

Correlation between overall survival and other endpoints in clinical trials of second-line chemotherapy for patients with advanced gastric cancer

Kohei Shitara · Keitaro Matsuo · Kei Muro ·
Toshihiko Doi · Atsushi Ohtsu

Received: 23 March 2013 / Accepted: 15 May 2013 / Published online: 5 June 2013
© The International Gastric Cancer Association and The Japanese Gastric Cancer Association 2013

Abstract

Background The correlation between progression-free survival (PFS) or time to progression (TTP) and overall survival (OS) has been evaluated in patients with advanced gastric cancer (AGC) who received first-line chemotherapy. No corresponding analysis has been done in patients who have undergone second-line chemotherapy.

Methods We evaluated the correlation between PFS, TTP, objective response rate (ORR), disease control rate (DCR), and OS in patients with AGC who underwent second-line chemotherapy. Correlations were evaluated by Spearman rank correlation coefficient (ρ).

Results Sixty-four trials, including 10 randomized studies, were selected for analysis. Median PFS/TTP moderately correlated with OS ($\rho = 0.56$). The correlation tended to be stronger in non-Asian trials ($\rho = 0.74$) than in Asian trials ($\rho = 0.37$). ORR and DCR did not strongly correlate with OS ($\rho = 0.38$ for ORR; $\rho = 0.54$ for DCR). The hazard ratio of PFS and OS in each of the arms of the

10 randomized studies also showed a low correlation ($\rho = 0.36$).

Conclusions PFS/TTP, ORR, and DCR did not correlate sufficiently with OS to be used as surrogate endpoints in patients with AGC who have undergone second-line chemotherapy. Further research is needed based on individual patient data from ongoing randomized trials.

Keywords Chemotherapy · Gastric cancer · Second-line chemotherapy · Surrogate endpoint · Progression-free survival · Time-to-progression

Introduction

Gastric cancer remains one of the most common malignancies and leading causes of cancer death worldwide [1]. The prognosis of patients with advanced or recurrent gastric cancer (AGC) remains poor, with median overall survival (OS) of only 1 year with commonly used first-line combination chemotherapy regimens (fluoropyrimidine plus a platinum agent with or without docetaxel or anthracyclines) [2–7]. Trastuzumab, a humanized monoclonal antibody that targets human epidermal growth factor receptor 2 (HER2), has recently been shown to improve the prognosis of HER2-positive AGC [7], although these cases account for fewer than 20 % of all AGCs. Because median progression-free survival (PFS) associated with these first-line chemotherapies is around 6 months and most patients ultimately experience disease progression, development of effective second-line chemotherapy is critical. Several phase II studies of second-line chemotherapy have suggested that taxanes (paclitaxel or docetaxel) or irinotecan can be effective, with corresponding objective response rates (ORRs) of approximately 10–20 %. Recently, a small

Electronic supplementary material The online version of this article (doi:10.1007/s10120-013-0274-6) contains supplementary material, which is available to authorized users.

K. Shitara · T. Doi · A. Ohtsu
Department of Gastrointestinal Oncology, National Cancer Center Hospital East, Chiba, Japan

K. Shitara (✉) · K. Muro
Department of Clinical Oncology, Aichi Cancer Center Hospital,
1-1 Kanokoden, Chikusa-ku, Nagoya, Aichi 464-8681, Japan
e-mail: kouheis0824@yahoo.co.jp

K. Matsuo
Division of Epidemiology and Prevention, Aichi Cancer Center
Research Institute, Nagoya, Japan

randomized study suggested that irinotecan improved outcomes in patients with pretreated AGC [8]. Another randomized study that compared docetaxel or irinotecan and best supportive care for AGC patients with one or two previous lines of chemotherapy also showed the survival benefit of salvage chemotherapy (OS, 5.8 vs. 3.8 months) [9].

Correlations between PFS or other endpoints and OS have been analyzed in an effort to identify surrogate endpoints of OS [10–15]. A validated shorter-term surrogate endpoint would likely both reduce drug development costs and facilitate the assessment of efficacy [16]. Previously, a literature-based analysis and an individual patient data meta-analysis evaluated PFS as surrogate endpoint for OS in patients with AGC who underwent first-line chemotherapy [14, 15]. However, no corresponding analysis had been done in patients who underwent second-line chemotherapy for AGC. Thus, the goal of the present study was to conduct a comprehensive analysis of the correlation between PFS or other endpoints and OS in patients with AGC who underwent second-line chemotherapy.

Materials and methods

Search for studies

We conducted a literature search for trials through computer-based searches of the Medline database (January 2002 and January 2013) and of abstracts from conference proceedings of the American Society of Clinical Oncology (2002–2012), Gastrointestinal Cancer Symposium (2002–2013), and European Cancer Conference and European Society for Medical Oncology (2002–2012). To avoid publication bias, both published and unpublished trials were identified. Data were gathered as possible from presentations in meeting as well as abstracts.

Search keywords included “gastric cancer” and “second-line chemotherapy.” The search was also guided by a thorough examination of reference lists of original and review articles. No limitation based on language was defined. We included unpublished data if sufficient information on study design, characteristics of participants, interventions, and outcomes was available from an abstract or meeting presentation.

Procedures

The data were abstracted in accordance with the Quality of Reporting of Meta-analyses (QUORUM) guidelines [17]. Prospective trials (single-arm or randomized trials) of chemotherapy for chemotherapy-pretreated adenocarcinoma (metastatic disease or unresectable locally advanced

disease or recurrent) of the stomach or gastroesophageal junction were included in the analysis. Because some trials included patients who received experimental treatments as second-line or third-line chemotherapy, these studies were also included. However, we excluded studies in which all patients received experimental treatments as third-line chemotherapy. Trials that compared chemotherapy with best supportive care were also included, as were those that included patients with adenocarcinoma of the distal esophagus. Eligibility was limited to trials that reported data on OS with either or both PFS and TTP. Exclusion criteria included trials designed to assess combined modality treatments, including radiotherapy and surgery (neoadjuvant or adjuvant chemotherapy).

For each trial, the following information was extracted: first author’s name; year of publication or report; trial design; trial region; number of enrolled patients; treatment regimens. The following data were also extracted if reported: previous treatment regimens, and proportion of patients with measurable lesions. For trials with more than two treatment arms, we constructed multiple pairs of each investigational arm and the reference arm.

Statistical methods

For each trial, median PFS, TTP, ORR, disease control rate (DCR; proportion of patients who achieved complete or partial response or stable disease), and OS were abstracted. In the case of randomized studies, hazard ratio (HR) with 95 % confidence intervals (CI) for clinical outcome (PFS/TTP and OS) was also abstracted. If the HR was not provided, we estimated HR and 95 % CI as relevant effect measures directly or indirectly from the given data [18]. The nonparametric Spearman rank correlation coefficient (ρ) was used as a measure of correlation between the median PFS/TTP and OS and of correlation between HR of PFS/TTP and HR of OS. As the number of subject studies was limited, we applied bootstrap resampling [19] using 10,000 bootstrap samples to estimate 95 % CI for correlation coefficients.

To investigate possible reasons for heterogeneity of correlation, subgroup analyses were conducted according to trial region (Asian vs. non-Asian), reported data (old trials; before 2009 vs. recent trials; 2009 or later), status of publication (published vs. presentation only), endpoint for progression (PFS vs. TTP), previous chemotherapy regimens [fluoropyrimidine plus platinum (FP) mandatory vs. not defined], treatment line (second-line only vs. second-line and third-line) and treatment regimens (taxane-based vs. irinotecan-based). In the case of global trials, data were classified as both Asian and non-Asian unless suitable subset analysis results were provided. Median values of each endpoint were calculated, and differences in subsets

were evaluated using the Mann–Whitney test. Statistical analyses were performed using STATA ver. 10 (Stata Corp., College Station, TX, USA). All tests were two sided, and p values less than 0.05 were considered statistically significant.

Results

Selection of studies

A total of 640 potentially relevant reports were identified, of which 472 were initially excluded by title view (Fig. 1). After review of the remaining studies, 64 trials were identified as eligible for this meta-analysis, including a total of 75 treatment arms and 4,286 patients (Supplement 1). Forty-four trials were published, and another 20 trials were presentations or abstracts only. Table 1 shows the characteristics of the 64 trials. Only 10 trials were randomized trials (5 phase II and 5 phase III), and 54 were single-arm phase II studies. By region, 39 were conducted in Asia, 23 were conducted in non-Asia regions, and 2 were global studies that included Asia. Sixteen trials included only patients who received a previous regimen that included FP as first-line chemotherapy. Forty-nine trials included only patients with measurable lesions. Forty-one studies described disease progression with previous chemotherapy as inclusion criteria. The most common primary endpoint was ORR ($n = 39$), followed by OS ($n = 10$). Only 16 studies assessed tumor response by independent review. Most commonly used regimens were taxanes followed by irinotecan or platinum-based therapy. As a time to event for progression, more studies reported PFS ($n = 41$) than TTP ($n = 23$), whereas no trial reported both PFS and TTP. Subset analysis according to region (Asia and non-Asia) was reported in one global phase II

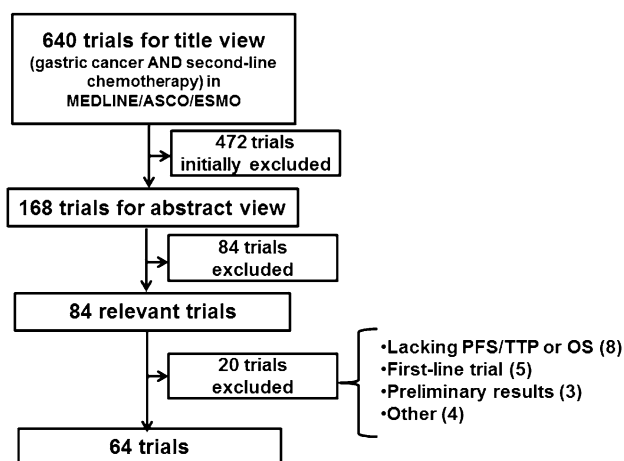


Fig. 1 Selection process for trials. PFS/TTP progression-free survival/time to progression; OS overall survival

trial, and these subset data were accordingly included in analyses that focused on comparing Asian and non-Asian trials.

Results of each endpoint according to subsets

Median value of reported OS among the 64 trials was 7.6 months, and median PFS or TTP was 3.0 months (Table 2). Median OS tended to be longer in Asian trials than in non-Asian trials (8.1 vs. 6.0 months; $p < 0.001$). In contrast, median PFS or TTP were not significantly different when comparing Asian and non-Asian trials (3.0 vs. 3.1 months; $p = 0.19$). Unpublished trials were associated with longer OS than published trials (8.1 vs. 6.7 months; $p = 0.02$). No other subset analysis showed significant

Table 1 Characteristics of the 60 clinical trials analyzed in the present study

Characteristic	<i>n</i>	%
Reported year		
Before 2009	28	44
2009–2012	36	56
Trial setting		
Single-arm phase II	54	84
Randomized phase II	5	8
Phase III	5	8
Trial area		
Asia	39	61
Non-Asia	22	34
Global, including Asia	3	5
Previous chemotherapy		
Fluoropyrimidine and platinum agents mandatory	16	25
Various	48	75
Inclusion criteria		
Measurable lesion mandatory	49	77
Primary endpoint		
Objective response rate	39	61
Overall survival	10	16
Progression-free survival or time to progression	5	8
Disease control rate	2	3
Not reported or not available	8	13
Treatment line		
Second-line only	49	77
Second- and third-line	15	23
Investigated agents ^a		
Taxanes	32	50
Irinotecan	26	41
Fluoropyrimidine	21	33
Platinum agents	18	28
Others	25	39

^a Among 75 treatment arms, some overlapped

Table 2 Results of each endpoint according to subsets

Subset	Number of arms	Median OS (months)	<i>p</i> value	Median PFS/TTP (months)	<i>p</i> value	ORR (%)	<i>p</i> value	DCR (%)	<i>p</i> value
All	75	7.6		3.0		17.9		53.8	
Trial area ^a									
Asia	47	8.1	<0.001	3.0	0.19	20.2	0.25	55.4	0.21
Non-Asia	23	6.0		3.1		15.5		50.0	
Reported year									
Before 2009	37	7.2	0.08	3.5	0.26	19.2	0.15	53.1	0.72
2009 or later	38	7.8		2.9		16.1		54.6	
Publication									
Published	46	6.7	0.02	3.1	0.86	18.9	0.11	55.0	0.98
Presentation only	29	8.1		3.0		20.0		52.2	
Endpoint									
PFS	51	7.7	0.08	3.0	0.50	17.0	0.08	52.2	0.42
TTP	24	7.0		3.6		20.6		55.0	
Measurable lesion									
Mandatory	52	7.0	0.07	3.0	0.63	18.2	0.36	55.0	0.35
Not mandatory	23	8.2		3.0		17.0		46.2	
Previous chemotherapy									
FP mandatory	19	6.6	0.13	2.9	0.22	14.8	0.29	50.5	0.48
Not defined or other	56	7.7		3.1		18.3		54.8	
Treatment line									
Second-line only	59	7.6	0.41	3.3	0.11	18.4	0.30	55.0	0.09
Second- and third-line	16	6.6		2.6		17.0		48.0	
Regimen ^b									
Taxane-based	31	8	0.31	3.6	0.78	17.5	0.41	58.0	0.34
Irinotecan-based	26	7.6		3.4		18.5		53.0	

OS overall survival, PFS progression-free survival, TTP time to progression, ORR objective response rate, DCR disease control rate

Statistical analyses were performed using the Mann–Whitney test, with the level of significance set at $p < 0.05$ (italicized)

^a Excluded two global trials

^b Excluded arm of taxane plus irinotecan or other regimens

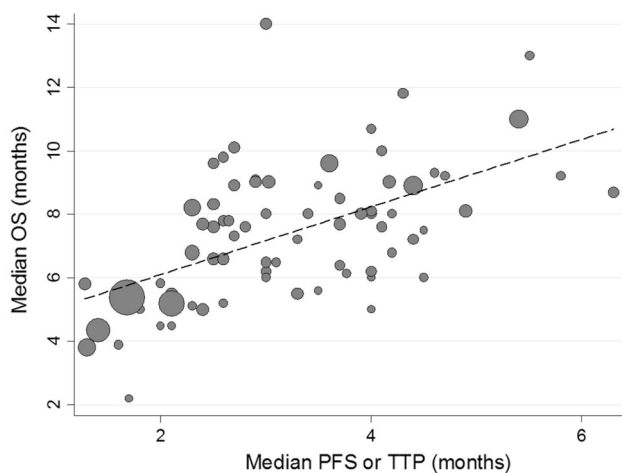


Fig. 2 Correlation between median progression-free survival/time to progression (PFS/TTP) and overall survival (OS). Size of gray markers (circles) corresponds to the number of randomized patients in the trial in this analysis. Median PFS or TTP and OS were moderately correlated ($r = 0.51$, 95 % CI 0.31–0.71)

differences in OS or PFS/TTP. Median reported ORR and DCR were 17.9 % and 53.8 %, respectively. DCR tend to be higher in trials of second-line only therapy when compared with trials of second- and third-line therapy ($p = 0.09$), although no other subset showed significant differences in DCR.

Correlation between PFS or TTP and OS

Median PFS or TTP and OS were moderately correlated ($\rho = 0.56$, 95 % CI 0.34–0.74; Fig. 2; Table 3). The correlation tended to be stronger with PFS ($\rho = 0.65$) than with TTP ($\rho = 0.28$), stronger in non-Asian trials ($\rho = 0.74$) than in Asian trials ($\rho = 0.37$; Fig. 3; Table 3), and stronger in trials with second-line and third-line chemotherapy ($\rho = 0.47$) than in trials of second-line therapy only ($\rho = 0.77$). The correlation was almost similar when comparing published trials vs. presentation only ($\rho = 0.52$, $\rho = 0.60$).

Table 3 Correlation between PFS/TTP, ORR, DCR, and OS

Subset	PFS/TTP and OS		ORR and OS		DCR and OS	
	ρ	95 % CI	ρ	95 % CI	ρ	95 % CI
All	0.56	0.37 to 0.74	0.38	0.16 to 0.61	0.54	0.33 to 0.75
Trial area ^a						
Asia	0.37	0.10 to 0.63	0.27	−0.01 to 0.55	0.43	0.12 to 0.74
Non-Asia	0.74	0.50 to 0.98	0.35	−0.10 to 0.80	0.66	0.35 to 0.97
Reported year						
Before 2009	0.47	0.16 to 0.77	0.13	−0.22 to 0.47	0.49	0.13 to 0.76
2009 or later	0.64	0.43 to 0.86	0.59	0.32 to 0.86	0.63	0.33 to 0.92
Publication						
Published	0.52	0.29 to 0.75	0.31	0.02 to 0.61	0.55	0.30 to 0.79
Presentation only	0.60	0.38 to 0.93	0.75	0.55 to 0.96	0.47	0.02 to 0.93
Endpoint						
PFS	0.65	0.46 to 0.83	0.56	0.33 to 0.80	0.63	0.41 to 0.84
TTP	0.28	−0.17 to 0.73	0.07	−0.39 to 0.54	0.23	−0.28 to 0.73
Measurable lesion						
Mandatory	0.51	0.28 to 0.74	0.31	0.03 to 0.59	0.51	0.27 to 0.75
Not mandatory	0.59	0.27 to 0.92	0.69	0.35 to 1.00	0.78	0.36 to 1.00
Previous chemotherapy						
FP only	0.55	0.20 to 0.91	0.42	0.02 to 0.83	0.37	−0.09 to 0.82
Not defined or others	0.55	0.33 to 0.77	0.36	0.08 to 0.63	0.61	0.38 to 0.84
Treatment line						
Second-line only	0.47	0.24 to 0.70	0.39	0.16 to 0.63	0.39	0.12 to 0.66
Second- and third-line	0.77	0.49 to 1.00	0.23	−0.37 to 0.84	0.89	0.72 to 1.00
Regimen ^b						
Taxane-based	0.35	−0.01 to 0.71	0.39	0.05 to 0.73	0.27	−0.20 to 0.75
Irinotecan-based	0.46	0.10 to 0.81	0.09	−0.28 to 0.46	0.56	0.20 to 0.92

OS overall survival, PFS progression-free survival, TTP time to progression, ORR objective response rate, FP fluoropyrimidine and platinum agents

^a Excluded one global trial

^b Excluded arm of taxane plus irinotecan or other regimens

Correlation between ORR, DCR, and OS

The ORR and DCR was not strongly correlated with OS ($\rho = 0.38$ for ORR, 95 % CI 0.16–0.61; $\rho = 0.54$ for DCR, 95 % CI 0.33–0.75; Fig. 4), although DCR was more strongly correlated with OS when compared with ORR vs. OS in the whole cohort or any subset (Table 3).

Correlation between HR for PFS/TTP and OS in randomized trials

A total of 11 pairs of HRs for PFS/TTP and OS between treatment arms were available from the 10 randomized trials (reported in 9 trials and estimated in 1 trial). The HR of PFS/TTP and OS in each arm showed a low correlation ($\rho = 0.36$, 95 % CI −0.30 to 1.00; Fig. 5). Wide 95 % CI indicated that the sample sizes were too small for this type of analysis.

Discussion

This is the first study to evaluate the correlation between PFS, TTP, or other endpoints and OS in patients with AGC who underwent second-line chemotherapy for AGC. Our results suggests that PFS/TTP, ORR, and DCR did not correlate sufficiently with OS to be used as surrogate endpoints for OS in patients with AGC who underwent second-line chemotherapy. We should interpret our results cautiously because this study is of exploratory nature and has the following several limitations. (1) Our analysis is based on literature-based data without individual patient data. (2) Most of the included studies were single-arm studies, and only ten of the studies were randomized trials. (3) Little information was available about subsequent treatment including crossover treatment, which may weaken the surrogacy. Against these limitations, we consider that our work could convey important aspects with

Fig. 3 Correlation between median PFS/TTP and OS according to trial area. The correlation tended to be stronger in non-Asian trials ($\rho = 0.74$) than in Asian trials ($\rho = 0.37$)

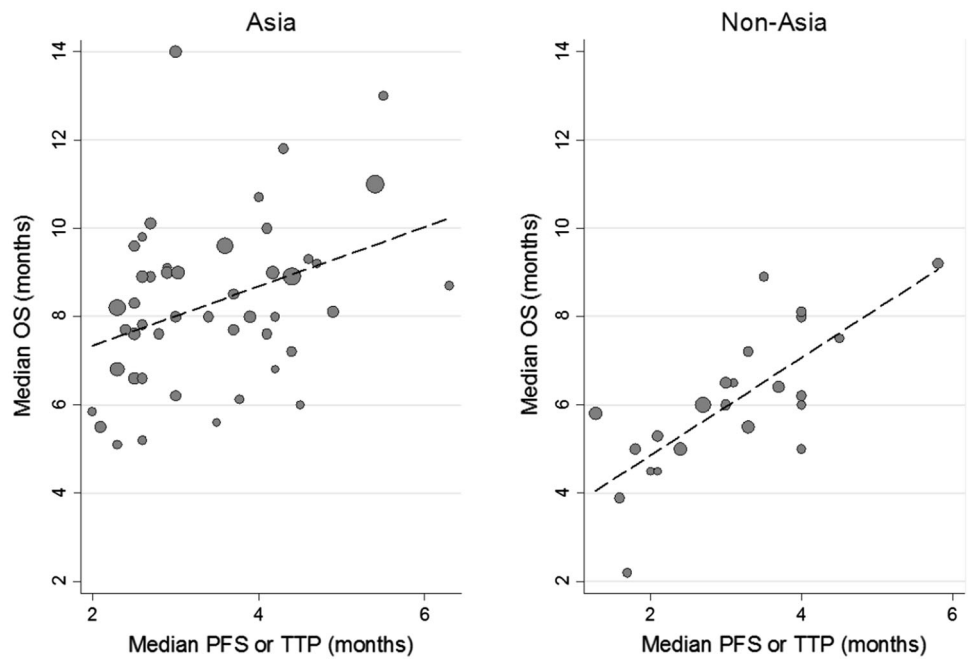
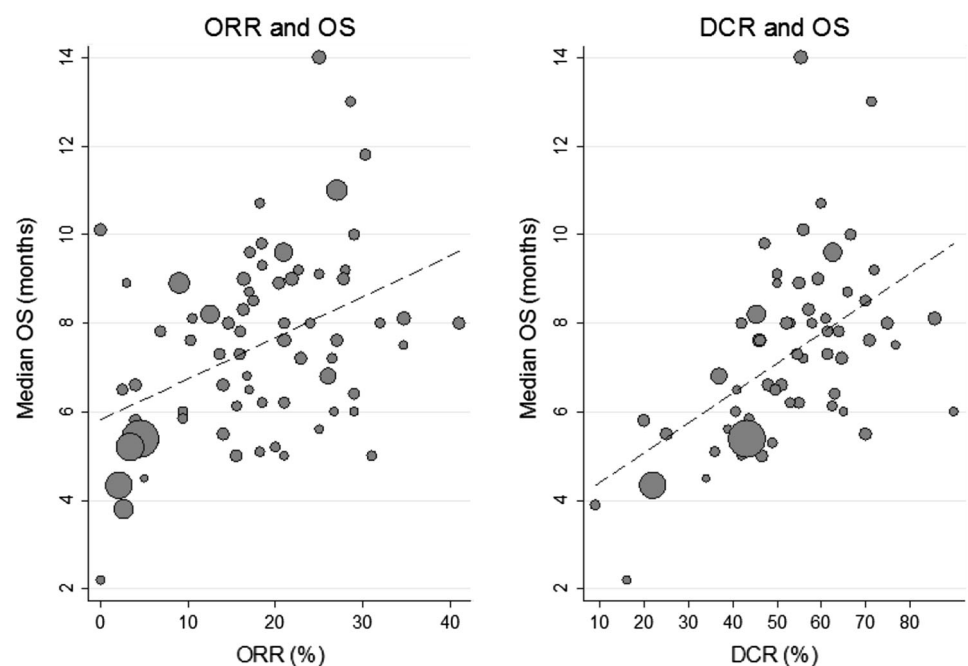


Fig. 4 Correlation between objective response rate (ORR) or disease control rate (DCR) and OS. ORR and DCR were not strongly correlated with OS ($\rho = 0.38$ for ORR, 95 % CI 0.16–0.61; $\rho = 0.54$ for DCR, 95 % CI 0.33–0.75)



regard to the trial conduct and data collection for the future trials of second-line therapy for advanced gastric cancers.

Previously, two meta-analyses studied whether PFS could be a surrogate endpoint for OS in patients with AGC who underwent first-line chemotherapy [14, 15]. According to a literature-based analysis of 36 randomized trials [14], median PFS or TTP moderately correlated with median OS ($\rho = 0.70$). The correlation coefficient between HR of PFS or TTP and OS was 0.80. Another meta-analysis called the GASTRIC project (Global Advanced/Adjuvant Stomach

Tumor Research through International Collaboration) analyzed data from 4,102 AGC patients included in 20 randomized trials [15]. The correlation between treatment effects on PFS and OS in each trial was only moderate (trial-level decision coefficient R^2 adjusted for estimation errors was 0.61), which is the same strength of relationship seen in the literature-based analysis [14]. Correlations between PFS and OS were lower for AGC than for those in patients with advanced colorectal cancer [10] or for those seen in studies of adjuvant treatment for colorectal cancer

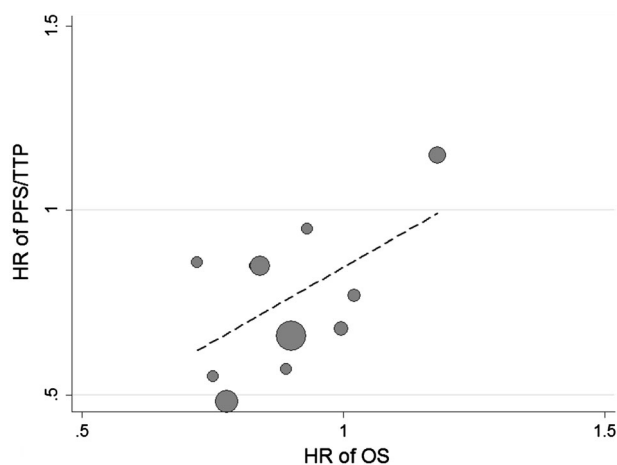


Fig. 5 Hazard ratio (HR) of PFS/TTP and OS in ten randomized studies. The HR/TTP of PFS was moderately correlated with OS in each arm ($\rho = 0.36$, 95 % CI -0.30 to 1.00)

or gastric cancer [20, 21]. These results suggest that PFS is not a good surrogate for OS in patients undergoing first-line chemotherapy for AGC.

Recently, it has been suggested that second-line chemotherapy prolonged the OS of patients with AGC, according to two randomized studies [8, 9]. Therefore, we conducted a literature-based analysis of endpoint of clinical trials patients who underwent second-line chemotherapy for AGC. The present analysis showed that there was an insufficient correlation between OS and other endpoints, which is similar to data observed in the first-line setting. There are several possible reasons for these results. First, heterogeneity of treatment, especially in terms of subsequent chemotherapy, may affect the results. In this analysis, median PFS was almost the same when comparing Asian trials and non-Asian trials, whereas OS was significantly longer in Asian trials when compared with non-Asian trials. One possible reason for this difference in survival after progression is the effect of subsequent treatment, as already suggested in the first-line setting [22]. Indeed, the proportion of patients who receive subsequent chemotherapy is higher in Asian trials than in Western trials [22, 23]; in the AVAGAST (a study of bevacizumab in combination with capecitabine and cisplatin as first-line therapy in patients with AGC) study, 66 % of Asian patients received second-line chemotherapy compared with 31 % of patients in Europe and 21 % in America [23]. Although the proportion of patients who can receive subsequent therapy is expected to be lower in second-line trials than in first-line trials, 40 % of patients in Korean randomized studies received subsequent therapy after second- and third-line chemotherapy [9]. Also, in the West Japan Oncology Group (WJOG) 4407 study, which compared irinotecan and weekly paclitaxel as second-line chemotherapy, more than 70 % of patients received third-line chemotherapy in both arms [24]. Therefore, subsequent

therapy may contribute to the difference in OS according to trial area and confound the correlation in the current analysis, similar to the phenomenon seen in a previous analysis [14].

Another possible reason of moderate correlation of PFS and OS may be heterogeneity in inclusion criteria and patient characteristics. Types of prior chemotherapy before enrollment or investigational agents were quite variable in this population. Also, the definition of failure of prior chemotherapy varied between source studies. Although subset analysis according to prior treatment or treatment regimens did not show a strong correlation between each endpoint, these heterogeneities may contribute to the weak correlation between each endpoint in our analysis. Further, although most studies included patients with measurable lesions, the Japan Clinical Oncology Group (JCOG) 0407 study included patients with peritoneal metastasis, which is associated with a low frequency of measurable lesions [25]. By contrast, the WJOG4007 study excluded patients with apparent peritoneal metastasis [24]. These variations in inclusion criteria might affect the results of correlation.

Although this study showed that there was an insufficient correlation between OS and all endpoints examined, the correlation between ORR and OS was much weaker than that between PFS, TTP or DCR, and OS. These results suggest that a single-arm phase II study with a primary endpoint of ORR may not be adequate to evaluate the efficacy of second-line chemotherapy for AGC. Randomised phase II studies that compare standard treatments and investigational treatments may be better methods of screening for effective treatments to include within phase III trials [26].

This study has several methodological limitations. First, as already described, most of the component studies were single-arm studies, and only ten of the studies were randomized trials. Although there is no consensus in terms of what defines a valid surrogate endpoint, any candidate endpoint must correlate with the true endpoint, and effects on the surrogate endpoint must correlate with those on the true endpoint [27, 28]. However, the effect of each treatments on the surrogate endpoints may be difficult to analyze in this case, as there were relatively few randomized trials available. Second, the present study was not based on an analysis of data from individual patients, which is a confirmatory method of evaluating individual-level measures of agreement between the two endpoints (PFS/TTP and OS) [29]. Additional individual data analysis, especially using ongoing randomized studies, might therefore be necessary to characterize the surrogacy of endpoints. Finally, most trials analyzed in this study provided little information on disease progression after prior chemotherapy, and only a few studies evaluated patient responses by external review. Also, interval to evaluation imaging is also

varied. Therefore, it is impossible to confirm whether the evaluation of disease progression was consistent among the trial arms.

In conclusion, our exploratory analysis suggests that PFS/TTP, ORR, and DCR do not correlate sufficiently with OS to be used as surrogate endpoints in patients with AGC who have undergone second-line chemotherapy. Further research is needed based on individual patient data from ongoing randomized trials to evaluate an optimal surrogate endpoint.

Conflict of interest None of the authors has financial or personal conflicts of interest to disclose.

References

- Kamangar F, Dores GM, Anderson WF. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *J Clin Oncol*. 2006;24:2137–50.
- Van Cutsem E, Moiseyenko VM, Tjulandin S, Majlis A, Constenla M, Boni C, et al. Phase III study of docetaxel and cisplatin plus fluorouracil compared with cisplatin and fluorouracil as first-line therapy for advanced gastric cancer: a report of the V325 Study Group. *J Clin Oncol*. 2006;24:4991–7.
- Cunningham D, Starling N, Rao S, Iveson T, Nicolson M, Coxon F, et al. Capecitabine and oxaliplatin for advanced esophagogastric cancer. *N Engl J Med*. 2008;358:36–46.
- Koizumi W, Narahara H, Hara T, Takagane A, Akiya T, Takagi M, et al. S-1 plus cisplatin versus S-1 alone for first-line treatment of advanced gastric cancer (SPIRITS trial): a phase III trial. *Lancet Oncol*. 2008;9:215–21.
- Kang YK, Kang WK, Shin DB, Chen J, Xiong J, Wang J, et al. Capecitabine/cisplatin versus 5-fluorouracil/cisplatin as first-line therapy in patients with advanced gastric cancer: a randomised phase III noninferiority trial. *Ann Oncol*. 2009;20:666–73.
- Ajani JA, Rodriguez W, Bodoky G, Moiseyenko V, Lichinitser M, Gorbunova V, et al. Multicenter phase III comparison of cisplatin/S-1 with cisplatin/infusional fluorouracil in advanced gastric or gastroesophageal adenocarcinoma study: the FLAGS trial. *J Clin Oncol*. 2010;28:1547–53.
- Bang YJ, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet*. 2010;376:687–97.
- Thuss-Patience PC, Kretzschmar A, Deist T, Hinke A, Bichev D, Lebedinzew B, et al. Irinotecan versus best supportive care (BSC) as second-line therapy in gastric cancer: a randomized phase III study of the Arbeitsgemeinschaft Internistische Onkologie (AIO). *J Clin Oncol* 2009;27:15s (suppl abstr 4540).
- Kang JH, Lee SI, Lim do H, Park KW, Oh SY, Kwon HC, et al. Salvage chemotherapy for pretreated gastric cancer: a randomized phase III trial comparing chemotherapy plus best supportive care with best supportive care alone. *J Clin Oncol*. 2012;30:1513–8.
- Buyse M, Burzykowski T, Carroll K, Michiels S, Sargent DJ, Miller LL, et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. *J Clin Oncol*. 2007;25:5218–24.
- Burzykowski T, Buyse M, Piccart-Gebhart MJ, Sledge G, Carmichael J, Lück HJ, et al. Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. *J Clin Oncol*. 2008;26:1987–92.
- Johnson KR, Ringland C, Stokes BJ, Anthony DM, Freemantle N, Irs A, et al. Response rate or time to progression as predictors of survival in trials of metastatic colorectal cancer or non-small-cell lung cancer: a meta-analysis. *Lancet Oncol*. 2006;7:741–6.
- Hotta K, Fujiwara Y, Matsuo K, Kiura K, Takigawa N, Tabata M, et al. Time to progression as a surrogate marker for overall survival in patients with advanced non-small cell lung cancer. *J Thorac Oncol*. 2009;4:311–7.
- Shitara K, Ikeda J, Yokota T, Takahari D, Ura T, Muro K, et al. Progression-free survival and time to progression as surrogate markers of overall survival in patients with advanced gastric cancer: analysis of 36 randomized trials. *Invest New Drugs*. 2012;30:1224–31.
- Shitara K, Burzykowski T; on behalf of GASTRIC project. Progression-free survival as surrogate endpoint of overall survival in patients with advanced/recurrent gastric cancer: Individual patient data analysis on 4,102 patients from 20 randomized trials. *J Clin Oncol* 2011;29(suppl abstr 4095).
- Saad ED, Katz A, Hoff PM, Buyse M. Progression-free survival as surrogate and as true end point: insights from the breast and colorectal cancer literature. *Ann Oncol*. 2010;21:7–12.
- Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet*. 2008;371:281–3.
- Parmar MB, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Stat Med*. 1998;17:2815–34.
- Efron B. Bootstrap methods: another look at the jackknife. *Ann Stat*. 1979;7:1–26.
- Sargent DJ, Patiyil S, Yothers G, Haller DG, Gray R, Benedetti J, et al. End points for colon cancer adjuvant trials: observations and recommendations based on individual patient data from 20,898 patients enrolled onto 18 randomized trials from the ACCENT Group. *J Clin Oncol*. 2007;25:4569–74.
- Burzykowski T, Bang Y; on behalf of the GASTRIC project. Disease-free survival as a surrogate endpoint for overall survival in an adjuvant trial of curatively resected stomach cancer using individual patient data meta-analysis. *J Clin Oncol* 2009;27:15s(suppl abstr 4517).
- Shitara K, Ikeda J, Kondo C, Takahari D, Ura T, Muro K, et al. Reporting patient characteristics and stratification factors in randomized trials of systemic chemotherapy for advanced gastric cancer. *Gastric Cancer*. 2012;15:137–43.
- Ohtsu A, Shah MA, Van Cutsem E, Rha SY, Sawaki A, Park SR, et al. Bevacizumab in combination with chemotherapy as first-line therapy in advanced gastric cancer: a randomized, double-blind, placebo-controlled phase III study. *J Clin Oncol*. 2011;29:3968–76.
- Ueda S, Hironaka S, Yasui H, Nishina T, Tsuda M, Tsumura T, et al. Randomized phase III study of irinotecan (CPT-11) versus weekly paclitaxel (wPTX) for advanced gastric cancer (AGC) refractory to combination chemotherapy (CT) of fluoropyrimidine plus platinum (FP): WJOG4007 trial. *J Clin Oncol* 2012;30(suppl abstr 4002).
- Takiuchi H, Fukuda H, Boku N, Shimada Y, Nasu J, Hamamoto Y, et al. Randomized phase II study of best-available 5-fluorouracil (5-FU) versus weekly paclitaxel in gastric cancer (GC) with peritoneal metastasis (PM) refractory to 5-FU-containing regimens (JCOG0407). *J Clin Oncol* 2010;28:15s (suppl abstr 4052).
- Tang H, Foster NR, Grothey A, Ansell SM, Goldberg RM, Sargent DJ. Comparison of error rates in single-arm versus randomized phase II cancer clinical trials. *J Clin Oncol*. 2010;28:1936–41.

27. Fleming TR. Surrogate endpoints and FDA's accelerated approval process. *Health Aff (Millwood)*. 2005;24:67–8.
28. Baker SG. Surrogate endpoints: wishful thinking or reality? *J Natl Cancer Inst*. 2006;98:502–3.
29. Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics*. 2000;1:49–67.