Correlations of survival with progression-free survival, response rate, and disease control rate in advanced biliary tract cancer: a meta-analysis of randomised trials of first-line chemotherapy

FULL PAPER

BJC

British Journal of Cancer (2016) 114, 881–888 | doi: 10.1038/bjc.2016.83

Keywords: biliary tract cancer; first-line chemotherapy; surrogate end point; overall survival; meta-analysis

Correlations of survival with progression-free survival, response rate, and disease control rate in advanced biliary tract cancer: a meta-analysis of randomised trials of first-line chemotherapy

Toshikazu Moriwaki*[,1], Yoshiyuki Yamamoto[1], Masahiko Gosho[2], Mariko Kobayashi[1], Akinori Sugaya[1], Takeshi Yamada[1], Shinji Endo[1] and Ichinosuke Hyodo[1]

[1]*Faculty of Medicine, Division of Gastroenterology, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8575, Japan and* [2]*Faculty of Medicine, Department of Clinical Trial and Clinical Epidemiology, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8575, Japan*

Background: The need to promote novel drug development for advanced biliary tract cancer (ABTC) has emphasised the importance of determining whether various efficacy end points can act as surrogates for overall survival (OS).

Methods: We conducted a literature search of randomised trials of first-line chemotherapy for ABTC and investigated correlations between efficacy end points and OS using weighted linear regression analysis. The ratios of the median OS, median progression-free survival (PFS), response rate, and disease control rate in each trial were used to summarise treatment effects. The surrogate threshold effect (STE), which was the minimum treatment effect on PFS required to predict a non-zero treatment effect on OS, was calculated.

Results: Seventeen randomised trials with 36 treatment arms were identified, and a sample size of 2148 patients with 19 paired arms was analysed. The strongest correlation between all evaluated efficacy end points was observed between median OS and median PFS ratios ($r^2 = 0.66$). In trials with gemcitabine-containing therapies and targeted agents, the r^2-values were 0.78. The STE was estimated at 0.83 for all trials and 0.81 for trials with gemcitabine-containing therapies, and was not calculated for trials with targeted agents.

Conclusions: The median PFS ratio correlated well with the median OS ratio, and may be useful for planning a clinical trial for novel drug development.

Biliary tract cancer is a rare disease that includes intrahepatic and extrahepatic cholangiocarcinoma, gallbladder carcinoma, and ampullary carcinoma, and patients with unresectable or recurrent forms of these cancers are generally treated with palliative chemotherapy. The key anti-cancer drug for advanced biliary tract cancer (ABTC) is gemcitabine. Recently, gemcitabine + cisplatin (GP) has been recognised as a standard first-line treatment for ABTC following the results of a randomised phase III trial (ABC-02 trial) that compared GP with gemcitabine alone (Valle *et al*, 2010). Despite the favourable outcome of that trial, the prognosis

*Correspondence: Dr T Moriwaki; E-mail: tmoriwak@gmail.com

Received 9 November 2015; revised 4 February 2016; accepted 3 March 2016; published online 31 March 2016

of patients with ABTC remains poor, with a median survival time of ~12 months. Accordingly, new drugs or combination therapies should be studied with the intent to improve survival.

Although overall survival (OS) is considered a gold-standard hard end point of trials assessing oncologic drugs, a validated shorter-term surrogate end point for OS would reduce the sample size, cost, and trial duration needed to demonstrate the benefit of a novel drug. A previous pooled analysis of 104 clinical trials for ABTC prior to establishing GP as a standard chemotherapy regimen was conducted to identify superior treatment regimens (Eckel and Schmid, 2007). This study identified the time to progression (TTP) was significantly associated with OS as a secondary objective. However, the analysis included mostly small sample size and non-randomised trials. In the present study, we investigated various efficacy end points as potential surrogates for OS in studies of first-line chemotherapy for ABTC in a limited pool of randomised trials to allow the accurate estimation of treatment effects.

MATERIALS AND METHODS

Registration. This study is registered in the PROSPERO database (CRD42014014526) and was conducted according to the Preferred Reporting Items for Systemic Reviews and Meta-Analysis (PRISMA) statement.

Selection of eligible studies. Eligible studies were randomised phase II and III trials involving patients with ABTC who were treated with first-line chemotherapy. Studies involving the following were excluded: hepatic infusional chemotherapy; combination therapy with a local control therapy such as radiotherapy, surgery or photodynamic therapy, or non-English language reports.

Search strategy. Trials involving ABTC that were published up to February 2015 were identified through a systemic search of the PubMed database, using the keywords 'biliary tract neoplasms' OR 'bile duct neoplasms' OR 'gallbladder neoplasms' OR 'cholangiocarcinoma' (all fields) AND 'chemotherapy' (all fields) AND 'clinical trial' (ptyp). A manual search was also performed for abstracts presented at the annual meetings of the American Society of Clinical Oncology (ASCO), Gastrointestinal Cancers Symposium, European Society of Medical Oncology, and World Congress of Gastrointestinal Cancer up to February 2015.

Data extraction. Two authors (TM and YY) independently extracted information from the selected literature using predefined data abstract forms. The following details were extracted: published or presented year, number of enroled patients, primary end point, chemotherapy regimen, and tumour location (intrahepatic cholangiocarcinoma, extrahepatic cholangiocarcinoma, gallbladder carcinoma, and ampullary carcinoma). The various efficacy parameters assessed as potential surrogate end points for and evaluated for correlations with OS were progression-free survival (PFS; defined as the time to initial progression or death by any cause), TTP (defined as the time to initial progression or cancer-related death), response rate (RR; defined as the rate of complete and partial responses), and disease control rate (DCR; defined as the rate of complete and partial responses and stable disease). The following efficacy data were collected: median values of OS, PFS, and TTP; hazard ratios of OS, PFS, and TTP; and RR and DCR. The control arm in each trial was determined by the consensus of three investigators (TM, YY, and TY) for randomised phase II studies with selection designs wherein all arms were considered experimental treatments.

Statistical analysis. The coefficient of determination (r^2) was used to evaluate correlations between the treatment effects on surrogate efficacy end points and treatment effects on OS. Treatment effects on OS were analysed using a linear regression model weighted according to the study sample size of each trial. This model included the treatment effect on each surrogate end point as an exploratory variable. The precision levels of predictions based on this model were demonstrated by 95% confidence intervals (CIs) around the regression line. The treatment effects on PFS and OS were analysed via conversion to a logarithmic scale. The ratios of the median PFS, TTP, and OS between the control and experimental arms in each trial were used to summarise treatment effects because the hazard ratios (HRs) were not always reported. For RR and DCR, the ratios between the control arm and experimental arm were evaluated. TTP and failure-free survival were reclassified as PFS (Schinzari *et al*, 2009; Sasaki *et al*, 2013). A ratio of less than 1 denotes a favourable result for PFS, RR, DCR, and OS in the experimental arm.

Several additional analyses, not pre-specified, were performed. The correlations between treatment effects on surrogate efficacy end points and treatment effects on OS among trials with gemcitabine-containing therapies and with targeted agents were analysed. Sensitivity analyses were also performed to support the correlation observed between the median PFS ratios and median OS ratios in all trials. The surrogate threshold effect (STE) on PFS was calculated in all trials and subgroups. The STE is derived from a vertical line that transects the upper 95% predictive limit and a median OS ratio equal to 1; this represents the minimum PFS effect that could predict a positive OS effect (Burzykowski and Buyse, 2006). The predicted median PFS ratio associated with a 20% improvement in the median OS was calculated from the slope of each regression line because an improvement in the median OS of at least 20% has been generally agreed upon as a clinically meaningful outcome improvement in cancer clinical trials (Ellis *et al*, 2014). We used SAS software version 9.4 (SAS Institute Inc., Cary, NC, USA) for the statistical analyses.

RESULTS

Trials included in the analysis. Among the 426 retrieved papers/abstracts, 17 randomised trials (14 phase II studies and 3 phase III studies) with 36 treatment arms were identified (Figure 1 and Table 1) (Kornek *et al*, 2004; Ducreux *et al*, 2005; Rao *et al*, 2005; Schinzari *et al*, 2009; Okusaka *et al*, 2010; Sharma *et al*, 2010; Valle *et al*, 2010; Lee *et al*, 2012; Kang *et al*, 2012; Morizane *et al*, 2013; Sasaki *et al*, 2013; Malka *et al*, 2014; Moehler *et al*, 2014; Valle *et al*, 2014; Chen *et al*, 2015; Leone *et al*, 2015; Santoro *et al*, 2015). A total of 2148 patients with 19 paired arms were analysed. PFS was reported in 14 trials, and TTP was reported in 2 trials. Although RR was reported in all trials, a best supportive care arm in one trial reported no response (Sharma *et al*, 2010). DCR was reported in 15 trials. Two trials were terminated early. A phase III trial comparing 5-fluorouracil (5-FU) + epirubicin + leucovorin with epirubicin + cisplatin + 5-FU and a phase II trial comparing GP + placebo with GP + cediranib were closed because of poor accrual and new drug development cessation, respectively (Rao *et al*, 2005; Valle *et al*, 2014). The median values tended to favour the experimental arms over the standard arms with respect to the reported median PFS (5.8 months *vs* 4.9 months), median OS (10.1 months *vs* 9.5 months), RR (26.1% *vs* 15.5%), and DCR (71.7% *vs* 64.9%). A forest plot of the treatment effects on PFS and OS in trials reporting HR is shown in Figure 2. The median HR also tended to favour the experimental arms over the standard arms with respect to the reported PFS (0.80) and OS (0.82). In trials with targeted agents, however, the HRs for PFS and OS trended to favour the control arms. The median PFS and OS for gemcitabine-alone ranged widely from 3.7 to 5.0 months and 7.7 to 11.2

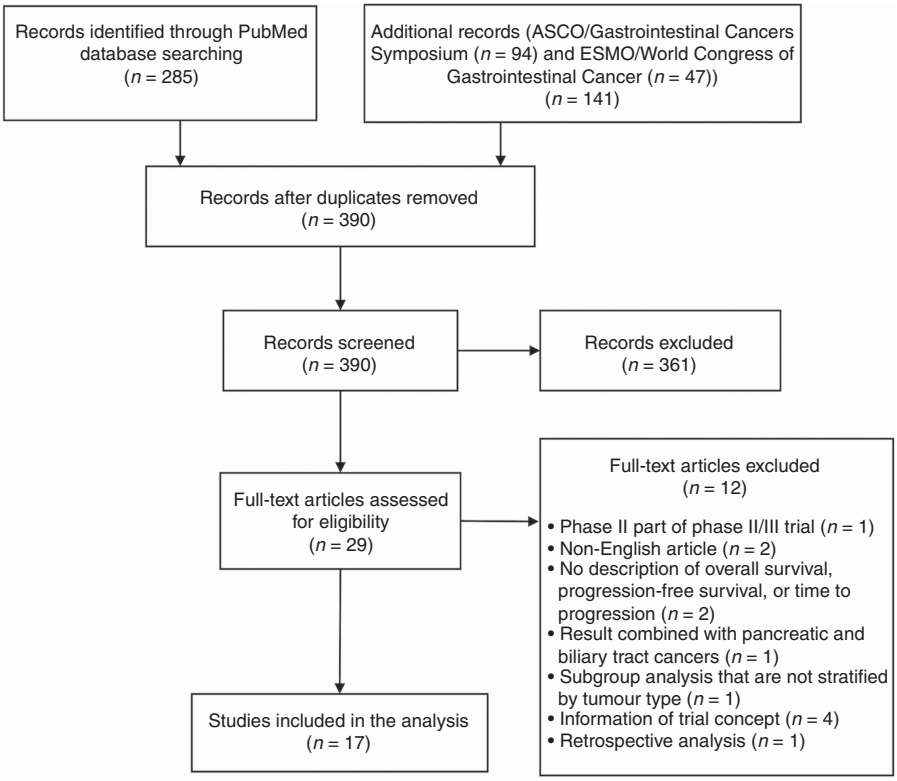

Figure 1. Study selection according to PRISMA (preferred reporting items for systemic reviews) diagram. Abbreviations: ASCO = American Society of Clinical Oncology; ESMO = European Society of Medical Oncology.

months, respectively. The median PFS and OS for gemcitabine + platinum combination therapies also exhibited wide distributions of 4.1–8.5 months and 9.5–12.4 months, respectively.

Correlation between treatment effects. Correlations between the surrogate end points and OS are summarised in Table 2. The median PFS ratio and OS ratio were moderately correlated ($r^2 = 0.66$; 95% CI, 0.32–0.85, $P < 0.001$; Figure 3A). The correlation between the RR ratio and median OS ratio in all 17 trials with 17 paired arms was weak ($r^2 = 0.29$; 95% CI, 0.01–0.65, $P = 0.021$; Figure 3B), as were the correlations between the DCR ratio and median OS ratio in 15 trials with 17 paired arms ($r^2 = 0.34$; 95% CI, 0.02–0.69, $P = 0.011$; Figure 3C).

Fourteen trials with gemcitabine-containing therapies were identified, and correlations were analysed between the median PFS ratio and OS ratio in 15 paired arms, RR ratio and median OS ratio in 14 paired arms, and DCR ratio and median OS ratio in 14 paired arms. The median PFS ratio associated strongly with the median OS ratio ($r^2 = 0.78$; 95% CI, 0.46–0.92, $P < 0.001$; Figure 4A). However, the correlations of the RR ratio and DCR ratio with the median OS ratio were weak (Table 2). Only six trials with targeted agents were identified, and correlations were analysed between the seven paired arms of those end points and the median OS ratios. The median PFS ratio associated strongly with the median OS ratio ($r^2 = 0.78$; 95% CI, 0.14–0.96, $P = 0.004$; Figure 4B), and this correlation was stronger than that observed between the RR ratio and DCR ratio and the median OS ratio (Table 2).

Sensitivity analysis. Several sensitivity analyses were performed to support the strong correlations observed between the median PFS ratio and median OS ratio in all trials (Table 2). When 10 trials with total sample sizes of <100 patients were excluded, an r^2-value of 0.60 (95% CI, 0.02–0.92, $P = 0.015$) was calculated between the median PFS ratio and median OS ratio in 7 trials with 8 paired arms. The r^2-value was 0.64 (95% CI, 0.27–0.86, $P < 0.001$) after excluding 2 trials that ended early. Among the 8 trials with 9 paired arms that reported HRs for both PFS and OS, the r^2-value was 0.63 (95% CI, 0.07–0.91, $P = 0.006$). After excluding 3 trials that reported TTP or failure-free survival, an r^2-value of 0.62 (95% CI, 0.23–0.85, $P < 0.001$) was calculated between the median PFS ratio and median OS ratio in 14 trials with 16 paired arms.

Predicted treatment effect on PFS. The STE was 0.83 for all trials (Figure 3A), 0.81 for trials with gemcitabine-containing therapies (Figure 4A), and was not calculated for trials with targeted agents. The predicted median PFS ratio associated with a 20% improvement in the median OS was 0.71 for all trials, 0.69 for trials with gemcitabine-containing therapies, and 0.40 for trials with targeted agents.

DISCUSSION

Our analysis demonstrated that an improvement in PFS was moderately associated with an improvement in OS in randomised trials of first-line chemotherapy for ABTC. A strong correlation between PFS and OS was confirmed in both subgroups of gemcitabine-containing trials and targeted agent-combined trials, which have been actively investigated in ABTC recently.

In the trials we analysed, the significant improvement in OS with the significant improvement in PFS was observed in two paired arms, which involved a comparison between GP and gemcitabine in the ABC-02 trial and a comparison between modified gemcitabine + oxaliplatin and best supportive care (Sharma *et al*, 2010; Valle *et al*, 2010). Particularly, the ABC-02 trial had the largest sample size. One of the reasons for our positive results was that most analysed trials tended to have results similar to those of the ABC-02 trial, although the slight improvement in

Table 1. Characteristics of trials included in the analysis

Reference (year)	Treatment arms	Phase	No. of patients	Primary end point	RR (%)	DCR (%)	Median PFS (months)	HR (95% CI)	Median OS (months)	HR (95% CI)
Kornek *et al* (2004)	C arm: MMC + CAPE	II	26	RR	30.7	65.4	5.3	NR	9.3	NR
	E arm: MMC + GEM		25		20	56	4.2		6.7	
Rao *et al* (2005)	C arm: FELV	III	27	OS	15	60	7.2[a]	NR	12	NR
	E arm: ECF		27		19.2	65.4	5.2[a]		9	
Ducreux *et al* (2005)	C arm: HDFU	II	29	RR	7.1	NR	3.3	NR	5	NR
	E arm: FLP		29		18.5		3.3		8	
Schinzari *et al* (2009)	C arm: FU/LV	II	23	NR	21.7	56.5	2.7[b]	NR	6.7	NR
	E arm: FOLFOX4		25		28	72	5.0[b]		12.6	
Sharma *et al* (2010)	C arm: BSC	II	27	OS	0	3.7	2.8	1	4.5	1
	E arm: FUFA		28		14.3	21.4	3.5	0.72 (0.39–1.34)	4.6	0.82 (0.45–1.51)
	E arm: mGEMOX		26		30.7	68.7	8.5	0.28 (0.14–0.56)	9.5	0.44 (0.22–0.86)
Okusaka *et al* (2010)	C arm: GEM	II	42	OS	11.5	50	3.7	0.66 (0.41–1.05)	7.7	0.69 (0.42–1.13)
	E arm: GP		41		19.5	68.3	5.8		11.2	
Valle *et al* (2010)	C arm: GEM	III	206	OS	15.5	71.8	5	0.63 (0.51–0.77)	8.1	0.64 (0.52–0.80)
	E arm: GP		204		26.1	81.4	8		11.7	
Kang *et al* (2012)	C arm: SP	II	47	PFS	23.8	85.7	5.4	0.85 (0.52–1.36)	9.9	0.72 (0.45–1.17)
	E arm: GP		49		19.6	71.7	5.7		10.1	
Lee *et al* (2012)	C arm: GEMOX	III	133	PFS	16	66	4.2	0.80 (0.61–1.03)	9.5	0.93 (0.69–1.25)
	E arm: GEMOX + Erlotinib		135		30	66	5.8		9.5	
Sasaki *et al* (2013)	C arm: GEM	II	32	RR	9.4	62.5	4.3[b]	NR	9.2	NR
	E arm: GEM + S-1		30		20	70	5.6[b]		8.9	
Morizane *et al* (2013)	C arm: S-1	II	50	OS	17.4	NR	4.2	0.44 (0.29–0.67)	9	0.86 (0.54–1.36)
	E arm: GEM + S-1		51		36.4		7.1		12.5	
Malka *et al* (2014)	C arm: GEMOX	II	74	PFS	23	64.9	5.5	1.08 (0.75–1.54)	12.4	NR
	E arm: GEMOX + Cetuximab		76		23.6	81.6	6.1		11	
Valle *et al* (2014)	C arm: GP + Placebo	II	62	PFS	18.5	64.8	7.4	0.93 (0.65–1.35)	11.9	0.86 (0.58–1.27)
	E arm: GP + Cediranib		62		44.1	78	8		14.1	
Moehler *et al* (2014)	C arm: GEM + Placebo	II	48	PFS	10	90	4.9	1.28 (0.81–2.02)	11.2	1.2 (0.75–1.93)
	E arm: GEM + Sorafenib		49		14.3	85.7	3		8.4	
Santoro *et al* (2015)	C arm: GEM + Placebo	II	52	PFS	13.5	38.5	4.9	1	10.1	NR
	E arm: Vandetanib		56		3.6	25	3.4	1.3 (0.86–1.96)	7.5	
	E arm: GEM + Vandetanib		57		19.3	29.8	3.7	1.3 (0.75–1.70)	9.3	
Chen *et al* (2015)	C arm: GEMOX	II	60	RR	15	36.7	4.1	0.70 (0.48–1.01)	9.8	NR
	E arm: GEMOX + Cetuximab		62		27.4	58.1	6.7		10.6	
Leone *et al* (2015)	C arm: GEMOX	II	44	PFS	18.2	63.6	5.5	NR	9.9	NR
	E arm: GEMOX + Panitumumab		45		24.4	73.3	7.7		9.5	

Abbreviations: BSC = best supportive care; C arm = control arm; CAPE = capecitabine; CI = confidence interval; DCR = disease control rate; E arm = experimental arm; ECF = epirubicin + cisplatin + 5-fluorouracil; 5-FU = 5-fluorouracil; FELV = 5-FU + epirubicin + leucovorin; FOLFOX4 = infusional 5-FU + leucovorin + oxaliplatin; FLP = 5-FU + leucovorin + cisplatin; FUFA = 5-FU + folic acid; FU/LV = 5-FU + leucovorin; GEM = gemcitabine; GEMOX; gemcitabine + oxaliplatin; GP = gemcitabine + cisplatin; HDFU = high-dose 5-FU; HR = hazard ratio; mGEMOX = modified gemcitabine + oxaliplatin; MMC = mitomycin C; NR = not reported; OS = overall survival; PFS = progression-free survival; RR = response rate; SP = S-1 + cisplatin.
[a]Failure-free survival.
[b]Time to progression.

OS was not statistically significant. Furthermore, it was suggested that the impact of the post-progression survival to OS was weak, compared with the association between PFS and OS. The validity of PFS as a surrogate end point for OS has been investigated in various types of cancers, but remains controversial in some (Burzykowski *et al*, 2008; Paoletti *et al*, 2013; Blumenthal *et al*, 2015; Johnson *et al*, 2015). The increased proportions of subsequent treatments and crossovers, as well as prolonged OS in recent cancer trials, might partly explain why the strong correlation between PFS and OS was not detected in first-line trials. The efficacy of subsequent treatment after the first progression has not been established in ABTC, and indeed, only a few active drugs are used as second-line chemotherapy for ABTC. The correlation between PFS and OS in second-line chemotherapy for ABTC was moderate and the evidence remains insufficient to recommend it in a systematic review (Lamarca *et al*, 2014), although 15–80% of

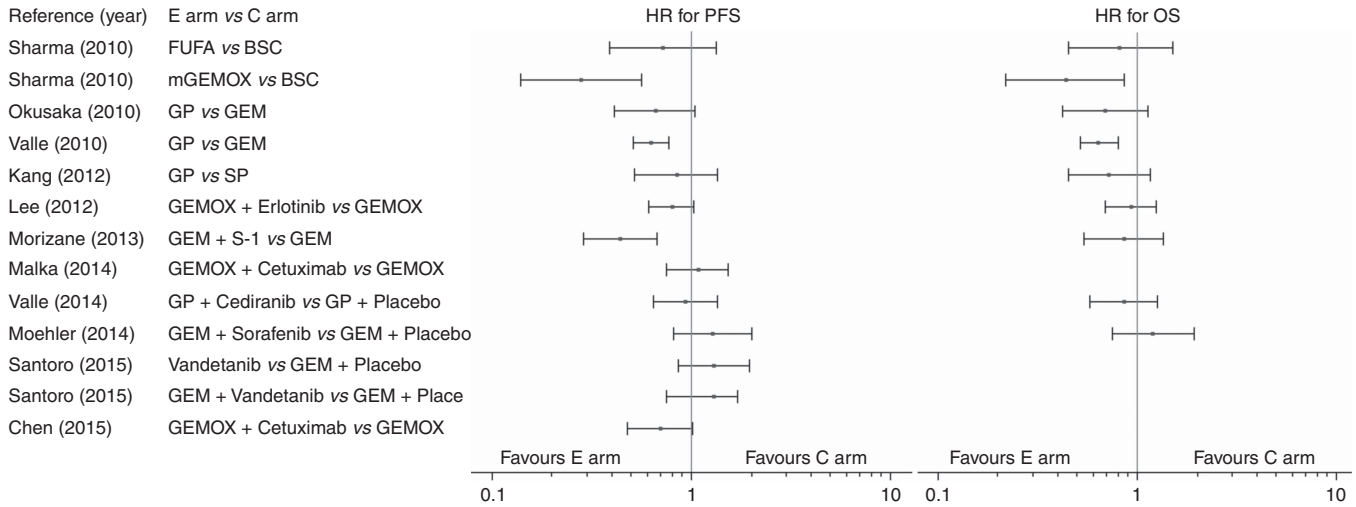

Figure 2. A forest plot of treatment effects on progression-free survival (PFS) and overall survival (OS) in trials reporting hazard ratios (HRs). Abbreviations: BSC = best supportive care; C arm = control arm; CI = confidence interval; E arm = experimental arm; FUFA = 5-FU + folic acid; GEM = gemcitabine; GEMOX = gemcitabine + oxaliplatin; GP = gemcitabine + cisplatin; mGEMOX = modified gemcitabine + oxaliplatin; SP = S-1 + cisplatin.

Table 2. Weighted linear regression analyses of correlations between surrogate end points and OS

	No. of patients (paired arms)	Intercept	Slope	r^2 (95% CI)	P-value
Median OS ratio *vs* median PFS ratio					
All trials	2148 (19)	0.032	0.624	0.66 (0.32–0.85)	<0.001
Trials with gemcitabine-containing therapies	1933 (15)	0.050	0.623	0.78 (0.46–0.92)	<0.001
Trials with targeted agents	953 (7)	0.112	0.328	0.78 (0.14–0.96)	0.004
Median OS ratio *vs* RR ratio					
All trials	2040 (17)	0.013	0.282	0.29 (0.01–0.65)	0.021
Trials with gemcitabine-containing therapies	1880 (14)	0.020	0.268	0.39 (0.02–0.75)	0.013
Trials with targeted agents	953 (7)	0.119	0.155	0.43 (0.03–0.89)	0.090
Median OS ratio *vs* DCR ratio					
All trials	1989 (17)	−0.038	0.227	0.34 (0.02–0.69)	0.011
Trials with gemcitabine-containing therapies	1832 (14)	−0.037	0.293	0.60 (0.17–0.86)	<0.001
Trials with targeted agents	953 (7)	0.094	0.312	0.44 (0.03–0.89)	0.086
Sensitivity analyses between PFS and OS					
Trials reporting median PFS	1984 (16)	0.026	0.596	0.62 (0.23–0.85)	<0.0001
Trials reporting both HRs for PFS and OS[a]	1287 (9)	−0.075	0.528	0.63 (0.07–0.91)	0.006
Trials with total sample size ⩾100 enroled patients	1392 (8)	0.037	0.597	0.60 (0.02–0.92)	0.015
Not early closed trials	1970 (17)	0.042	0.630	0.64 (0.27–0.86)	<0.001

Abbreviations: CI = confidence interval; DCR = disease control rate; HR = hazard ratio; OS = overall survival; PFS = progression-free survival; RR = response rate; r^2 = coefficient of determination.
[a]The correlation between PFS HRs and OS HRs.

patients received second-line chemotherapy in our analysed trials. Thus, this might not affect the close correlation of PFS and OS in the present study.

It is difficult to interpret whether a high correlation coefficient of a parameter with a true end point is clinically meaningful as a surrogate (Burzykowski and Buyse, 2006). The STE was introduced as a concept, which was defined as the minimum treatment effect on the surrogate necessary to predict a non-zero effect on the true end point. In several studies that demonstrated a strong correlation between PFS and OS, the STE on PFS exceeded 0.80 (Mauguen *et al*, 2013; Paoletti *et al*, 2013; Sidhu *et al*, 2013). Our result is similar, and the high surrogate thresholds were obtained both in all trials (STE 0.83) and in trials with gemcitabine-containing therapies (0.81). Therefore, PFS may be used as a surrogate end point for OS in first-line chemotherapy for ABTC. However, among trials with targeted agents, the STE on PFS could not be calculated and the PFS did not appear to be a valid surrogate end point for OS, despite its strong correlations with OS. The reason for this discrepancy may be attributable to the small number of trials and lack of a significant OS improvement in those trials.

According to the discussions recently held by working groups of the ASCO Research Committee, an improvement in the median OS of at least 20% was generally agreed upon as a clinically meaningful improvement in the outcome of a cancer clinical trial (Ellis *et al*, 2014). Using this benchmark of a minimum 20% improvement in OS, the calculated median PFS ratio in our study was ∼0.70. This value will become the targeted median PFS ratio when planning a randomised clinical trial of ABTC treatment with a superiority design.

This analysis did not reveal a strong correlation between RR and OS or DCR and OS. Recent meta-analyses of other cancers also failed to report strong correlations between these parameters (Burzykowski *et al*, 2008; Sidhu *et al*, 2013; Blumenthal *et al*, 2015). Some recent reports have described correlations between OS and novel parameters such as changes in the tumour volume, depth of response, and early tumour shrinkage in some types of cancers

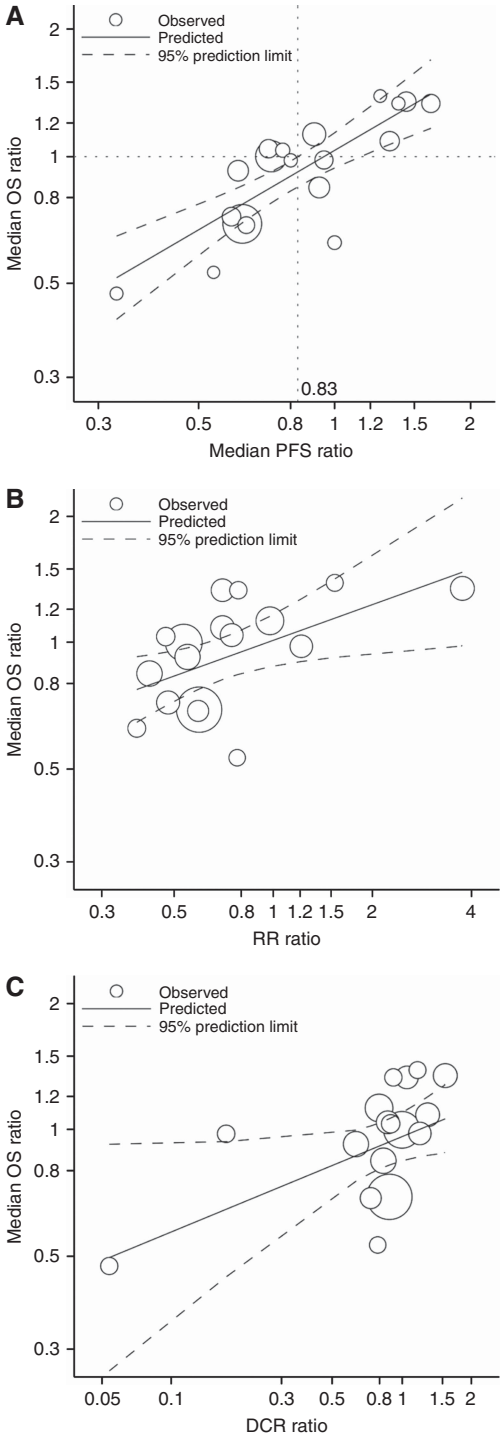

Figure 3. Correlations between treatment effects on surrogacy end points and overall survival (OS) in all selected trials. (**A**) Correlation between median progression-free survival (PFS) ratios and median OS ratios. The point where the horizontal and vertical dotted line crosses indicates the surrogate threshold effect. (**B**) Correlation between median OS ratios and response rate (RR) ratios. (**C**) Correlation between median OS ratios and disease control rate (DCR) ratios. Circle size is proportional to sample size.

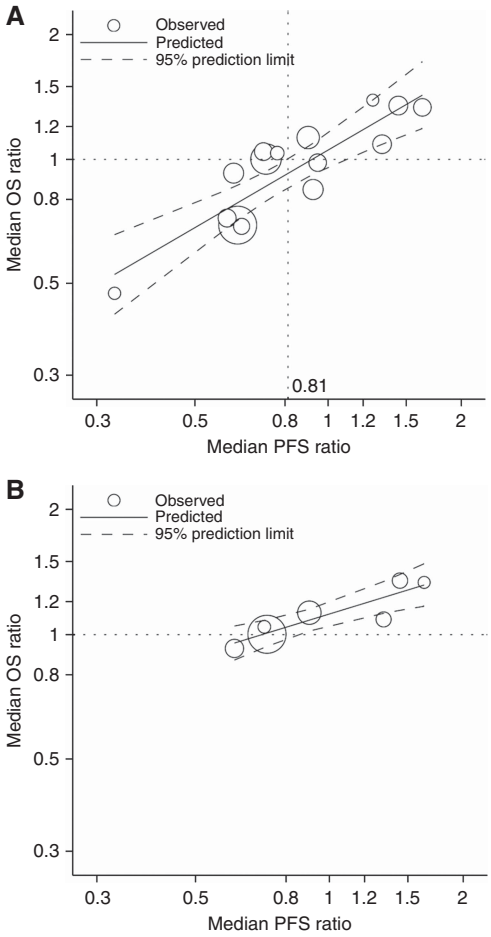

Figure 4. Correlations between median progression-free survival (PFS) ratios and median overall survival (OS) ratios. (**A**) Trials with gemcitabine-containing therapies. (**B**) Trials with targeted agents. The point where the horizontal and vertical dotted line crosses indicates the surrogate threshold effect. Circle size is proportional to sample size.

(Jain *et al*, 2012; Petrelli *et al*, 2015; Sahani *et al*, 2015; Sharma *et al*, 2015; Venook and Tabernero, 2015). However, the usefulness of those parameters as surrogate end points for OS remains unclear.

Our study has a few limitations. First, a majority of the main trials included in this analysis were phase II trials. Only three randomised phase III trials were included, one of which was closed early because of poor accrual. The accuracy decreased in trials with a small sample size. Therefore, we analysed trials with total sample size of $\geqslant 100$ patients and trials excluding early closed trials, and confirmed that the correlations between PFS and OS were similar to those reported in all trials. Second, the correlations between the median PFS and OS ratios were analysed mainly because only a few HRs were reported for OS and PFS. However, the correlation between the reported HRs for PFS and reported HRs for OS was similar to the results of the primary analysis. In addition, the HRs for PFS and OS were found to be reasonably well-represented by the ratios of the medians (Redman *et al*, 2013). Third, the tumour locations varied among biliary tract cancer cases, and poorly visible perihilar tumours, in which progression was difficult to assess, were often included. These might explain the differences in efficacy observed between trials. The reported median PFS and OS in our analysed trials were uneven among the gemcitabine-alone arms and among the gemcitabine + platinum combination therapy arms. Additional analysis involving the proportions of tumour locations among trials might be needed in future. Finally, this study was based on literature, and thus future analyses of individual patient data are needed to confirm the strength of the correlation in ABTC (Buyse *et al*, 2010). To validate our prediction externally, we searched for eligible studies from March 2015 to January 2016. Only two small randomised phase II trials, both of which were presented as abstracts at the ASCO annual meeting 2015, were

identified (Jensen *et al*, 2015; Vogel *et al*, 2015). These trials included both gemcitabine-containing regimens and targeted agents. The reported median OS ratios were 1.39 and 0.60; our calculated OS ratios were 1.24 and 0.91, respectively. There were a small number of trials for the external validation, and it should be done in future.

In conclusion, we found a moderate correlation between PFS and OS, and a relatively good STE value. Our regression model could provide benchmarks to calculate the PFS in order to demonstrate the clinically required improvement of OS. In patients with ABTC, OS could easily be considered the primary end point in a phase III trial, because the post-progression survival is short. Accordingly, our results indicate that PFS is an appropriate end point in a phase II trial of a newly developed drug.

ACKNOWLEDGEMENTS

We thank Ms Kazuko Shinozaki for helping to collect relevant literature. We thank Editage (www.editage.jp) for English language editing.

CONFLICT OF INTEREST

TM has received honoraria from Takeda, Chugai, Yakult, Taiho, and Merck Serono, and has received research funding from Taiho and Boehringer Ingelheim. YY has received honoraria from Takeda and Taiho. MG has received honoraria from Novartis and Taiho. SE has received honoraria from Chugai and Taiho. IH has received honoraria and research funding from Chugai, Yakult, and Taiho, and has received research funding from Dai-ichi Sankyo. The remaining authors declare no conflict of interest.

REFERENCES

Blumenthal GM, Karuri SW, Zhang H, Zhang L, Khozin S, Kazandjian D, Tang S, Sridhara R, Keegan P, Pazdur R (2015) Overall response rate, progression-free survival, and overall survival with targeted and standard therapies in advanced non-small-cell lung cancer: US Food and Drug Administration trial-level and patient-level analyses. *J Clin Oncol* **33**: 1008–1014.

Burzykowski T, Buyse M (2006) Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. *Pharm Stat* **5**: 173–186.

Burzykowski T, Buyse M, Piccart-Gebhart MJ, Sledge G, Carmichael J, Luck HJ, Mackey JR, Nabholtz JM, Paridaens R, Biganzoli L, Jassem J, Bontenbal M, Bonneterre 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**: 1987–1992.

Buyse M, Sargent DJ, Grothey A, Matheson A, de Gramont A (2010) Biomarkers and surrogate end points—the challenge of statistical validation. *Nat Rev Clin Oncol* **7**: 309–317.

Chen JS, Hsu C, Chiang NJ, Tsai CS, Tsou HH, Huang SF, Bai LY, Chang IC, Shiah HS, Ho CL, Yen CJ, Lee KD, Chiu CF, Rau KM, Yu MS, Yang Y, Hsieh RK, Chang JY, Shan YS, Chao Y, Chen LT. Taiwan Cooperative Oncology G (2015) A KRAS mutation status-stratified randomized phase II trial of gemcitabine and oxaliplatin alone or in combination with cetuximab in advanced biliary tract cancer. *Ann Oncol* **26**: 943–949.

Ducreux M, Van Cutsem E, Van Laethem JL, Gress TM, Jeziorski K, Rougier P, Wagener T, Anak O, Baron B, Nordlinger B. Group EGITC (2005) A randomised phase II trial of weekly high-dose 5-fluorouracil with and without folinic acid and cisplatin in patients with advanced biliary tract carcinoma: results of the 40955 EORTC trial. *Eur J Cancer* **41**: 398–403.

Eckel F, Schmid RM (2007) Chemotherapy in advanced biliary tract carcinoma: a pooled analysis of clinical trials. *Br J Cancer* **96**: 896–902.

Ellis LM, Bernstein DS, Voest EE, Berlin JD, Sargent D, Cortazar P, Garrett-Mayer E, Herbst RS, Lilenbaum RC, Sima C, Venook AP, Gonen M, Schilsky RL, Meropol NJ, Schnipper LE (2014) American Society of Clinical Oncology perspective: raising the bar for clinical trials by defining clinically meaningful outcomes. *J Clin Oncol* **32**: 1277–1280.

Jain RK, Lee JJ, Ng C, Hong D, Gong J, Naing A, Wheler J, Kurzrock R (2012) Change in tumor size by RECIST correlates linearly with overall survival in phase I oncology studies. *J Clin Oncol* **30**: 2684–2690.

Jensen L, Fernebro E, Ploen J, Eberhard E, Lindebjerg J, Moeller Jakobsen A (2015) Randomized phase II crossover trial exploring the clinical benefit from targeting EGFR or VEGF with combination chemotherapy in patients with non-resectable biliary tract cancer. *J Clin Oncol* **33**: abstract 4071.

Johnson KR, Liauw W, Lassere MN (2015) Evaluating surrogacy metrics and investigating approval decisions of progression-free survival (PFS) in metastatic renal cell cancer: a systematic review. *Ann Oncol* **26**: 485–496.

Kang MJ, Lee JL, Kim TW, Lee SS, Ahn S, Park do H, Lee SS, Seo DW, Lee SK, Kim MH (2012) Randomized phase II trial of S-1 and cisplatin *versus* gemcitabine and cisplatin in patients with advanced biliary tract adenocarcinoma. *Acta Oncol* **51**: 860–866.

Kornek GV, Schuell B, Laengle F, Gruenberger T, Penz M, Karall K, Depisch D, Lang F, Scheithauer W (2004) Mitomycin C in combination with capecitabine or biweekly high-dose gemcitabine in patients with advanced biliary tract cancer: a randomised phase II trial. *Ann Oncol* **15**: 478–483.

Lamarca A, Hubner RA, David Ryder W, Valle JW (2014) Second-line chemotherapy in advanced biliary cancer: a systematic review. *Ann Oncol* **25**: 2328–2338.

Lee J, Park SH, Chang HM, Kim JS, Choi HJ, Lee MA, Jang JS, Jeung HC, Kang JH, Lee HW, Shin DB, Kang HJ, Sun JM, Park JO, Park YS, Kang WK, Lim HY (2012) Gemcitabine and oxaliplatin with or without erlotinib in advanced biliary-tract cancer: a multicentre, open-label, randomised, phase 3 study. *Lancet Oncol* **13**: 181–188.

Leone F, Marino D, Filippi R, Cereda S, Belli C, Spadi R, Nasti G, Montano M, Amatu A, Aprile G, Cagnazzo C, Fasola G, Siena S, Ciuffreda L, Reni M, Aglietta M (2015) A phase II, open-label, randomized clinical trial of panitumumab plus gemcitabine and oxaliplatin (GEMOX) *versus* GEMOX alone as first-line treatment in advanced biliary tract cancer: the Vecti-BIL study. *J Clin Oncol* **33**: abstract 281.

Malka D, Cervera P, Foulon S, Trarbach T, de la Fouchardiere C, Boucher E, Fartoux L, Faivre S, Blanc JF, Viret F, Assenat E, Seufferlein T, Herrmann T, Grenier J, Hammel P, Dollinger M, Andre T, Hahn P, Heinemann V, Rousseau V, Ducreux M, Pignon JP, Wendum D, Rosmorduc O, Greten TF. investigators B (2014) Gemcitabine and oxaliplatin with or without cetuximab in advanced biliary-tract cancer (BINGO): a randomised, open-label, non-comparative phase 2 trial. *Lancet Oncol* **15**: 819–828.

Mauguen A, Pignon JP, Burdett S, Domerg C, Fisher D, Paulus R, Mandrekar SJ, Belani CP, Shepherd FA, Eisen T, Pang H, Collette L, Sause WT, Dahlberg SE, Crawford J, O'Brien M, Schild SE, Parmar M, Tierney JF, Le Pechoux C, Michiels S. Surrogate Lung Project Collaborative G (2013) Surrogate endpoints for overall survival in chemotherapy and radiotherapy trials in operable and locally advanced lung cancer: a re-analysis of meta-analyses of individual patients' data. *Lancet Oncol* **14**: 619–626.

Moehler M, Maderer A, Schimanski C, Kanzler S, Denzer U, Kolligs FT, Ebert MP, Distelrath A, Geissler M, Trojan J, Schutz M, Berie L, Sauvigny C, Lammert F, Lohse A, Dollinger MM, Lindig U, Duerr EM, Lubomierski N, Zimmermann S, Wachtlin D, Kaiser AK, Schadmand-Fischer S, Galle PR, Woerns M. Working Group of Internal O (2014) Gemcitabine plus sorafenib *versus* gemcitabine alone in advanced biliary tract cancer: a double-blind placebo-controlled multicentre phase II AIO study with biomarker and serum programme. *Eur J Cancer* **50**: 3125–3135.

Morizane C, Okusaka T, Mizusawa J, Takashima A, Ueno M, Ikeda M, Hamamoto Y, Ishii H, Boku N, Furuse J (2013) Randomized phase II study of gemcitabine plus S-1 *versus* S-1 in advanced biliary tract cancer: a Japan Clinical Oncology Group trial (JCOG 0805). *Cancer Sci* **104**: 1211–1216.

Okusaka T, Nakachi K, Fukutomi A, Mizuno N, Ohkawa S, Funakoshi A, Nagino M, Kondo S, Nagaoka S, Funai J, Koshiji M, Nambu Y, Furuse J, Miyazaki M, Nimura Y (2010) Gemcitabine alone or in combination with cisplatin in patients with biliary tract cancer: a comparative multicentre study in Japan. *Br J Cancer* **103**: 469–474.

Paoletti X, Oba K, Bang YJ, Bleiberg H, Boku N, Bouche O, Catalano P, Fuse N, Michiels S, Moehler M, Morita S, Ohashi Y, Ohtsu A, Roth A, Rougier P, Sakamoto J, Sargent D, Sasako M, Shitara K, Thuss-Patience P, Van Cutsem E, Burzykowski T, Buyse M. GASTRIC group (2013) Progression-free survival as a surrogate for overall survival in advanced/recurrent gastric cancer trials: a meta-analysis. *J Natl Cancer Inst* **105**: 1667–1670.

Petrelli F, Pietrantonio F, Cremolini C, Di Bartolomeo M, Coinu A, Lonati V, de Braud F, Barni S (2015) Early tumour shrinkage as a prognostic factor and surrogate end-point in colorectal cancer: a systematic review and pooled-analysis. *Eur J Cancer* **51**: 800–807.

Rao S, Cunningham D, Hawkins RE, Hill ME, Smith D, Daniel F, Ross PJ, Oates J, Norman AR (2005) Phase III study of 5FU, etoposide and leucovorin (FELV) compared to epirubicin, cisplatin and 5FU (ECF) in previously untreated patients with advanced biliary cancer. *Br J Cancer* **92**: 1650–1654.

Redman MW, Goldman BH, LeBlanc M, Schott A, Baker LH (2013) Modeling the relationship between progression-free survival and overall survival: the phase II/III trial. *Clin Cancer Res* **19**: 2646–2656.

Sahani DV, Hayano K, Galluzzo A, Zhu AX (2015) Measuring treatment response to systemic therapy and predicting outcome in biliary tract cancer: comparing tumor size, volume, density, and metabolism. *AJR Am J Roentgenol* **204**: 776–781.

Santoro A, Gebbia V, Pressiani T, Testa A, Personeni N, Arrivas Bajardi E, Foa P, Buonadonna A, Bencardino K, Barone C, Ferrari D, Zaniboni A, Tronconi MC, Carteni G, Milella M, Comandone A, Ferrari S, Rimassa L (2015) A randomized, multicenter, phase II study of vandetanib monotherapy *versus* vandetanib in combination with gemcitabine *versus* gemcitabine plus placebo in subjects with advanced biliary tract cancer: the VanGogh study. *Ann Oncol* **26**: 542–547.

Sasaki T, Isayama H, Nakai Y, Ito Y, Yasuda I, Toda N, Kogure H, Hanada K, Maguchi H, Sasahira N, Kamada H, Mukai T, Okabe Y, Hasebe O, Maetani I, Koike K (2013) A randomized phase II study of gemcitabine and S-1 combination therapy *versus* gemcitabine monotherapy for advanced biliary tract cancer. *Cancer Chemother Pharmacol* **71**: 973–979.

Schinzari G, Mambella G, Cangiano R, Cassano A, Basso M, Ricciotti A, Astone A, De Vizia R, Trigila N, Pozzo C, Barone C (2009) A randomized phase II study comparing 5-FU/LV (deGramont regimen) *versus* 5-FU/LV + oxaliplatin (Folfox 4) in advanced adenocarcinoma of extra- or intrahepatic bile ducts. *Gastrointestinal Cancers Symposium*. abstract 237.

Sharma A, Dwary AD, Mohanti BK, Deo SV, Pal S, Sreenivas V, Raina V, Shukla NK, Thulkar S, Garg P, Chaudhary SP (2010) Best supportive care compared with chemotherapy for unresectable gall bladder cancer: a randomized controlled study. *J Clin Oncol* **28**: 4581–4586.

Sharma MR, Gray E, Goldberg RM, Sargent DJ, Karrison TG (2015) Resampling the N9741 trial to compare tumor dynamic *versus* conventional end points in randomized phase II trials. *J Clin Oncol* **33**: 36–41.

Sidhu R, Rong A, Dahlberg S (2013) Evaluation of progression-free survival as a surrogate endpoint for survival in chemotherapy and targeted agent metastatic colorectal cancer trials. *Clin Cancer Res* **19**: 969–976.

Valle J, Wasan H, Jitlal M, Backen A, Palmer D, Duggan M, Cunningham D, Anthoney A, Corrie P, Madhusudan S, Maraveyas A, Ross P, Waters J, Steward W, Rees C, Beare S, Dive C, Bridgewater J (2014) ABC-03: a randomized phase II trial of cediranib (AZD2171) or placebo in combination with cisplatin/gemcitabine (CisGem) chemotherapy for patients (pts) with advanced biliary tract cancer (ABC). *J Clin Oncol* **32**: abstract 4002.

Valle J, Wasan H, Palmer DH, Cunningham D, Anthoney A, Maraveyas A, Madhusudan S, Iveson T, Hughes S, Pereira SP, Roughton M, Bridgewater J. Investigators ABCT (2010) Cisplatin plus gemcitabine *versus* gemcitabine for biliary tract cancer. *N Engl J Med* **362**: 1273–1281.

Venook AP, Tabernero J (2015) Progression-free survival: helpful biomarker or clinically meaningless end point? *J Clin Oncol* **33**: 4–6.

Vogel A, Kasper S, Weichert W, Bitzer M, Block A, Riess H, Schulze-Bergkamen H, Moehler M, Merx K, Endris V, Schnoy E, Siveke J, Michl P, Waldschmidt D, Kuhlmann J, Geissler M, Kahl C, Kubicka S (2015) Panitumumab in combination with gemcitabine/cisplatin (GemCis) for patients with advanced kRAS WT biliary tract cancer: a randomized phase II trial of the Arbeitsgemeinschaft Internistische Onkologie (AIO). *J Clin Oncol* **33**: abstract 4082.