37	8.1	5.2 (3.7-7.4)	17.5 (12.0-24.6)	55	1.30	-1.1	3.7	1.00 (0.61-1.65) ^d	1.00 (0.61-1.65) ^d			
Atzpodien (2001) ^e	Mixed	PFS	y	TAM	37	0.0	6.0 (4.0-8.0)	13.0 (8.0-18.0)	IFNA+IL+	41	39.0	7.0 (3.0-11.0)	24.0 (11.0-37.0)	78	15.42	1.0	11.0	0.31 (0.19-0.52) ^d	0.56 (0.32-0.97) ^d			
Atzpodien (2002)	Mixed	PFS		IFNA+IL	97	0.0	6.0 (4.0-8.0)	17.0 (13.0-21.0)	IFNA+IL	260	0.0	6.0 (4.0-8.0)	19.0 (16.0-22.0)	227	0.38	0.0	2.0	0.84 (0.63-1.12) ^d	1.03 (0.80-1.31) ^d			
				IFNA+IL+5FU	260	0.0	6.0 (4.0-8.0)	19.0 (16.0-22.0)	IFNA+IL+5FU+CRA	86	0.0	12.0 (8.0-16.0)	32.0 (20.0-44.0)	216	2.98	6.0	13.0	0.69 (0.51-0.94) ^d	0.81 (0.47-0.81) ^d			
Atzpodien (2004)	Mixed	PFS		IFNA+VBL	63	20.6	5.0 (3.6-7.8)	16.0 (16.0-22.0)	IFNA+IL	132	31.1	6.0 (3.6-7.8)	25.0 (20.0-44.0)	129	1.46	1.0	9.0	0.66 (0.46-0.95) ^d	0.78 (0.53-1.14) ^d			
Atzpodien (2006)	Mixed	PFS		IFNA+IL	132	31.1	6.0 (3.6-7.8)	16.0 (16.0-22.0)	IFNA+IL+5FU	146	26.0	7.0 (3.6-7.8)	27.0 (20.0-44.0)	212	0.84	1.0	2.0	1.00 (0.74-1.34) ^d	0.94 (0.70-1.27) ^d			
				IFNA+IL+CRA	78	29.5	5.0 (3.6-7.8)	22.0 (16.0-22.0)	IFNA+IL+CRA	65	30.8	4.0 (3.6-7.8)	18.0 (16.0-22.0)	143	1.04	-1.0	-4.0	0.82 (0.55-1.24) ^d	1.26 (0.82-1.95) ^d			
Bulcowski (2007)	None	PFS		IFNA+IL+SFU+CRA	116	19.0	0.0 (4.0-8.0)	18.0 (13.0-21.0)	IFNA+IL+SFU+CRA	120	26.7	4.0 (3.0-11.0)	16.0 (11.0-37.0)	236	1.41	4.0	-2.0	0.74 (0.55-0.99) ^d	0.96 (0.69-1.32) ^d			
Donskov (2005)	Mixed	PFS		BEV+PLC	53	13.2	8.5 (4.0-8.0)	38.0 (13.0-21.0)	CRA+CAP	51	14.0	9.9 (4.0-8.0)	20.0 (16.0-22.0)	104	1.06	1.4	-18.0 ^f	0.86 (0.50-1.49)	1.57 (0.84-2.94)			
				IL	30	3.3	2.2 (4.0-8.0)	11.4 (13.0-21.0)	IL+HD	33	12.1	4.5 (8.0-16.0)	18.3 (20.0-44.0)	63	3.64	2.3	6.9	0.68 (0.40-1.13) ^d	0.61 (0.36-1.05)			
				IL	20	15.0	4.5 (4.0-8.0)	12.9 (13.0-21.0)	IL+HD	21	14.3	4.3 (8.0-16.0)	13.2 (20.0-44.0)	41	0.95	-0.2	0.3	1.15 (0.61-2.19) ^d	1.24 (0.64-2.39) ^d			
Dutcher (2003)	None	PFS		IFNG	39	0.0	1.4 (4.0-6.0)	7.0 (13.0-21.1)	IFNA+IFNG	49	10.2	2.9 (10.0-12.0)	10.9 (19.5-NR)	88	1.5	1.5	3.9	0.59 (0.37-0.92) ^d	0.84 (0.52-1.35) ^d			
Escudier (2007) ^g	None	PFS	y	PLC	452	1.8	2.8 (4.0-6.0)	15.9 (9.7-21.1)	SOR	451	9.8	5.5 (10.0-12.0)	19.3 (28.1)	903	5.51	2.7	3.4	0.51 (0.43-0.60)	0.77 (0.63-0.95)			
Escudier (2007) ^h	None	PFS	y	IFNA+PLC	322	12.8	5.5 (4.0-6.0)	21.3 (9.7-21.1)	IFNA+BEV	327	31.4	10.4 (4.0-6.0)	23.3 (19.5-NR)	649	2.45	4.9	2.0	0.63 (0.52-0.75)	0.79 (0.62-1.02)			
Figlin (2008)	None	PFS	y	IFNA	162	12.0	5.0 (4.0-6.0)	14.1 (9.7-21.1)	SUT	193	46.9	11.0 (10.0-12.0)	28.1 (19.5-NR)	355	3.91	6.0	14.0	0.43 (0.32-0.54)	0.65 (0.48-0.87)			
Metzer (2009) ⁱ	None	PFS		IFNA	502	14.0	5.5 (4.0-6.0)	18.9 (17.0-23.2)	IFNA+IL	504	21.0	5.3 (4.8-6.0)	18.6 (16.5-20.6)	1006	1.50	-0.2	-0.3	1.02 (0.89-1.16)	1.05 (0.09-1.21)			
Gene (2010)	None	PFS		IFNA	207	4.8	3.1 (2.2-3.8)	7.3 (6.1-8.8)	TEM	209	8.6	5.5 (3.9-7.0)	10.9 (8.6-12.7)	3115	1.79	2.4	3.6	0.67 (0.55-0.82) ^d	0.73 (0.58-0.92)			
Hudes (2007)	None	PFS		TEM	209	8.6	5.5 (3.9-7.0)	10.9 (8.6-12.7)	IFNA+TEM	210	8.1	4.7 (3.9-5.8)	8.4 (6.6-10.3)	3145	0.94	-0.8	-2.5	1.02 (0.83-1.26) ^d	1.32 (0.62-1.98) ^d			
Jaxson (1998)	None	PFS		IL	30	6.7	5.1 (3.9-7.0)	14.6 (8.6-12.7)	IFNA+IL	30	0.0	3.7 (3.9-5.8)	12.5 (6.6-10.3)	60	0.00	-1.4	-2.1	1.57 (0.88-2.63) ^d	1.11 (0.62-1.98) ^d			
Kruit (1997)	None	TTP		IL	17	23.5	6.0 (3.9-7.0)	13.9 (8.6-12.7)	IFNA+IL	55	37.3	5.9 (3.9-5.8)	16.9 (12.5)	72	1.58	-0.1	3.0					
McDemott (2005)	None	PFS		IFNA+IL	96	9.9	3.1 (2.2-3.8)	13.0 (8.6-12.7)	IL2	96	23.2	3.1 (3.9-5.8)	17.1 (12.5)	192	2.34	0.0	4.1	0.78 (0.58-1.06) ^d	0.81 (0.59-1.13)			
Metzer (2000)	None	PFS		IFNA	145	6.8	5.3 (3.9-7.0)	15.0 (8.6-12.7)	IFNA+CRA	139	12.4	4.7 (3.9-5.8)	15.0 (8.6-12.7)	284	1.82	-0.6	0.0	0.76 (0.60-0.97) ^d	0.82 (0.64-1.06) ^d			
Metzer (2008)	Targeted	PFS	y	PLC	138	0.0	1.9 (3.9-5.8)	10.0 (6.6-10.3)	EVE	272	1.0	4.9 (3.9-5.8)	14.8 (6.6-10.3)	410	3.00	3.0	4.8	0.33 (0.25-0.43)	0.55 (0.31-0.97)			
MRCRCC (1999)	Mixed	PFS		MPA	168	2.0	3.0 (2.2-3.8)	6.0 (6.1-8.8)	IFNA	167	14.0	4.0 (3.9-5.8)	8.5 (6.6-10.3)	335	7.00	1.0	2.5	0.72 (0.56-0.92)	0.72 (0.55-0.94)			
Negrier (1998)	None	PFS		IL	138	7.7	4.0 (3.9-5.8)	12.0 (8.6-12.7)	IFNA	147	8.1	4.0 (3.9-5.8)	13.0 (8.6-12.7)	2115	1.05	0.0	1.0	0.87 (0.69-1.12) ^d	1.13 (0.85-1.51) ^d			
				IFNA	147	8.1	4.0 (3.9-5.8)	13.0 (8.6-12.7)	IFNA+IL	140	21.7	5.0 (3.9-5.8)	17.0 (12.5)	2135	2.68	1.0	4.0	0.75 (0.59-0.96) ^d	0.82 (0.62-1.09) ^d			

Table 1 (Continued)

Author (Year)	Prior treatment	PFS/TTP end point	Cross-over	Treatment 1						Treatment 2								
				Treatment	N	ORR (%)	PFS	OS	Median months (95% CI)	Treatment	N	ORR (%)	PFS	OS	Median months (95% CI)			
Negrier (2000)	None	PFS	IFNA + IL	70	1.4	2.0	13.0	IFNA + IL + OTH	61	8.2	3.0	13.0	131	5.74	1.0	0.0	0.81 (0.56-1.17) ^d	1.04 (0.69-1.57) ^d
Negrier (2007)	None	PFS	Non-IFNA	248	3.2	3.2 (3.0-3.9)	15.1 (13.2-17.8)	IFNA	244	7.0	3.5 (3.1-5.3)	15.4 (14.5-18.7)	246	2.16	0.3	0.3	1.00 (0.81-1.24)	
	None	PFS	Non-IL	245	3.3	3.2 (3.0-3.8)	14.9 (12.8-18.7)	IL	247	6.9	3.5 (3.1-5.3)	15.7 (14.4-18.4)	246	2.11	0.3	0.8	1.07 (0.87-1.33)	
Negrier (2008)	None	PFS	IFNA + IL-IV	80	17.9	7.2 (6.0-9.6)	37.7 (28.2-55.6)	IFNA + IL-SC	75	21.3	6.2 (5.1-8.5)	30.1 (25.1-34.5)	155	1.19	-1.0	-7.6	1.16 (0.83-1.63)	1.20 (0.78-1.83)
Purtonen (1999)	Mixed	PFS	VBL	81	2.4	2.1	8.7	IFNA + VBL	79	16.5	3.0	15.5	160	6.88	0.9	6.9	0.52 (0.38-0.72) ^d	0.66 (0.48-0.92) ^d
Ravaud (2008)	Mixed	TTP	HT	207	0.5	3.5	9.9	LAP	209	1.4	3.5	10.8	416	2.80	0.0	0.9	0.94 (0.75-1.18)	0.88 (0.68-1.12)
Rini (2010)	None	PFS	IFNA	363	13.1	5.2 (3.1-5.6)	17.4 (14.4, 20.0)	BEV + IFNA	369	25.5	8.5 (7.5, 9.7)	18.3 (16.3, 22.5)	732	1.95	3.3	0.9	0.71 (0.61-0.83)	0.86 (0.73-1.01)
Sternberg (2009/10) ^k	Mixed	PFS	PLC	145	3.4	4.2	18.7	PAZ	290	30.3	9.2	21.1	435	8.80	5.0	2.4	0.46 (0.34-0.62)	0.43 (0.13-1.44)
Tamir (2006)	Mixed	PFS	IFNA-0.5 MU bid	59	6.8	3.7 (2.0-5.5)	25.5 (15.9-NR)	IFNA-5 MU qd	59	6.8	3.4 (2.1-5.5)	17.5 (14.7-26.6)	118	1.00	-0.3	-8.0	1.15 (0.79-1.66)	0.96 (0.59-1.54)
Yang (2003)	Mixed	TTP	PLC	40	0.0	2.5	13.2	BEV-3mg	37	0.0	3.0	15.1	57	0.5	1.9	1.9	0.79 ^d	0.82 (0.50-1.34)
			PLC	40	0.0	2.5	13.2	BEV-10mg	39	10.3	4.8	15.5	59	2.3	2.3	2.3	0.39 ^d	0.81 (0.50-1.31)

Abbreviations: BEV = bevacizumab; CAP = capecitabine; CI = confidence interval; CRA = 13-cis-retinoic acid; ERL = erlotinib; EVE = everolimus; 5FU = 5-fluorouracil; GEF = gefitinib; HD = histamine dithydrochloride; HR = hazard ratio; HT = hormonal therapy; IFNA = interferon- α ; IFNG = Interferon- γ ; IL = interleukin-2; IL-INH = inhaled interleukin-2; IL-IV = intravenous interleukin-2; IL-SC = subcutaneous interleukin-2; IMA = imatinib; LAP = lapatinib; MPA = medroxyprogesterone acetate; MRCRCC = Medical Research Council Renal Cancer Collaborators; NR = not reached; ORR = overall response rate; OS = overall survival; PAZ = pazopanib; PLC = placebo; PFS = progression-free survival; RD = recommended dose; RR = relative risk; TAM = tamoxifen; TTP = time to disease progression; TEM = temsirolimus; TTP = time to progression; SOR = sorafenib; SUT = sunitinib; VBL = vinblastine. For each trial with number of treatment arms (k), k - 1 comparisons shown. ^aN for comparison adjusted for number of times each arm contributes to a comparison. ^bNo continuity correction. ^cTwo patients from the imatinib arm crossed over to the gefitinib arm after experiencing disease progression. ^dHR (95% CI) not reported; estimated from Kaplan-Meier survival curves. ^eSixteen patients from tamoxifen arm crossed over to immunotherapy regimen after progressive disease. ^fMedians not reported; values reported were obtained by fitting Weibull survival functions to published survival curve data. Values base on Weibull functions were used in sensitivity analysis only. ^gCrossover permitted from placebo to sorafenib after single planned analysis of PFS showed a statistically significant benefit of sorafenib over placebo. ^hCrossover not allowed, but analyses potentially confounded by off-label use of bevacizumab. ⁱIFN- γ patients allowed to crossover to sunitinib following documented disease progression. ^j> 50% patients in both the groups received other therapies post progression; Ns and results for OS based on patients who did not receive any post-study treatment. ^kPlacebo patients allowed to crossover to everolimus after documented progression. Median OS was based on analysis using rank-preserving structural failure time model (Korhonen (2009)); HR for OS was based on analysis using inverse probability of censoring weighted (IPCW) analysis (Weiderkehr (2009)). ^lPlacebo patients allowed to crossover to pazopanib after documented progression. HR for OS was based on analysis using rank-preserving structural failure time model (Sternberg (2010)); 95% CI: for HR based P-value. Median OS based on unadjusted analysis.

et al, 2009, 2010a, b; Gore *et al*, 2010; Korhonen and Malangone, 2010; Rini *et al*, 2010; Korhonen *et al*, 2011; Wiederkehr *et al*, 2011). The great majority of the studies that were excluded for lack of information on both PFS or TTP and OS.

Study characteristics

Fifteen studies (48%) were published before 2006; 17 (55%) were in treatment-naïve patients; seven (23%) allowed crossover to active treatment after disease progression. Ten studies (32%) included one or more targeted treatments. For the phase III trial of sunitinib vs IFN, several analyses of OS were conducted, which might be differentially affected by crossover from IFN to sunitinib. In our base case, we used the results from the analysis in which patients who received any post-study treatment were excluded (HR 0.647, 95% CI: 0.483–0.870, median OS 28.1 months vs 14.1 months for sunitinib ($n=193$) vs IFN ($n=162$)) (Figlin *et al*, 2008; Motzer *et al*, 2009). The HR from this analysis was virtually identical to that reported in the interim analysis of the ITT population before patients were allowed to crossover (HR 0.65, 95% CI: 0.449–0.942, median OS not reached for sunitinib ($n=375$) or IFN ($n=375$)) (Motzer *et al*, 2007). We used the values from the former because median OS was not reached for the latter. For the phase III trial of everolimus, placebo patients were allowed to crossover to everolimus after documented progression (McDermott *et al*, 2005; Korhonen and Malangone, 2010; Motzer *et al*, 2010; Korhonen *et al*, 2011; Wiederkehr *et al*, 2011). For this study we used median OS based on analysis using the RPSFT model to control for crossover (Korhonen and Malangone, 2010; Korhonen *et al*, 2011); the HR for OS was based on analysis using IPCW analysis (Wiederkehr *et al*, 2011). For phase III trial of pazopanib, the HR for OS was based on the analysis using RPSFT to control for crossover (Sternberg *et al*, 2010b).

Thirty studies representing 40 treatment comparisons reported median PFS/TTP and median OS for one or more comparisons. Median OS was estimated based on fitting of Weibull survival functions to Kaplan–Meier curves for one treatment arm (bevacizumab plus placebo arm in the study by Bukowski *et al* (2007). This arm was represented in one comparison. Across all studies, median PFS/TTP and OS averaged 4.9 and 16.6 months, respectively. The median difference between the treatment groups in PFS/TTP averaged 1.4 months (s.d. 2.1 months, range -1.4 – 7.1 months); the median difference between the treatment groups in OS averaged 2.0 months (s.d. 5.7 months, range -18.0 – 14.0 months). Twenty-eight studies representing 36 treatment comparisons reported sufficient information for the analysis of $-\ln HR_{PFS/TTP}$ vs $-\ln HR_{OS}$. The $-\ln HR_{PFS/TTP}$ averaged 0.31 (s.d. 0.36, range -0.42 – 1.17); the $-\ln HR_{OS}$ averaged 0.15 (s.d. 0.27, range -0.45 – 0.84). HRs were estimated from Kaplan–Meier curves or log-rank statistics and event counts in 40 treatment arms represented in 23 comparisons.

The funnel plot of estimates of $-\ln HR_{OS}$ vs corresponding s.e.'s provided no strong evidence of publication bias (Figure 1). The estimated intercept on a regression of the inverse s.d. vs the standardized effect size (Egger's test) was 0.17 ($P=0.7658$); this also suggests no evidence of publication bias.

Association between treatment effects on PFS/TTP and treatment effects on OS

The weighted Pearson correlation coefficient for the difference in median PFS/TTP and the difference in median OS was 0.54 ($P=0.0002$). In linear regression analysis, a 1-month difference in median PFS/TTP was associated with a 1.17-month difference in median OS (95% CI: 0.59, 1.76; adjusted $R^2=0.28$) (Figure 2).

The weighted Pearson correlation coefficient for $-\ln HR_{PFS/TTP}$ and $-\ln HR_{OS}$ was 0.80 ($P<0.0001$). The coefficient on $-\ln HR_{PFS/TTP}$ vs $-\ln HR_{OS}$ was 0.64 (95% CI: 0.47, 0.81; adjusted

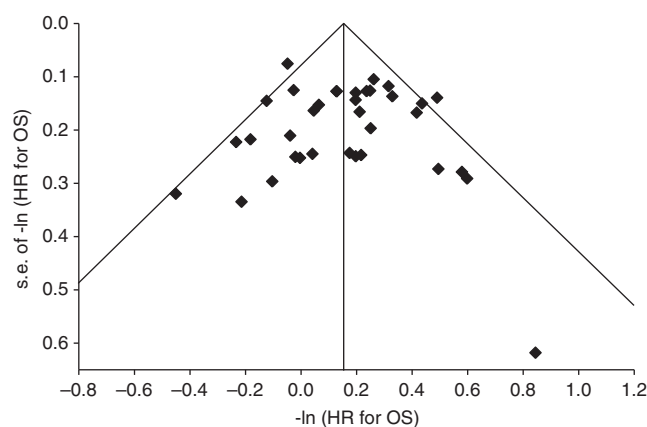


Figure 1 Funnel plot of negative log of HR for OS vs corresponding s.e. for each comparison. The funnel plot shows an assessment of publication bias. If there is no publication bias, the coordinates should be scattered symmetrically around the pooled estimate. The vertical line represents the fixed effects pooled estimate of $-\ln HR_{OS}$. The diagonal lines describing the funnel represent the 95% CI for each value of the s.e. The outlier is the coordinate for the pivotal study of pazopanib ($-\ln HR_{OS}=0.84$, s.e. ($-\ln HR_{OS}$)=0.62) (Sternberg *et al*, 2010b). The relatively high degree of imprecision associated with this estimate was due to the RPSFT method used to analyse OS to control for crossover.

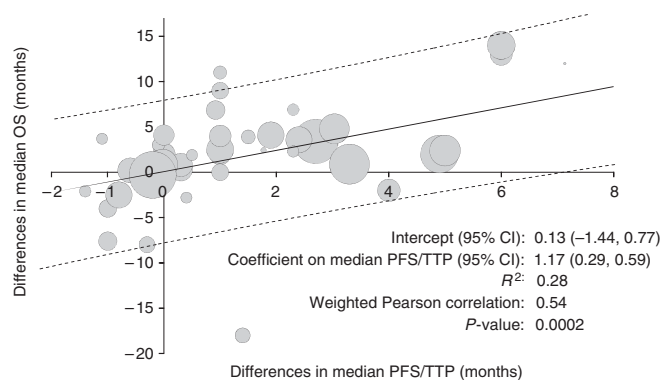


Figure 2 Association between differences in median PFS/TTP and differences in median OS. Abbreviation: R^2 = adjusted R -squared. Area of bubbles is proportional to the number of patients. Solid line is predicted value. Dashed lines are prediction intervals.

$R^2=0.63$) (Figure 3), suggesting that a 10% increase in the RRR for PFS/TTP is associated with an $\sim 6\%$ increase in the RRR for OS.

Subgroup and sensitivity analyses

Results in subgroups of studies are presented in Table 2. The correlation between treatment effects on PFS/TTP and treatment effects on OS was greater in studies that did not allow/require crossover, studies that used PFS rather than TTP, and in studies published before 2005 (studies before 2005 were less likely to have allowed crossover). There was no significant association between the treatment effects on PFS/TTP and OS in the subset of trials of vascular endothelial growth factor (VEGF) inhibitors, although there was a trend in the linear regression for $-\ln HR_{PFS/TTP}$ vs $-\ln HR_{OS}$ ($P=0.0510$). Results were similar to those of primary analysis when all potential comparisons from trials with multiple treatment arms were included. The adjusted R^2 for the analysis of differences in median PFS/TTP vs differences in median OS was greater with the exclusion of the study by Bukowski *et al* (2007), a randomized phase II trial comparing bevacizumab plus erlotinib

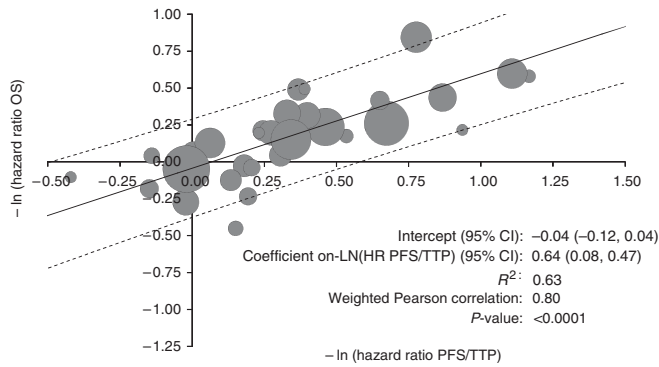


Figure 3 Association between negative log of HR for PFS/TTP and negative log of HR for OS. Abbreviation: R^2 = adjusted R -squared. Area of bubbles is proportional to the number of patients. Solid line is predicted value. Dashed lines are 95% prediction intervals.

vs bevacizumab plus placebo that was an extreme outlier, with a positive treatment effect on PFS/TTP and a negative treatment effect on OS (difference in median PFS/TTP 1.4 vs difference in median OS -18.0 (the latter was estimated based on fitting a Weibull survival function to Kaplan–Meier curves) and HR for PFS/TTP 0.86 vs HR for OS 1.57). The observed negative effect of erlotinib on OS in this study may have been due to the relatively high utilisation of non-study treatment post progression in the placebo group (Bukowski *et al*, 2007). The associations between treatment effects on PFS/TTP and treatment effects on OS were less strong when we used the results for OS from trials of sunitinib, everolimus, and pazopanib that were not adjusted for crossover from placebo to active therapy. The weighted Pearson correlation coefficient for the natural log of the relative risk of the ORR (i.e., $\ln RR_{ORR}$) vs $-\ln HR_{OS}$ was 0.78 ($P < 0.0001$). In linear regression, the coefficient on $\ln RR_{ORR}$ vs $-\ln HR_{OS}$ was 0.30 (95% CI: 0.20, 0.39, adjusted $R^2 = 0.59$) (Figure 4). In the analysis of $-\ln HR_{PFS/TTP}$ vs $-\ln HR_{OS}$ in which comparisons were weighted by the inverse variance of $-\ln HR_{OS}$ (35 comparisons), the weighted Pearson correlation coefficient for $\ln HR_{PFS/TTP}$ vs $-\ln HR_{OS}$ was 0.76 ($P < 0.0001$). The coefficient on $-\ln HR_{PFS/TTP}$ was 0.53 (95% CI: 0.37, 0.68, adjusted $R^2 = 0.56$). These results are qualitatively similar to those in which the results are weighted by the numbers of subjects.

DISCUSSION

Advances in understanding the biology and genetics of renal cell carcinoma have led to novel approaches for treatment of mRCC that target the VEGF receptor. With the growing therapeutic arsenal against mRCC, it is now feasible for patients to receive multiple lines of potentially beneficial treatment. Indeed, a recent trial reported on a study population that had received three to five prior lines of therapy (Motzer *et al*, 2010). With the increasing number of effective treatments available (Soulieres, 2009), the effect of first-line therapies on OS are more likely to be confounded by the effects of subsequent therapies. The question of whether PFS/TTP rather than OS should be employed as a primary outcome measure in pivotal studies of new treatments for mRCC is therefore important. This situation is similar to that with metastatic colorectal cancer, in which there was rapid development of novel treatments, necessitating the consideration of using PFS as a surrogate for OS in pivotal studies (Buyse *et al*, 2007). Although several novel treatments for mRCC have been approved for use in the United States with TTP or PFS as the primary end point in pivotal studies, and results of population-based historical cohort

studies of sunitinib and sorafenib have demonstrated that the introduction of these treatment has resulted in increased survival (Heng *et al*, 2009a; Warren *et al*, 2009), a rigorous examination of the association between PFS/TTP end points and OS has yet to be undertaken.

The analysis presented here suggests that treatment effects on measures of PFS/TTP are strongly associated with treatment effects on OS in patients with mRCC. However, the proportion of variability in treatment effects on OS that was explained by treatment effects on PFS/TTP was modest. In particular, the adjusted R^2 was 0.63 for the association between $-\ln HR_{PFS/TTP}$ and $-\ln HR_{OS}$. This value is within the range reported in other prior analyses of the relationship between treatment effects on PFS/TTP and OS (Sherrill *et al*, 2008). A high R^2 is not a necessary criterion for surrogacy, however, as some of the unexplained variation may reflect the sampling error in each trial due to small sample size. Even for a perfect surrogate end point, therefore, R^2 will be less than one in a set of trials with small samples (Tang *et al*, 2007). The trials examined in this evaluation were relatively small (median of 96 patients per arm). Moreover, there is no standard value above which an R^2 (or correlation coefficient) can be claimed to be sufficient. The adjusted R^2 for the association between differences in median PFS/TTP and differences in median OS was only 0.28. While the difference in median survival times may be a more appropriate measure of treatment effect than HRs if the proportional hazards assumption is violated, median survival times represent only a single point on the survival distribution and are potentially imprecise. It is not surprising therefore that amount of unexplained variation is greater when treatment effects are measured in terms of differences in median survival. Despite the relatively low R^2 from this regression, it is useful to note that the results from the regression analysis presented here suggest that, on average, there is a slightly better than 1-month gain in median OS associated with a 1 month gain in median PFS/TTP. This is consistent with the hypothesis that treatment effects on post-progression survival are uncorrelated with treatment effects on PFS/TTP (Bowater *et al*, 2008).

Not surprisingly, the association between treatment effects was stronger in studies that did not allow crossover to active treatment. Additionally, the association between treatment effects on PFS/TTP and OS were less in trials conducted after 2005, when targeted therapies for treatment of mRCC were more likely to be available as potential off-study second-line treatments. Estimates of the association between treatment effects on PFS/TTP and OS based on the entire sample of trials may therefore be conservative. An increase in response rate was also correlated with OS, although the association was not as strong as that with treatment effects on PFS/TTP measured in terms of $-\ln(HR)$.

Limitations of this study should be noted. First, this study was based on published results of controlled trials which may be subject to publication bias. To the extent that only studies showing positive effects on both PFS and OS were published, then our estimates may overstate the true association between PFS and OS. However, a funnel plot analysis of the $-\ln HR_{OS}$ provided no strong evidence of publication bias (the plot was symmetric around the mean effect size and Egger's test was not significant).

Ideally, the assessment of association of PFS/TTP and OS should be demonstrated over different stages of the disease (as the causal pathways of the disease process might differ depending on the stage) and across classes of drug (as drugs with different modes of action may have different pathways of intervention) (Fleming and DeMets, 1996). It is possible that the association reported here could only apply to specific recognised prognostic groups, but analyses by prognostic groups were unfeasible based on data reported in study publications (Molina and Motzer, 2008; Heng *et al*, 2009b). The majority of studies included in this analysis involved comparisons of two or more cytokine therapies. The association between treatment effects on PFS/TTP and those on OS

Table 2 Sensitivity and subgroup analyses

Subgroup/Sensitivity analysis	N	Δ Median OS (months) vs Δ Median PFS/TTP (months)					– ln HR for OS vs – ln HR for PFT/TTP					
		Weighted Pearson correlation	Weighted linear regression			N	Weighted Pearson correlation	Weighted linear regression				
			Coefficient on Δ in median PFS/TTP (months)					Coefficient on – ln HR for PFS/TTP				
			Estimate	95% CI	R ²			Estimate	95% CI	R ²		
All	41	0.54	1.17	0.59	1.76	0.28	36	0.80	0.64	0.47	0.81	0.63
Prior treatment												
None	20	0.57	1.22	0.35	2.08	0.29	17	0.84	0.61	0.39	0.82	0.69
Any	21	0.49	1.04	0.14	1.94	0.20	19	0.78	0.62	0.37	0.88	0.58
Targeted therapy												
No	27	0.65	1.42	0.74	2.10	0.40	24	0.80	0.60	0.40	0.80	0.62
Yes	14	0.38	0.85	–0.44	2.13	0.08	12	0.79	0.70	0.32	1.09	0.59
Measure of disease progression												
PFS	35	0.55	1.21	0.56	1.86	0.28	31	0.81	0.68	0.49	0.86	0.65
TTP	6	–0.10	–0.21	–2.98	2.56	–0.24	5	0.64	0.17	–0.20	0.53	0.21
Crossover allowed												
No	33	0.50	1.29	0.47	2.11	0.23	30	0.70	0.69	0.42	0.97	0.47
Yes	8	0.28	0.82	–1.95	3.59	–0.07	6	0.61	0.63	–0.49	1.76	0.22
Year of publication												
≤ 2005	22	0.80	1.78	1.15	2.41	0.62	21	0.69	0.55	0.27	0.83	0.45
> 2005	19	0.59	1.22	0.37	2.08	0.31	15	0.84	0.68	0.41	0.95	0.68
Number of subjects												
< 200	20	0.33	2.15	–0.95	5.26	0.06	17	0.70	0.51	0.23	0.80	0.46
≥ 200	21	0.66	1.13	0.51	1.74	0.40	19	0.82	0.67	0.43	0.90	0.65
HR estimated from Kaplan–Meier curves												
No							12	0.82	0.61	0.31	0.91	0.64
Yes							24	0.72	0.69	0.40	0.98	0.50
Drug class												
Cytokines	27	0.67	1.75	0.94	2.55	0.42	22	0.76	0.66	0.40	0.93	0.56
VEGF inhibitors	9	0.50	1.43	–0.80	3.65	0.14	9	0.66	0.65	0.00	1.30	0.36
MTOR inhibitors	5	0.87	1.65	–0.06	3.36	0.68	5	0.93	0.70	0.21	1.19	0.83
Exclude Bukowski <i>et al</i> (2007)	40	0.60	1.16	0.65	1.67	0.34	35	0.81	0.63	0.47	0.79	0.64
No adjustment for crossover in sunitinib, everolimus, and pazopanib trials	41	0.44	0.82	0.27	1.38	0.17	36	0.62	0.37	0.20	0.53	0.36
No intercept	41	0.54	1.20	0.76	1.65		36	0.80	0.58	0.46	0.69	
All potential comparisons	48	0.55	1.25	0.68	1.82		42	0.79	0.64	0.48	0.80	
All comparisons, no intercept	48	0.55	1.33	0.90	1.76		42	0.79	0.56	0.46	0.67	
All comparisons, no intercept, exclude Bukowski <i>et al</i> (2007)	47	0.59	1.24	0.73	1.74		41	0.80	0.63	0.47	0.78	

Abbreviations: CI = confidence interval; HR = hazard ratio; MTOR = mammalian target of rapamycin; N = number of comparisons; OS = overall survival; PFS = progression-free survival; R² = adjusted R-squared; TTP = time to progression; VEGF = vascular endothelial growth factor. Adjusted R² for regressions without intercept may not be comparable to those with intercept and are not reported.

were significant in trials evaluating targeted and non-targeted therapies. The association between treatment effects on PFS/TTP and OS was not significant for comparisons involving VEGF inhibitors, although there was a trend towards an association ($P=0.0510$). The number of such comparisons was small, however, and these comparisons may have been more likely to have been confounded by crossover and receipt of other non-study therapies post progression. It is reasonable to assume that results presented here can be generalised to evaluations of agents such as axitinib, that have similar mechanisms of action to the therapies included in this analysis (Rugo *et al*, 2005; Rini *et al*, 2007; Rixe *et al*, 2007).

For studies that allowed for crossover from control to active therapy, we used the reported measure of treatment effect that was considered to be least likely to be subject to confounding by such crossover. While it would be desirable to use a common measure of treatment effect for all studies, it is well established that crossover from control to active treatment may attenuate observed treatment effects on OS relative to what would have been observed in the absence of such crossover (Finkelstein and Schoenfeld, 2011; Saad and Buyse, 2012). To include results of studies with extensive crossover without controlling for crossover would add no useful information to the analyses. The RPSFT and IPCW methods used in the analyses of everolimus (Korhonen and Malangone, 2010;

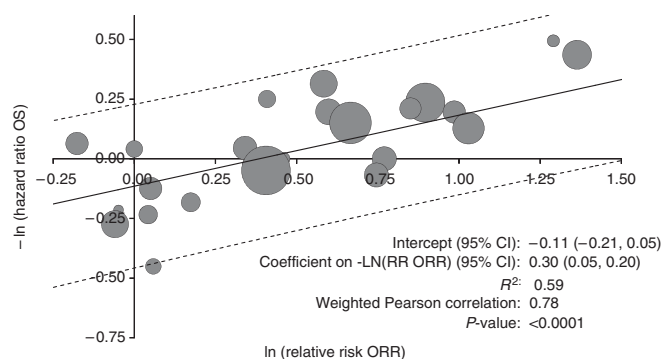


Figure 4 Association between the log of relative risk of overall response and the negative log of the hazard ratio of OS. R^2 = adjusted R -squared. Area of bubbles is proportional to the number of patients. Solid line is predicted value. Dashed lines are 95% prediction intervals.

Korhonen *et al*, 2011; Wiederkehr *et al*, 2011) and pazopanib (Sternberg *et al*, 2010b) are useful methods for analysing OS in the context of selective crossover (Finkelstein and Schoenfeld, 2011; Morden *et al*, 2011; Rimawi and Hilsenbeck, 2012).

In unblinded trials, there may be a motivation for clinicians to call a patient's disease progression earlier if the patient is in the control arm than if the same patient had been in the experimental arm (Dodd *et al*, 2008). To the extent that this inflates the treatment effects on PFS, the association between treatment effects on PFS/TTP and treatment effect on OS might be attenuated (because OS is not impacted by this bias). The use of blinded independent central review (BICR) may reduce any such bias. However, retrospective BICR may necessitate informative censoring on local assessment of progression, which may bias the comparison in favour of control patients (Dodd *et al*, 2008). This also would attenuate the observed association between treatment effect on PFS/TTP and treatment effect on OS. Treatment assignment was blinded in only six of the studies included in the analyses. Independent review of progression was employed in six studies. As studies that used blinded treatment assignment and/or review of progression tended to be those evaluating novel targeted agents, assessment of the independent effects of blinding of treatment assignment and/or BICR on the association between treatment effects on PFS and treatment effects on OS was infeasible.

REFERENCES

- Aass N, De Mulder PH, Mickisch GH, Mulders P, van Oosterom AT, van PH, Fossa SD, de PL, Sylvester RJ (2005) Randomized phase II/III trial of interferon alfa-2a with and without 13-cis-retinoic acid in patients with progressive metastatic renal cell carcinoma: the European Organisation for Research and Treatment of Cancer Genito-Urinary Tract Cancer Group (EORTC 30951). *J Clin Oncol* 23(18): 4172–4178
- Amato RJ, Jac J, Hernandez-McClain J (2008) Interferon-alpha in combination with either imatinib (Gleevec) or gefitinib (Iressa) in metastatic renal cell carcinoma: a phase II trial. *Anticancer Drugs* 19(5): 527–533
- Atkins MB, Hidalgo M, Stadler WM, Logan TF, Dutcher JP, Hudes GR, Park Y, Liou SH, Marshall B, Boni JP, Dukart G, Sherman ML (2004) Randomized phase II study of multiple dose levels of CCI-779, a novel mammalian target of rapamycin kinase inhibitor, in patients with advanced refractory renal cell carcinoma. *J Clin Oncol* 22(5): 909–918
- Atzpodien J, Hoffmann R, Franzke M, Stief C, Wandert T, Reitz M (2002) Thirteen-year, long-term efficacy of interferon 2alpha and interleukin 2-based home therapy in patients with advanced renal cell carcinoma. *Cancer* 95(5): 1045–1050
- Atzpodien J, Kirchner H, Illiger HJ, Metzner B, Ukena D, Schott H, Funke PJ, Gramatzki M, Jurgenson S, Wandert T, Patzelt M (2001) IL-2 in combination with IFN- alpha and 5-FU versus tamoxifen in metastatic

renal cell carcinoma: long-term results of a controlled randomized clinical trial. *Br J Cancer* 85(8): 1130–1136

Atzpodien J, Kirchner H, Jonas U, Bergmann L, Schott H, Heynemann H, Fornara P, Loening SA, Roigas J, Muller SC, Bodenstein H, Pomer S, Metzner B, Rebmann U, Oberneder R, Siebels M, Wandert T, Puchberger T, Reitz M (2004) Interleukin-2- and interferon alfa-2a-based immunotherapy in advanced renal cell carcinoma: a Prospectively Randomized Trial of the German Cooperative Renal Carcinoma Chemoimmunotherapy Group (DGCIN). *J Clin Oncol* 22(7): 1188–1194

Atzpodien J, Kirchner H, Rebmann U, Soder M, Gertenbach U, Siebels M, Roigas J, Raschke R, Salm S, Schwindl B, Muller SC, Hauser S, Leiber C, Huland E, Heinzer H, Siemer S, Metzner B, Heynemann H, Fornara P, Reitz M (2006) Interleukin-2/interferon-alpha2a/13-retinoic acid-based chemoimmunotherapy in advanced renal cell carcinoma: results of a prospectively randomised trial of the German Cooperative Renal Carcinoma Chemoimmunotherapy Group (DGCIN). *Br J Cancer* 95(4): 463–469

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(1): 48–53

Broglio KR, Berry DA (2009) Detecting an overall survival benefit that is derived from progression-free survival. *J Natl Cancer Inst* 101(23): 1642–1649

Information from the trial reports on the frequency of assessments, the criteria used to assess response and/or progression, or the duration of treatment was not extracted. It therefore was not feasible in this analysis to assess how these and other unmeasured factors might affect the association between treatment effects on PFS and treatment effects on OS. Differences in these factors might help explain some variability in observed associations between treatment effects on PFS/TTP and on OS.

As the searches upon which this study was based were conducted in 2010, results of randomized controlled trials of systemic therapies for mRCC may have been published since the original literature search for this study was conducted. One such trial is the Renal EFFECT trial, a randomized controlled trial of intermittent vs continuous sunitinib (Motzer *et al*, 2012). It may be worthwhile in future research to update these analyses using results of this and other recently published studies, and to explore in multivariate analysis the independent effects of study design and other factors on the associations between treatment effects on PFS/TTP and treatment effects on OS.

In conclusion, results presented in this study suggest that treatment effects on disease progression end points are strongly associated with treatment effects on OS. Further research is required to establish whether disease progression end points may be used as surrogate end points for OS in clinical trials of novel treatments for mRCC.

ACKNOWLEDGEMENTS

This work was supported by Novartis Pharmaceuticals Corp., Florham Park, NJ, USA. We thank Andrea Kay, Jean Francois Baladi, and Oleg Sofrygin for their assistance in the conduct of this study.

Conflict of interest

TED has received research support and consulting fees from Novartis, GlaxoSmithKline, Amgen, and other companies, which manufacture or market treatments for patients with metastatic renal cell carcinoma. TH is an employee of Novartis and owns stock or stock options in Novartis. DS has participated on Advisory Boards and received Honoraria from Novartis. The remaining authors declare no conflict of interest.

- Bukowski RM, Kabbinavar FF, Figlin RA, Flaherty K, Srinivas S, Vaishampayan U, Drabkin HA, Dutcher J, Ryba S, Xia Q, Scappaticci FA, McDermott D (2007) Randomized phase II study of erlotinib combined with bevacizumab compared with bevacizumab alone in metastatic renal cell cancer. *J Clin Oncol* 25(29): 4536–4541
- Burzykowski T, Buyse M, Piccart-Gebhart MJ, Sledge G, Carmichael J, Luck HJ, Mackey JR, Nabholz JM, Paridaens R, Biganzoli L, Jassem J, Bontenbal M, Bonnetterre J, Chan S, Basaran GA, Therasse P (2008) 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(12): 1987–1992
- Buyse M, Burzykowski T, Carroll K, Michiels S, Sargent DJ, Miller LL, Elfring GL, Pignon JP, Piedbois P (2007) Progression-free survival is a surrogate for survival in advanced colorectal cancer. *J Clin Oncol* 25(33): 5218–5224
- Carroll KJ (2003) On the use and utility of the Weibull model in the analysis of survival data. *Control Clin Trials* 24(6): 682–701
- Coppin C (2008) Immunotherapy for renal cell cancer in the era of targeted therapy. *Expert Rev Anticancer Ther* 8(6): 907–919
- Coppin C, Le L, Porzolt F, Wilt T (2008) Targeted therapy for advanced renal cell carcinoma. *Cochrane Database Syst Rev* 2: CD006017
- Coppin C, Porzolt F, Autenrieth M, Kumpf J, Coldman A, Wilt T (2005) Immunotherapy for advanced renal cell cancer. *Cochrane Database Syst Rev* 1: CD001425
- Dodd LE, Korn EL, Freidlin B, Jaffe CC, Rubinstein LV, Dancy J, Mooney MM (2008) Blinded independent central review of progression-free survival in phase III clinical trials: important design element or unnecessary expense? *J Clin Oncol* 26(22): 3791–3796
- Donskov F, Middleton M, Fode K, Meldgaard P, Mansoor W, Lawrance J, Thatcher N, Nellemann H, von der Maase H (2005) Two randomised phase II trials of subcutaneous interleukin-2 and histamine dihydrochloride in patients with metastatic renal cell carcinoma. *Br J Cancer* 93(7): 757–762
- Dutcher JP, Fine JP, Krigel RL, Murphy BA, Schaefer PL, Ernstoff MS, Loehrer PJ (2003) Stratification by risk factors predicts survival on the active treatment arm in a randomized phase II study of interferon-gamma plus/minus interferon-alpha in advanced renal cell carcinoma (E6890). *Med Oncol* 20(3): 271–281
- Egger M, Davey Smith G, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315(7109): 629–634
- Escudier B, Eisen T, Stadler WM, Szczylik C, Oudard S, Siebels M, Negrier S, Chevreau C, Solska E, Desai AA, Rolland F, Demkow T, Hutson TE, Gore M, Freeman S, Schwartz B, Shan M, Simantov R, Bukowski RM, for the TARGET Study Group (2007a) Sorafenib in advanced clear-cell renal-cell carcinoma. *N Engl J Med* 356(2): 125–134
- Escudier B, Pluzanska A, Koralewski P, Ravaud A, Bracarda S, Szczylik C, Chevreau C, Filipek M, Melichar B, Bajetta E, Gorbunova V, Bay JO, Bodrogi I, Jagiello-Gruszfeld A, Moore N, for the AVOREN Trial Investigators (2007b) Bevacizumab plus interferon alfa-2a for treatment of metastatic renal cell carcinoma: a randomised, double-blind phase III trial. *Lancet* 370(9605): 2103–2111
- Figlin RA, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, Negrier S, Huang X, Kim ST, Chen I, Motzer RJ (2008) Overall survival with sunitinib versus interferon (IFN)-alpha as first-line treatment of metastatic renal cell carcinoma (mRCC). *J Clin Oncol* 26: 15
- Finkelstein DM, Schoenfeld DA (2011) Correcting for discretionary treatment crossover in an analysis of survival in the Breast International Group BIG 1-98 trial by using the inverse probability of censoring weighted method. *J Clin Oncol* 29(9): 1093–1095
- Fleming TR, DeMets DL (1996) Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 125(7): 605–613
- Gore ME, Griffin CL, Hancock B, Patel PM, Pyle L, Aitchison M, James N, Oliver RT, Mardiak J, Hussain T, Sylvester R, Parmar MK, Royston P, Mulders PF (2010) Interferon alfa-2a versus combination therapy with interferon alfa-2a, interleukin-2, and fluorouracil in patients with untreated metastatic renal cell carcinoma (MRC RE04/EORTC GU 30012): an open-label randomised trial. *Lancet* 375(9715): 641–648
- Heng DY, Chi KN, Murray N, Jin T, Garcia JA, Bukowski RM, Rini BI, Kollmannsberger C (2009a) A population-based study evaluating the impact of sunitinib on overall survival in the treatment of patients with metastatic renal cell cancer. *Cancer* 115(4): 776–783
- Heng DY, Xie W, Regan MM, Warren MA, Golshayan AR, Sahi C, Eigel BJ, Ruether JD, Cheng T, North S, Venner P, Knox JJ, Chi KN, Kollmannsberger C, McDermott DF, Oh WK, Atkins MB, Bukowski RM, Rini BI, Choueiri TK (2009b) Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: results from a large, multi-center study. *J Clin Oncol* 27(34): 5794–5799
- Hudes G, Carducci M, Tomczak P, Dutcher J, Figlin R, Kapoor A, Staroslawska E, Sosman J, McDermott D, Bodrogi I, Kovacevic Z, Lesovoy V, Schmidt-Wolf IG, Barbarash O, Gokmen E, O'Toole T, Lustgarten S, Moore L, Motzer RJ (2007) Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. *N Engl J Med* 356(22): 2271–2281
- Johnson KR, Ringland C, Stokes BJ, Anthony DM, Freemantle N, Irs A, Hill SR, Ward RL (2006) 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 Oncol* 7(9): 741–746
- Knox JJ (2008) Progression-free survival as endpoint in metastatic RCC? *Lancet* 372(9637): 427–429
- Korhonen P, Malangone E (2010) Overall survival (OS) of mRCC patients corrected for crossover using inverse probability of censoring weights (IPCW) and rank preserving structural failure time (RPSFT) models: two analyses from the RECORD-1 trial. *J Clin Oncol* 28(Suppl): 15S (abstract 4595)
- Korhonen P, Malangone E, Sherman S, Casciano R, Motzer RJ, Baladi J, Haas T, Zuber E, Hollaender N, Lebwohl DE (2011) Overall survival (OS) of mRCC patients corrected for crossover using inverse probability of censoring weights (IPCW) and rank preserving structural failure time (RPSFT) models: two analyses from the RECORD-1 trial. In: *10th International Kidney Cancer Symposium*. 14–15 October 2011: Kidney Cancer Association. Chicago, IL, USA.
- Kruit WH, Goey SH, Lamers CH, Gratama JW, Visser B, Schmitz PI, Eggermont AM, Bolhuis RL, Stoter G (1997) High-dose regimen of interleukin-2 and interferon-alpha in combination with lymphokine-activated killer cells in patients with metastatic renal cell cancer. *J Immunother* 20(4): 312–320
- Louvet C, de GA, Tournigand C, Artru P, Maindrault-Goebel F, Krulik M (2001) Correlation between progression free survival and response rate in patients with metastatic colorectal carcinoma. *Cancer* 91(11): 2033–2038
- McDermott DF, Regan MM, Clark JI, Flaherty LE, Weiss GR, Logan TF, Kirkwood JM, Gordon MS, Sosman JA, Ernstoff MS, Tretter CPG, Urba WJ, Smith JW, Margolin KA, Mier JW, Gollob JA, Dutcher JP, Atkins MB (2005) Randomized phase III trial of high-dose interleukin-2 versus subcutaneous interleukin-2 and interferon in patients with metastatic renal cell carcinoma. *J Clin Oncol* 23(1): 133–141
- Medical Research Council Renal Center Collaborators (1999) Interferon-alpha and survival in metastatic renal carcinoma: early results of a randomised controlled trial. Medical Research Council Renal Cancer Collaborators. *Lancet* 353(9146): 14–17
- Molina AM, Motzer RJ (2008) Current algorithms and prognostic factors in the treatment of metastatic renal cell carcinoma. *Clin Genitourin Cancer* 6(Suppl 1): S7–S3
- Morden JP, Lambert PC, Latimer N, Abrams KR, Wailoo AJ (2011) Assessing methods for dealing with treatment switching in randomised controlled trials: a simulation study. *BMC Med Res Methodol* 11: 4
- Motzer RJ, Escudier B, Oudard S, Hutson TE, Porta C, Bracarda S, Grunwald V, Thompson JA, Figlin RA, Hollaender N, Kay A, Ravaud A (2010) Phase 3 trial of everolimus for metastatic renal cell carcinoma: final results and analysis of prognostic factors. *Cancer* 116(18): 4256–4265
- Motzer RJ, Escudier B, Oudard S, Hutson TE, Porta C, Bracarda S, Grunwald V, Thompson JA, Figlin RA, Hollaender N, Urbanowitz G, Berg WJ, Kay A, Lebwohl D, Ravaud A, for the RECORD-1 Study Group (2008) Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial. *Lancet* 372(9637): 449–456
- Motzer RJ, Hutson TE, Olsen MR, Hudes GR, Burke JM, Edenfield WJ, Wilding G, Agarwal N, Thompson JA, Cella D, Bello A, Korytowsky B, Yuan J, Valota O, Martell B, Hariharan S, Figlin RA (2012) Randomized Phase II trial of sunitinib on an intermittent versus continuous dosing schedule as first-line therapy for advanced renal cell carcinoma. *J Clin Oncol* 30(12): 1371–1377
- Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, Oudard S, Negrier S, Szczylik C, Pili R, Bjarnason GA, Garcia-del-Muro X, Sosman JA, Solska E, Wilding G, Thompson JA, Kim ST, Chen I, Huang X, Figlin RA (2009) Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma. *J Clin Oncol* 27(22): 3584–3590

- Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, Rixe O, Oudard S, Negrier S, Szczylik C, Kim ST, Chen I, Bycott PW, Baum CM, Figlin RA (2007) Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N Engl J Med* 356(2): 115–124
- Motzer RJ, Murphy BA, Bacik J, Schwartz LH, Nanus DM, Mariani T, Loehrer P, Wilding G, Fairclough DL, Cella D, Mazumdar M (2000) Phase III trial of interferon alfa-2a with or without 13-cis-retinoic acid for patients with advanced renal cell carcinoma. *J Clin Oncol* 18(16): 2972–2980
- Negrier S, Caty A, Lesimple T, Douillard JY, Escudier B, Rossi JF, Viens P, Gomez F (2000) Treatment of patients with metastatic renal carcinoma with a combination of subcutaneous interleukin-2 and interferon alfa with or without fluorouracil. Groupe Francais d'Immunotherapie, Federation Nationale des Centres de Lutte Contre le Cancer. *J Clin Oncol* 18(24): 4009–4015
- Negrier S, Escudier B, Lasset C, Douillard JY, Savary J, Chevreau C, Ravaud A, Mercatello A, Peny J, Mousseau M, Philip T, Tursz T, for the Groupe Francais d'Immunotherapie (1998) Recombinant human interleukin-2, recombinant human interferon alfa-2a, or both in metastatic renal-cell carcinoma. *N Engl J Med* 338(18): 1272–1278
- Negrier S, Perol D, Ravaud A, Bay JO, Oudard S, Chabaud S, Fargeot P, Delva R, Deplanque G, Gravis G, Escudier B (2008) Randomized study of intravenous versus subcutaneous interleukin-2, and IFNalpha in patients with good prognosis metastatic renal cancer. *Clin Cancer Res* 14(18): 5907–5912
- Negrier S, Perol D, Ravaud A, Chevreau C, Bay JO, Delva R, Sevin E, Caty A, Escudier B, for the French Immunotherapy Intergroup (2007) Medroxyprogesterone, interferon alfa-2a, interleukin 2, or combination of both cytokines in patients with metastatic renal carcinoma of intermediate prognosis: results of a randomized controlled trial. *Cancer* 110(11): 2468–2477
- Pyrhonen S, Salminen E, Ruutu M, Lehtonen T, Nurmi M, Tammela T, Juusela H, Rintala E, Hietanen P, Kellokumpu-Lehtinen PL (1999) Prospective randomized trial of interferon alfa-2a plus vinblastine versus vinblastine alone in patients with advanced renal cell cancer. *J Clin Oncol* 17(9): 2859–2867
- Rimawi M, Hilsenbeck SG (2012) Making sense of clinical trial data: is inverse probability of censoring weighted analysis the answer to crossover bias? *J Clin Oncol* 30(4): 453–458
- Rini B, Wilding GT, Hudes G (2007) Axitinib (AG-013736; AG) in patients (pts) with metastatic renal cell cancer (RCC) refractory to sorafenib. *J Clin Oncol* 25(18S, 20 June Supplement): 5032
- Rini BI, Halabi S, Rosenberg JE, Stadler WM, Vaena DA, Archer L, Atkins JN, Picus J, Czaykowski P, Dutcher J, Small EJ (2010) Phase III trial of bevacizumab plus interferon alfa versus interferon alfa monotherapy in patients with metastatic renal cell carcinoma: final results of CALGB 90206. *J Clin Oncol* 28(13): 2137–2143
- Rixe O, Bukowski RM, Michaelson MD, Wilding G, Hudes GR, Bolte O, Motzer RJ, Bycott P, Liau KF, Freddo J, Trask PC, Kim S, Rini BI (2007) Axitinib treatment in patients with cytokine-refractory metastatic renal-cell cancer: a phase II study. *Lancet Oncol* 8(11): 975–984
- Rugo HS, Herbst RS, Liu G, Park JW, Kies MS, Steinfeldt HM, Pithavala YK, Reich SD, Freddo JL, Wilding G (2005) Phase I trial of the oral antiangiogenesis agent AG-013736 in patients with advanced solid tumors: pharmacokinetic and clinical results. *J Clin Oncol* 23(24): 5474–5483
- Saad ED, Buyse M (2012) Overall survival: patient outcome, therapeutic objective, clinical trial end point, or public health measure? *J Clin Oncol* 30(15): 1750–1754
- Sargent DJ, Hayes DF (2008) Assessing the measure of a new drug: is survival the only thing that matters? *J Clin Oncol* 26(12): 1922–1923
- Sherrill B, Amonkar M, Wu Y, Hirst C, Stein S, Walker M, Cuzick J (2008) Relationship between effects on time-to-disease progression and overall survival in studies of metastatic breast cancer. *Br J Cancer* 99(10): 1572–1578
- Soulieres D (2009) Review of guidelines on the treatment of metastatic renal cell carcinoma. *Curr Oncol* 16(Suppl 1): S67–S70
- Sternberg CN, Davis ID, Mardiak J, Szczylik C, Lee E, Wagstaff J, Barrios CH, Salman P, Gladkov OA, Kavina A, Zarba JJ, Chen M, McCann L, Pandite L, Roychowdhury DF, Hawkins RE (2010a) Pazopanib in locally advanced or metastatic renal cell carcinoma: results of a randomized phase III trial. *J Clin Oncol* 28(6): 1061–1068
- Sternberg CN, Hawkins RE, Szczylik C, Davis ID, Wagstaff J, McCann L, Chen M, Rubin SD (2010b) Randomized, double-blind phase III study of pazopanib in patients with advanced/metastatic renal cell carcinoma (MRCC): final overall survival (OS) results. *Annals Oncol* 21(Suppl 8): viii1–viii12
- Sternberg CN, Szczylik C, Lee E, Salman PV, Mardiak J, Davis ID, Pandite L, Chen M, McCann L, Hawkins R (2009) A randomized, double-blind phase III study of pazopanib in treatment-naïve and cytokine-pretreated patients with advanced renal cell carcinoma (RCC). Presentation available at http://www.asco.org/ascov2/Meetings/Abstracts?&vmview=abst_detail_view&confID=65&abstractID=30883 (Accessed 16 April, 2012). *J Clin Oncol* 27: (15s (suppl; abstract 50211))
- Tang PA, Bentzen SM, Chen EX, Siu LL (2007) 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* 25: 4562–4568
- Tannir NM, Cohen L, Wang X, Thall P, Mathew PF, Jonasch E, Siefker-Radtke A, Pagliaro LC, Ng CS, Logothetis C (2006) Improved tolerability and quality of life with maintained efficacy using twice-daily low-dose interferon-alpha-2b: results of a randomized phase II trial of low-dose versus intermediate-dose interferon-alpha-2b in patients with metastatic renal cell carcinoma. *Cancer* 107(9): 2254–2261
- Thompson Coon J, Hoyle M, Green C, Liu Z, Welch K, Moxham T, Stein K (2010) Bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma: a systematic review and economic evaluation. *Health Technol Assess* 14(2): 1–184, iii–iv
- Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR (2007) Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials* 8: 16
- Warren M, Venner PM, North S, Cheng T, Venner C, Ghosh S, Venner AA, Finch D (2009) A population-based study examining the effect of tyrosine kinase inhibitors on survival in metastatic renal cell carcinoma in Alberta and the role of nephrectomy prior to treatment. *Can Urol Assoc J* 3(4): 281–289
- Wiederkehr D, Howe CJ, Casciano R, Motzer R, Zheng J, Baladi JF (2011) Overall survival among metastatic renal cell carcinoma patients corrected for crossover using inverse probability of censoring weights: analyses from the RECORD phase 3 trial. Presented at the Joint 15th Congress of the European CanCer Organisation (ECCO) and 34th Congress of the European Society for Medical Oncology (ESMO), 20–24 September, 2009, Berlin, Germany
- Yang JC, Haworth L, Sherry RM, Hwu P, Schwartzentruber DJ, Topalian SL, Steinberg SM, Chen HX, Rosenberg SA (2003) A randomized trial of bevacizumab, an anti-vascular endothelial growth factor antibody, for metastatic renal cancer. *N Engl J Med* 349(5): 427–434

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License.