ELMER PRESS

Original Article

World J Oncol. 2022;13(3):126-135

Surrogate Endpoints for Overall Survival in Immune-Oncology Trials of Advanced Gastro-Esophageal Carcinoma

Yuan Fang Li[a, b, c, g], Yun Wang[b, c, d, g], Jie Zhou[b, c, e, g], Yi Cheng Wei[a, b, c, g], Jun Lin[a, b, c], Yi Xin Yin[b, c, e], Guo Ming Chen[a, b, c], Fei Yang Zhang[a, b, c], Shi Chen[f], Zhi Wei Zhou[a, b, c], Ying Bo Chen[a, b, c, h], Run Cong Nie[a, b, c, h]

Abstract

Background: We aimed to assess whether the Response Evaluation Criteria in Solid Tumors (RECIST)-based objective response rate (ORR), disease control rate (DCR) and progression-free survival (PFS) could serve as surrogate endpoints for overall survival (OS) in immune-oncology (IO) trials of advanced gastro-esophageal (GE) carcinoma.

Methods: Randomized controlled trials (RCTs) of IO that reported RECIST-based endpoints and OS in advanced GE carcinoma were screened. Surrogacy of endpoints for OS was assessed based on the correlation between endpoints with OS (arm-level), and between treatment effects on endpoints (trial-level). The correlations were quantified by Pearson correlation coefficient (R). Leave-one-out cross-validation was used to assess the prediction accuracy of surrogate model.

Results: Seventeen RCTs (9,657 subjects) with 20 comparisons were included. The correlations between DCR and OS were not strong at arm- (R = 0.80) and trial-levels (R = 0.45), but strong correlations between ORR (R = 0.91), PFS (R = 0.89) and OS at arm-level were observed. Treatment effect on ORR and PFS (both R = 0.71) was moderately correlated with treatment effect on OS. Leave-one-out cross-validation approach further validated the surrogacy of PFS. Our analysis showed that 3-month PFS could reliably predict 6-month OS, 6-month PFS could reliably predict 12-month OS, and 12-month PFS could reliably predict 18-month OS. The conservative minimum threshold effect of HR_{PFS} was 0.73.

Conclusions: PFS may be the appropriate surrogate for OS in IO trials of GE carcinoma. A conservative minimum threshold effect of $HR_{PFS} \leq 0.73$ has the potential to predict a significant improvement in OS.

Keywords: PD-1; PD-L1; Immune checkpoint inhibitor; Surrogate endpoint; Overall survival; Gastro-esophageal carcinoma

Manuscript submitted March 31, 2022, accepted May 19, 2022
Published online June 16, 2022

[a]Department of Gastric Surgery, Sun Yat-sen University Cancer Center, Guangzhou, China
[b]State Key Laboratory of Oncology in South China, Guangzhou, China
[c]Collaborative Innovation Center for Cancer Medicine, Guangzhou, China
[d]Department of Hematologic Oncology, Sun Yat-sen University Cancer Center, Guangzhou, China
[e]Sun Yat-sen University Cancer Center, Guangzhou, China
[f]Department of Gastrointestinal Surgery, the Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou, China
[g]These authors contributed equally to this study.
[h]Corresponding Author: Run Cong Nie and Ying Bo Chen, Department of Gastric Surgery, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong 510060, China. Email: nierc@sysucc.org.cn and chenyb@sysucc.org.cn

doi: https://doi.org/10.14740/wjon1481

Introduction

Despite that the incidences of gastric and esophageal carcinoma are broadly declining over the past decades, they remain the fifth (5.7% of total) and seventh (3.2% of total) most common cancer worldwide, respectively. According to the GLOBOCAN 2018 database, gastric and esophageal carcinoma, in total, accounted for 13.5% of all cancer deaths worldwide [1]. Patients with gastro-esophageal (GE) carcinoma commonly have advanced or metastatic disease at initial diagnosis [2, 3], and the treatment strategy is characterized by the use of cytotoxic regimens. However, although several randomized trials have demonstrated that advanced or metastatic GE carcinoma could benefit from systemic chemotherapy, the prognosis of GE carcinoma patients remains dismal, with a median overall survival (OS) of approximately 12 months [4-7]. Therefore, novel drugs are needed to improve clinical outcomes [8].

Over the past decades, immune checkpoint inhibitors (ICIs) that block the programmed death-1 (PD-1) axis have shown promising therapeutic efficacy in various solid tumors, including GE carcinoma [9-11]. So far, ICIs have shown superior survival over chemotherapy as first and later line treatment in advanced GE carcinoma [12-18]. Nonetheless, approximately 40% of GE carcinoma patients treated with ICIs still suffer from intrinsic or acquired drug resistance, and many immune-oncology (IO) trials are required to further improve their prognoses. To accelerate the approval of effective ICIs, development of surrogate endpoint for OS is an optional but promising strategy. In the era of chemotherapy, the conven-

tional RECIST-based endpoints have been widely applied to reflect the antitumor activity and validated as the robust surrogacy for OS in advanced GE carcinoma trials [19]. However, ICIs have distinct mechanisms of action (e.g., delayed clinical benefit [20], pseudoprogression [21] and hyper-progression [22]). Previous meta-analyses have shown that the conventional RECIST-based endpoints cannot serve as a primary endpoint for OS in pan-cancer IO trials [23, 24]. Nonetheless, significant heterogeneities among different solid tumors limit these applications in the IO trials of advanced or metastatic GE carcinoma.

Therefore, we used arm- and trial-level quantitative approaches to evaluate, for the first time, the correlation between RECIST-based endpoints (including progression-free survival (PFS), objective response rate (ORR) and disease control rate (DCR)) and OS in randomized controlled IO trials of GE carcinoma.

Materials and Methods

Search strategy and study selection

Two authors (RCN and YW) independently searched Medline (PubMed), Web of Science, Embase, ClinicalTrials.gov and Cochrane Library databases for eligible trials from January 1, 2000 to September 30, 2021, using the following search terms: nivolumab, pembrolizumab, avelumab, atezolizumab, durvalumab, PD-1, PD-L1, checkpoint inhibitors, gastro-esophageal carcinoma and randomized controlled trial. Supplementary Material 1 (www.wjon.org) shows the detailed search terms. Randomized controlled trials (RCTs) investigating anti-PD-1/programmed death ligand-1 (PD-L1) therapy in advanced GE carcinoma that reported treatment effect (hazard ratios (HRs)/odds ratios (ORs)) on OS and surrogate endpoints (PFS/ORR/DCR) were included. We excluded reviews, abstracts, case reports and studies with sample size less than 150 subjects. Conference abstracts of the 2021 American Society of Clinical Oncology (ASCO) annual meeting and the European Society for Medical Oncology (ESMO) Congress 2021 were manually searched to retrieve eligible trials.

Data extraction and endpoints

The following data for each eligible trial were extracted: population, study phase, treatment protocol, sample size, primary endpoint, results of OS and surrogate endpoints (PFS, ORR and DCR). For trials reporting on multiple populations, the largest population with reported primary endpoints was included. The survival rates of OS and PFS at different cut-off time points (3, 6, 9, 12, 15, 18 and 24 months) were measured using the Engauge Digitizer tool V.12.1 (http://markummitchell.github.io/engauge-digitizer/). The HRs for OS and PFS at different cut-off time points were calculated using the Kaplan-Meier curves, according to the description by Parmar et al [25]. OS was defined as the time from randomization to death from any cause. PFS was defined as the time from randomization to disease progression or any death. ORR was defined as the proportion of best confirmed complete response (CR) or partial response (PR). DCR was defined as the percentage of best-confirmed CR, PR or stable disease (SD).

Statistical analysis

Our quantitative evaluation used two correlation approaches (arm- and trial-level) to assess the potential surrogate endpoints for OS, as previously described [26, 27]. The strength of association between the surrogate endpoints (median PFS, ORR and DCR) and median OS of each experimental arm (arm treated with ICIs) at the arm-level was assessed. The correlation between HRs for PFS and ORs for ORR/DCR and HRs for OS at the trial-level was assessed via a linear regression model, weighted by trial arm or trial size. The sample size of trials that reported multiple arms was down-weighted based on the descriptions of A'Hern et al [28]. The arm- and trial-level correlations were quantified by weighted Pearson correlation coefficient (R). According to the criteria of the Institute for Quality and Efficiency in Health Care (IQWIG) [29], the strength of association between endpoints was categorized as weak ($R < 0.70$), moderate ($R = 0.70 - 0.85$) and strong ($R > 0.85$), based on the value of R.

For each meta-analysis, we used the leave-one-out cross-validation analysis to assess the prediction accuracy of the surrogate model [30]. Each trial was left out once and the surrogate model was built using the remaining trials. Predicted HRs for OS with 95% prediction intervals were calculated from the observed HR of PFS of that particular trial. To demonstrate typical conditions, the strength of associations between HRs for 3, 6, 12, and 15-month PFS, and HRs for 6, 12, 18, and 24-month OS were calculated, and several subgroup analyses of tumor type, trials line, treatment strategy and follow-up duration were performed. Statistical analyses were performed using the R software, version 4.2.0 (http://www.r-project.org).

Results

After screening 657 reports and conference abstracts, a total of 17 trials were found eligible (Fig. 1) [12-18, 31-41]. Two phase 2 RCTs were excluded because of small sample size [42, 43]. All the eligible studies were phase 3 randomized trials. Table 1 shows the detailed information of the eligible trials. We included the largest primary endpoint population of KEYNOTE-181 (combined positive score (CPS) $\geq$ 10) [15], CheckMate 649 (CPS $\geq$ 5) [16, 36], JAVELIN Gastric 100 (all patients) [37], KEYNOTE-062 (CPS $\geq$ 1) [35], CheckMate 648 (CPS $\geq$ 1) [31], ORIENT-15 (all patients) [39] and ORIENT-16 (all patients) [40]. CheckMate 649 was comprised of 2,031 patients, of whom 1,581 were randomly assigned (1:1) to nivolumab plus chemotherapy or chemotherapy, and 813 were randomly assigned (1:1) to nivolumab plus ipilimumab or chemotherapy. The former cohort was published in 2021 [16], and the latter was reported in the ESMO congress 2021 [36]; thus, two comparisons were included in our analysis. Likewise, two com-

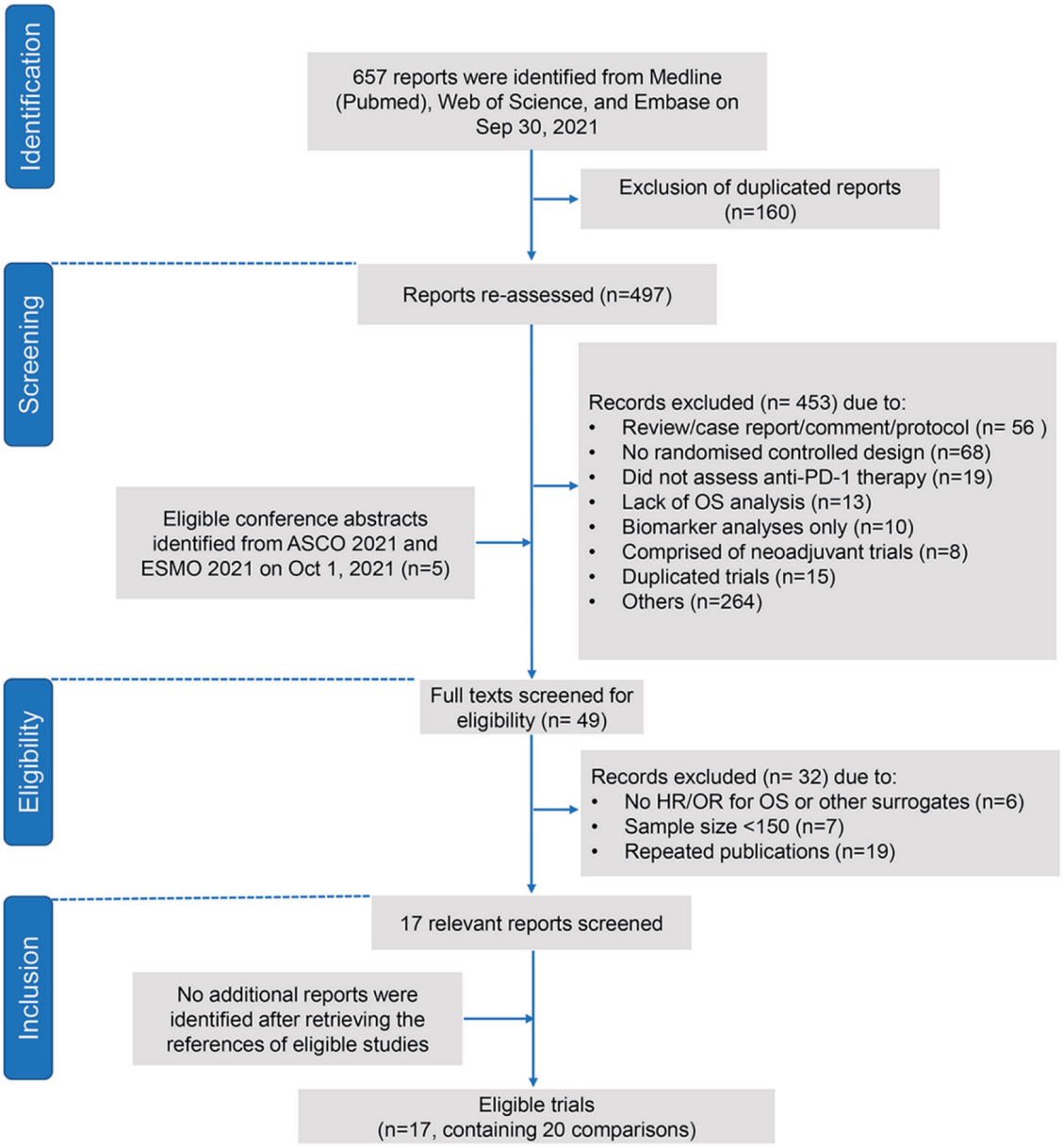

Figure 1. Study flow chart. PD-1: programmed death-1; ASCO: American Society of Clinical Oncology; ESMO: European Society for Medical Oncology; OR: odds ratio; HR: hazard ratio; OS: overall survival.

parisons of the KEYNOTE-062 [35] (pembrolizumab versus chemotherapy, and pembrolizumab plus chemotherapy versus chemotherapy) and CheckMate 648 [31] (nivolumab plus chemotherapy versus chemotherapy, and nivolumab plus ipilimumab versus chemotherapy) were included in our analysis. Overall, the 17 eligible trials yielded 20 treatment comparisons with a total of 9,657 subjects.

First, a total of 20 available arms were included to derive the arm-level correlations between potential endpoints and OS. ORR and DCR showed strong and moderate correlations with median OS ($R = 0.91$, $P < 0.001$, Fig. S1A; $R = 0.80$, $P < 0.001$, Fig. S1B) (Supplementary Material 2, www.wjon.org). Similarly, median PFS was strongly correlated with median OS ($R = 0.87$, $P < 0.001$, Fig. 2a).

We then derived the degree of association between treatment effect on potential endpoints and OS at trial-level. Since none of the 131 patients in the placebo group had an objective response in the ATTRACTION-2 trial [12], the OR for ORR (infinite) in the ATTRACTION-2 trial was not available. Eighteen comparisons of ORs for ORR and HRs for OS were available, among which 10 reported improvements in both ORR (lower limit of CI for OR > 1.0) and OS (upper limit of CI for HR < 1.0). Correlation between OR_{ORR} and HR_{OS} was moderate ($R = 0.71$, $P < 0.001$, Fig. S1C) (Supplementary Material 2, www.wjon.org). Including the ATTRACTION-2 trial, the correlation between OR_{ORR} and HR_{OS} was not significant.

Table 1. Characteristics of the Included Trials

Studies	Population	Line of treatments	Study phase	Primary endpoint	Treatment arms	N	DCR (%)	ORR (%)	OS		PFS	
									Median	HR	Median	HR
ATTRACTION-2 [12]	G/EGJ carcinoma	≥ 3	3	OS	Nivo; placebo	330; 163	40; 25	11; 0	5.26; 4.14	0.63 (0.51 - 0.78)	1.61; 1.45	0.60 (0.49 - 0.75)
JAVELIN Gastric 300 [32]	G/EGJ carcinoma	≥ 2	3	OS	Avelumab; chemotherapy	185; 186	22.2; 44.1	2.2; 4.3	4.6; 5.0	1.1 (0.90 - 1.40)	1.4; 2.7	1.73 (1.40 - 2.20)
KEYNOTE-061 [33]	G/EGJ carcinoma	2	3	OS; PFS	Pembrolizumab; paclitaxel	196; 199	59.5; 29.5	16; 14	9.1; 8.3	0.82 (0.66 - 1.03)	1.5; 4.1	1.27 (1.03 - 1.57)
ATTRACTION-3 [13]	ESCA	2	3	OS	Nivo; chemotherapy	210; 209	37; 66	19; 22	10.9; 8.4	0.77 (0.62 - 0.96)	1.7; 3.4	1.08 (0.87 - 1.34)
ATTRACTION-4 [34]	G/EGJ carcinoma	1	3	OS; PFS	Nivo + chemotherapy; placebo + chemotherapy	362; 362	71.8; 68.5	57.5; 47.8	17.45; 17.15	0.90 (0.75 - 1.08)	10.45; 8.34	0.68 (0.51 - 0.90)
ESCORT [14]	ESCA	2	3	OS	Camrelizumab; chemotherapy	228; 220	44.7; 34.5	20.2; 6.4	8.3; 6.2	0.71 (0.57 - 0.87)	1.9; 1.9	0.69 (0.56 - 0.86)
KEYNOTE-590 [18]	ESCA	1	3	OS; PFS	Pembrolizumab + chemotherapy; placebo + chemotherapy	373; 376	79.3; 75.6	45; 29.3	12.4; 9.8	0.73 (0.62 - 0.86)	6.3; 5.8	0.65 (0.55 - 0.76)
KEYNOTE-181 [15]*	ESCA	2	3	OS (CPS ≥ 10)	Pembrolizumab; chemotherapy	107; 115	49.5; 47.0	21.5; 6.1	9.3; 6.7	0.69 (0.52 - 0.93)	2.6; 3.0	0.73 (0.54-0.97)
CheckMate 649 [16, 36]**	G/EGJ carcinoma	1	3	OS; PFS (CPS ≥ 5)	Nivo + chemotherapy; chemotherapy; Nivo + ipi; chemotherapy	473; 482; 234; 239	88; 79; 54; 83	60; 45; 27; 47	14.4; 11.1; 11.2; 11.6	0.71 (0.59 - 0.86); 0.89 (0.71 - 1.10)	7.7; 6.0; 2.8; 6.3	0.68 (0.56 - 0.81); 1.42 (1.14 - 1.76)
JAVELIN Gastric 100 [37]***	G/EGJ carcinoma	1	3	OS (all patients or TPS ≥ 1)	Avelumab; chemotherapy	249; 250	50.2; 61.2	13.3; 14.4	10.4; 10.9	0.91 (0.74 - 1.11)	3.2; 4.4	1.04 (0.85 - 1.28)
KEYNOTE-062 [35]****	Gastric carcinoma	1	3	OS; PFS (CPS ≥ 1 or ≥ 10)	Pembrolizumab + chemotherapy; pembrolizumab; chemotherapy	257; 256; 250	78; 42; 79	49; 15; 37	13.9; 10.6; 11.1	0.85 (0.70 - 1.03); 0.91 (0.74 - 1.10)	6.9; 2.0; 6.4	0.84 (0.70 - 1.02); 1.66 (1.37 - 1.51)
RATIONALE 302 [38]	ESCA	2	3	OS	Tislelizumab; chemotherapy	256; 256	47; 41.8	20.4; 9.8	8.6; 6.3	0.70 (0.57 - 0.85)	1.6; 2.1	0.83 (0.67 - 1.01)
CheckMate 648 [31]*****	ESCA	1	3	OS; PFS (CPS ≥ 1)	Nivo + chemotherapy; Nivo + ipi; chemotherapy	158; 158; 157	78; 63; 66	53; 35; 20	15.4; 13.7; 9.1	0.54 (0.37 - 0.80); 0.64 (0.46 - 0.90)	6.9; 4.0; 4.4	0.65 (0.46 - 0.92); 1.02 (0.73 - 1.43)
ESCORT-1st [17]	ESCA	1	3	OS; PFS	Camrelizumab + chemotherapy; chemotherapy	298; 298	91.3; 88.9	72.1; 62.1	15.3; 12.0	0.70 (0.56 - 0.88)	6.9; 5.6	0.56 (0.46 - 0.68)
ORIENT-15 [39]***	ESCA	1	3	OS (CPS ≥ 10 and all patients)	Sintilimab + chemotherapy; chemotherapy	327; 332	NR	66.1; 45.5	16.7; 12.5	0.63 (0.51 - 0.78)	7.2; 5.7	0.56 (0.46 - 0.68)
ORIENT-16 [40]***	G/EGJ carcinoma	1	3	OS (CPS ≥ 5 and all patients)	Sintilimab + chemotherapy; chemotherapy	327; 323	NR	58.2; 48.4	15.2; 12.3	0.76 (0.63 - 0.94)	7.1; 5.7	0.64 (0.53 - 0.77)
JUPITER-06 [41]	ESCA	1	3	OS; PFS	Toripalimab + chemotherapy; chemotherapy	257; 257	NR	NR	17.0; 11.0	0.58 (0.43 - 0.78)	5.7; 5.5	0.58 (0.46 - 0.74)

*Patients with CPS ≥ 10 were included. **Patients with CPS ≥ 5 were included. ***All randomly assigned patients were included. ****Patients with CPS ≥ 1 were included. *****Patients with CPS ≥ 1 were included. G: gastric; EGJ: esophagogastric junction; ESCA: esophageal carcinoma; Nivo: nivolumab; Ipi: ipilimumab; NIVO3: nivolumab 3 mg/kg; NIVO1 + IPI3: nivolumab 1 mg/kg plus ipilimumab 3 mg/kg; NIVO3 + IPI1: NIVO3 plus ipilimumab 1 mg/kg; NR: not reported; CPS: combined positive score; TPS: tumor positive score; ORR: objective response rate; DCR: disease control rate; OS: overall survival; PFS: progression-free survival; HR: hazard ratio.

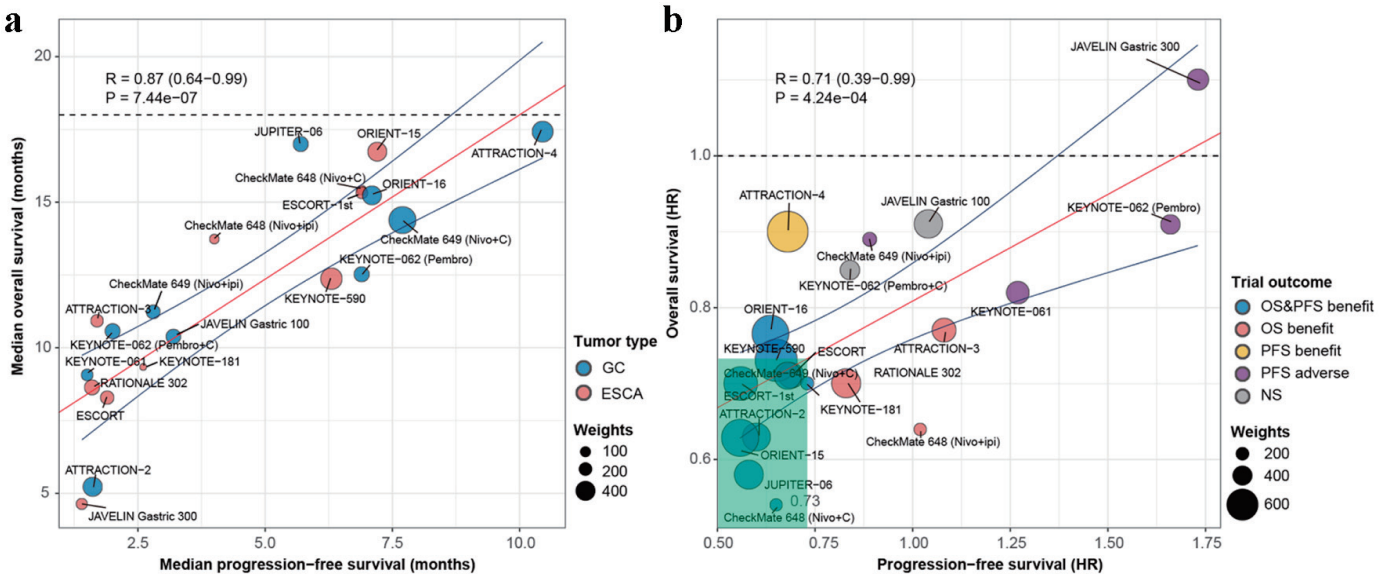

Figure 2. Performance of PFS as surrogate endpoint for OS in immuno-oncology trials of advanced gastro-esophageal carcinoma. (a) Correlation between PFS and OS at arm-level. Each dot represents one of the experimental arms of the phase 3 clinical trials, with size of the dot being proportional to the sample size. (b) Correlation between HRs for PFS and OS at trial-level. Size of dots is proportional to weighted sample size. The blue line represents the upper and lower 95% confidence intervals of the regression line (red line). Trials are colored based on whether the endpoint results were statistically significant. Nivo: nivolumab; Pembro: pembrolizumab; Ipi: ipilimumab; C: chemotherapy; GC: gastric carcinoma; ESCA: esophageal carcinoma; HR: hazard ratio; OS: overall survival; PFS: progression-free survival; NS: not significant; R: weighted Pearson correlation coefficient.

Sixteen pairs of ORs for DCR and HRs for OS were available, and the correlation between OR_{DCR} and HR_{OS} was weak (R = 0.45, P = 0.069, Fig. S1D) (Supplementary Material 2, www.wjon.org). Twenty pairs of HRs for PFS and OS were available. Apart from the comparison of the ATTRACTION-4 trial, other 10 comparisons that showed improvement in PFS reported improvement in OS (Table 1, Fig. 2b). Correlation between HR_{PFS} and HR_{OS} was moderate (R = 0.71, P < 0.001, Fig. 2b). A conservative minimum threshold effect of HR_{PFS} less than 0.73 demonstrated the potential to predict a significant improvement in OS.

Further, leave-one-out cross-validation analyses were performed to evaluate the accuracy of PFS in predicting OS. It was noted that the observed HR for OS fell within the 95% prediction intervals in 19 of 20 comparisons, indicating that the treatment effect on PFS could be a potential predictor of OS (Fig. 3).

Figure 4 shows the strength of association between PFS and OS at different cut-off time points. The arm- (Fig. 4a) and trial-level (Fig. 4b) correlations showed that 3-month PFS were strongly correlated with 6-month OS (R = 0.92, R = 0.90), 6-month PFS strongly correlated with 12-month OS (R = 0.88, R = 0.94), and 12-month PFS strongly correlated with 18-month OS (R = 0.86, R = 0.86). The strength of association was weakened as the OS increased.

Finally, subgroup analyses were performed to evaluate the correlation between treatment effect on PFS and OS in different tumor types, trial lines, treatment strategy and follow-up duration (Table 2). The strength of association between HR_{PFS} and HR_{OS} remained moderate in gastric or GE junction cancer (R = 0.71), but weak in esophageal cancer (R = 0.47). Notably, the correlation between HR_{PFS} and HR_{OS} became strong in trials of ≥ 2 lines (R = 0.96), monotherapy (R = 0.89) and shorter follow-up duration (R = 0.91).

Discussion

This is the first study to comprehensively evaluate the candidate surrogate endpoints for OS in IO trials of advanced or metastatic GE carcinoma. In the present study, we found that RECIST-based DCR could not serve as appropriate surrogate endpoint for OS. However, RECIST-based ORR and PFS correlated strongly with OS at arm-level and moderately with OS at trial-level. The leave-one-out cross-validation approach also confirmed that the effects observed on PFS were adequate to predict the treatment effect on OS. Therefore, we proposed the use of PFS as potential surrogate endpoint for OS in IO trials of advanced or metastatic GE carcinoma.

Recently, the KEYNOTE-590 [18], and ESCORT-1st [17] trials demonstrated that a combination of ICIs with chemotherapy was more effective than chemotherapy alone in previously untreated esophageal carcinoma. Furthermore, early reports from the CheckMate 648 trial in ASCO 2021 [31] suggest that chemo-free regimen (nivolumab plus ipilimumab) could represent a novel standard first-line treatment for esophageal carcinoma. Despite the unsuccessful exploration of pembrolizumab in second [33] and first-line [35] in gastric cancer, the CheckMate 649 trial [16] showed that nivolumab plus chemotherapy improved survival compared with chemotherapy

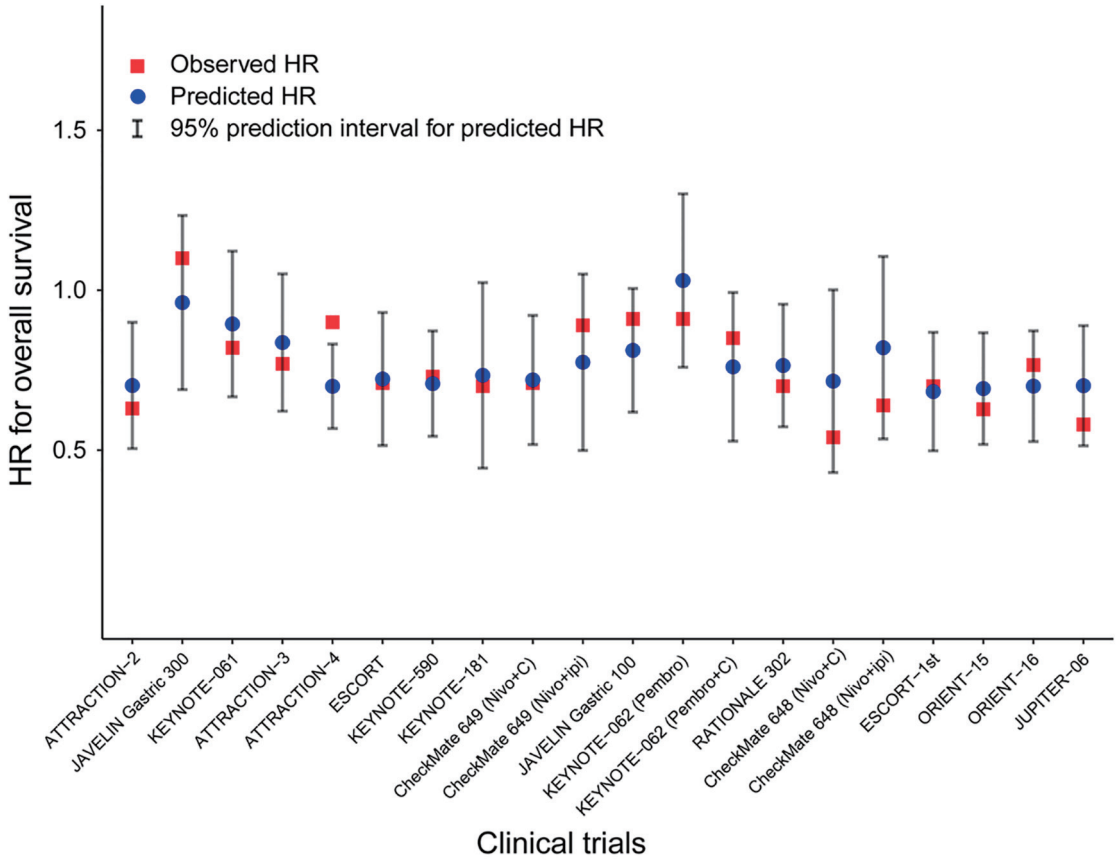

Figure 3. Leave-one-out cross-validation analysis of the prediction of OS by treatment effect on PFS. Predicted HRs for OS (blue circles) with 95% prediction intervals (vertical grey lines) were calculated from the observed HR on PFS of that particular trial and the surrogate model built on the remaining trials. Observed HRs are shown for OS (red squares). Nivo: nivolumab; Pembro: pembrolizumab; Ipi: ipilimumab; C: chemotherapy; HR: hazard ratio; OS: overall survival; PFS: progression-free survival.

alone. Therefore, the emerging of anti-PD-1/PD-L1 agents has unprecedentedly changed the treatment landscape of advanced GE carcinoma. However, not all patients have clinical response to ICIs, and several critical issues are required to be clarified, namely, identification of responders before the initial use of ICIs, and improvement of the therapeutic effect of ICIs through effective combination modality. Consequently, several randomized trials were in process to investigate the therapeutic effect of combinational regimens, such as a combination of ICIs with chemotherapy (KEYNOTE-859; RATIONAL-305; NCT03958890), anti-angiogenic (NCT03813784; NCT04949256), and targeted agents (KEYNOTE-811).

It is well recognized that OS is the golden standard primary endpoint for clinical trials of solid tumors. To reduce the sample size, shorten the follow-up duration and accelerate the approval of effective regimens, identification of surrogate endpoint for OS is an optional but important surrogate. Several clinical trials had set PFS (NCT03958890) as the unique primary endpoint or PFS and OS [16, 18, 33, 34] as the dual primary endpoints. Indeed, in the era of chemotherapy, RECIST-based endpoints had been commonly used as surrogate endpoints for OS in GE carcinoma; however, the use of these endpoints for OS in IO trials remains debatable because of the distinct anti-tumor mechanism of ICIs [20, 44], such as low-quality progression and delayed response [21]. Two previous meta-analyses showed that weak correlations did not support the surrogacy of RECIST-based endpoints for OS in pan-cancer advanced IO trials [23, 24]. Despite this, heterogeneity is pervasive and enormous across various cancer types [45], and the response patterns of cancer types treated with ICIs are diverse. Thus, the correlations in pan-cancer advanced IO trials cannot extrapolate to trials of particular cancer type [46]. Therefore, exploration of surrogate endpoints for OS in IO trials of GE carcinoma is still important.

In the present study, we applied rigid criteria and included a total of 17 large phase 3 trials with 9,657 patients to solve this issue. Firstly, we found that DCR and ORR did not strongly correlate with OS at both arm- and trial-level. We considered that not only the evaluation of targeted lesions, but also the follow-up duration is critical. In addition, the DCR and ORR at extreme condition (e.g., 0% and 100%) could not effectively predict outcome of OS. We found that the correlations between PFS and OS at arm- and trial-level were strong and moderate, respectively. The leave-one-out cross-validation analysis further confirmed the potential surrogacy of PFS for OS. Our study indicated a conservative minimum threshold effect of

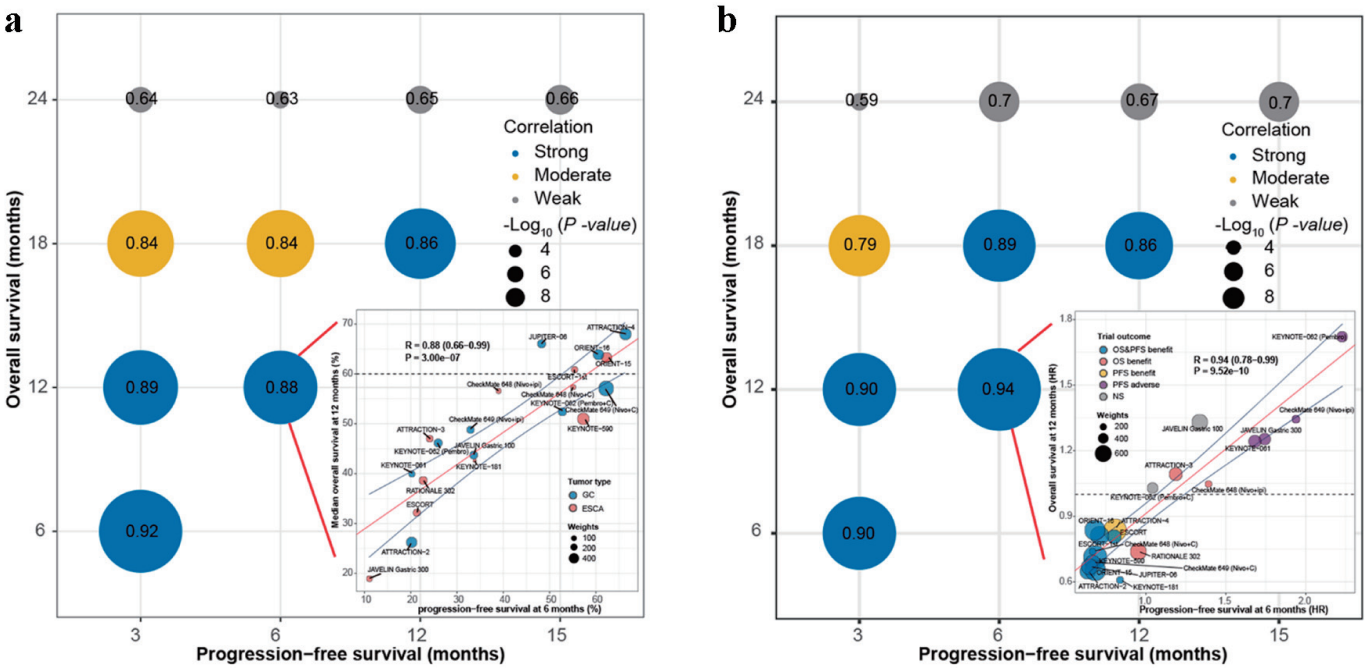

Figure 4. Correlation between PFS and OS at different cut-off time points. (a) Correlation between PFS and OS at arm-level. Bottom right: PFS at 6 months to predict OS at 12 months. (b) Correlation between HRs for PFS and OS at trial-level. Bottom right: HRs for PFS at 6 months to predict HRs for OS at 12 months. Nivo: nivolumab; Pembro: pembrolizumab; Ipi: ipilimumab; C: chemotherapy; GC: gastric carcinoma; ESCA: esophageal carcinoma; HR: hazard ratio; OS: overall survival; PFS: progression-free survival; NS: not significant; R: weighted Pearson correlation coefficient.

$HR_{PFS} \leq 0.73$ to highly predict a significant improvement in OS. It is believed that the acceptable correlation between PFS and OS in IO trials of GE carcinoma is largely ascribed to the condition that limited subsequent lines of therapy if patients with advanced GE carcinoma progressed after treating with ICIs. However, we should note that the heterogeneity is still obvious, including the heterogeneity of multiple cancer types (gastric cancer, gastroesophageal junction adenocarcinoma and esophageal squamous cell carcinoma) and line treatment. Therefore, our study should be interpretated cautiously.

In future IO trial, interest could be focused on predicting the treatment effects on OS by observing the effects on PFS at earlier time points. Kok et al reported that 6-month PFS could effectively predict 12-month OS in IO trials [47]. Similarly, our study found that 3-month PFS could reliably predict 6-month OS, 6-month PFS could reliably predict 12-month

Table 2. Subgroup Analysis of the Correlation Between PFS and OS as Trial Level

Subgroup analysis	No. of comparisons	Weighted correlation coefficients, R (95% CI)	P value
Tumor type			
ESCA [13-15, 17, 18, 31, 38, 39, 41]	10	0.47 (0.00 - 0.99)	0.174
G/EGJ cancer [12, 16, 32-37, 40]	10	0.71 (0.22 - 0.99)	0.021
Trials line			
First-line [16-18, 31, 34-37, 39-41]	13	0.57 (0.08 - 0.99)	0.043
≥ 2 lines [12-15, 32, 33, 38]	7	0.96 (0.72 - 0.99)	< 0.001
Treatment strategy			
Monotherapy [12-15, 32, 33, 35, 37, 38]	9	0.89 (0.56 - 0.99)	0.001
Combinational therapy [16-18, 31, 34-36, 39-41]	11	0.41 (0.00 - 0.99)	0.215
Median follow-up			
≥ 10 months [13, 16-18, 31, 32, 34-37, 39, 40]	14	0.71 (0.31 - 0.99)	0.005
< 10 months [12, 14, 15, 33, 38, 41]	6	0.91 (0.51 - 0.99)	0.011

G: gastric; EGJ: esophagogastric junction; ESCA: esophageal carcinoma; CI: confidence interval.

OS, and 12-month PFS could reliably predict 18-month OS in IO trials of advanced GE carcinoma. However, we noted weakened correlations between HR_{PFS} and HR_{OS} as the follow-up duration increased. We considered that this phenomenon could be mainly attributed to the disproportionate increase of HR_{PFS} and HR_{OS} because of delayed responses in the experimental arms.

Our study had several limitations. First, despite that the treatment modalities of gastric and esophageal carcinoma are similar, potential heterogeneity in terms of tumor type should be noted in our study. The combination of first line, later line and different treatment modalities also contributed to certain level of heterogeneity within eligible trials. Although we performed subgroup analyses to reduce these biases, the small number of comparisons (range: 6 - 13) in each analysis indicated a low power for statistical analysis. In addition, several endpoints modified based on RECIST criteria may better reflect the response pattern of ICIs, such as irRC [48], irRECIST [49] and iRECIST [50] criteria. However, the included trials of our studies had not reported these endpoints; thus, we could not explore the surrogacy of these endpoints for OS in IO trials of GE carcinoma. Lastly, our analysis was performed at arm- and trial-levels, and lacked patients-level analysis.

Conclusions

RECIST-based PFS may be the appropriate surrogate for predicting OS in IO trials of GE carcinoma. A conservative minimum threshold effect of HR_{PFS} less than 0.73 has the potential to predict a significant improvement in OS.

Supplementary Material

Suppl 1. PubMed search terms.

Suppl 2. Performance of ORR and DCR as surrogate endpoint for OS in immuno-oncology trials of advanced gastro-esophageal carcinoma. Correlation between ORR (A) and DCR (B) and OS at arm-level. Each dot represents one of the experimental arms of the phase 3 clinical trials, with size of the dot being proportional to the sample size. Correlation treatment effects on ORR (C) and DCR (D) and OS at trial-level. Size of dots is proportional to weighted sample size. The blue line represents the upper and lower 95% confidence intervals of the regression line (red line). Trials are colored based on whether the endpoint results were statistically significant. Nivo: nivolumab; Pembro: pembrolizumab; Ipi: ipilimumab; C: chemotherapy; GC: gastric carcinoma; ESCA: esophageal carcinoma; OR: odds ratio; HR: hazard ratio; ORR: objective response rate; DCR: disease control rate; OS: overall survival; NS: not significant; R: weighted Pearson correlation coefficient.

Acknowledgments

None to declare.

Financial Disclosure

This work was supported by the grants of the National Science Foundation of China (82103586), China Postdoctoral Science Foundation (2021M703728) and Guangdong Provincial Natural Science Foundation (2021A1515012369).

Conflict of Interest

The authors declare that they have no conflict of interest.

Informed Consent

Not applicable.

Author Contributions

YFL and YW searched and analyzed the data and contributed to drafting the typescript. RCN guided the statistical analyses. YW, JZ, YXY, GMC, FYZ, YCW, SC, and ZWZ contributed to collecting the data. YW and JL prepared the figures and tables. RCN and YBC edited and revised the typescript. RCN designed the study and takes responsibility for the integrity of the work.

Data Availability

All data generated or analyzed during this study are included in this published article.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
2. Pennathur A, Gibson MK, Jobe BA, Luketich JD. Oesophageal carcinoma. Lancet. 2013;381(9864):400-412.
3. Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. Lancet. 2020;396(10251):635-648.
4. Cunningham D, Starling N, Rao S, Iveson T, Nicolson M, Coxon F, Middleton G, et al. Capecitabine and oxaliplatin for advanced esophagogastric cancer. N Engl J Med. 2008;358(1):36-46.
5. Koizumi W, Narahara H, Hara T, Takagane A, Akiya T, Takagi M, Miyashita K, 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(3):215-221.
6. Kang YK, Chin K, Chung HC, Kadowaki S, Oh SC, Nakayama N, Lee KW, et al. S-1 plus leucovorin and ox-

aliplatin versus S-1 plus cisplatin as first-line therapy in patients with advanced gastric cancer (SOLAR): a randomised, open-label, phase 3 trial. Lancet Oncol. 2020;21(8):1045-1056.
7. Muro K, Lordick F, Tsushima T, Pentheroudakis G, Baba E, Lu Z, Cho BC, et al. Pan-Asian adapted ESMO Clinical Practice Guidelines for the management of patients with metastatic oesophageal cancer: a JSMO-ESMO initiative endorsed by CSCO, KSMO, MOS, SSO and TOS. Ann Oncol. 2019;30(1):34-43.
8. Peng Z, Liu T, Wei J, Wang A, He Y, Yang L, Zhang X, et al. Efficacy and safety of a novel anti-HER2 therapeutic antibody RC48 in patients with HER2-overexpressing, locally advanced or metastatic gastric or gastroesophageal junction cancer: a single-arm phase II study. Cancer Commun (Lond). 2021;41(11):1173-1182.
9. Ribas A, Wolchok JD. Cancer immunotherapy using checkpoint blockade. Science. 2018;359(6382):1350-1355.
10. Sanmamed MF, Chen L. A paradigm shift in cancer immunotherapy: from enhancement to normalization. Cell. 2018;175(2):313-326.
11. Wang FH, Zhang XT, Li YF, Tang L, Qu XJ, Ying JE, Zhang J, et al. The Chinese Society of Clinical Oncology (CSCO): Clinical guidelines for the diagnosis and treatment of gastric cancer, 2021. Cancer Commun (Lond). 2021;41(8):747-795.
12. Kang YK, Boku N, Satoh T, Ryu MH, Chao Y, Kato K, Chung HC, et al. Nivolumab in patients with advanced gastric or gastro-oesophageal junction cancer refractory to, or intolerant of, at least two previous chemotherapy regimens (ONO-4538-12, ATTRACTION-2): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet. 2017;390(10111):2461-2471.
13. Kato K, Cho BC, Takahashi M, Okada M, Lin CY, Chin K, Kadowaki S, et al. Nivolumab versus chemotherapy in patients with advanced oesophageal squamous cell carcinoma refractory or intolerant to previous chemotherapy (ATTRACTION-3): a multicentre, randomised, open-label, phase 3 trial. Lancet Oncol. 2019;20(11):1506-1517.
14. Huang J, Xu J, Chen Y, Zhuang W, Zhang Y, Chen Z, Chen J, et al. Camrelizumab versus investigator's choice of chemotherapy as second-line therapy for advanced or metastatic oesophageal squamous cell carcinoma (ESCORT): a multicentre, randomised, open-label, phase 3 study. Lancet Oncol. 2020;21(6):832-842.
15. Kojima T, Shah MA, Muro K, Francois E, Adenis A, Hsu CH, Doi T, et al. Randomized phase III KEYNOTE-181 study of pembrolizumab versus chemotherapy in advanced esophageal cancer. J Clin Oncol. 2020;38(35):4138-4148.
16. Janjigian YY, Shitara K, Moehler M, Garrido M, Salman P, Shen L, Wyrwicz L, et al. First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): a randomised, open-label, phase 3 trial. Lancet. 2021;398(10294):27-40.
17. Luo H, Lu J, Bai Y, Mao T, Wang J, Fan Q, Zhang Y, et al. Effect of camrelizumab vs placebo added to chemotherapy on survival and progression-free survival in patients with advanced or metastatic esophageal squamous cell carcinoma: the ESCORT-1st randomized clinical trial. JAMA. 2021;326(10):916-925.
18. Sun JM, Shen L, Shah MA, Enzinger P, Adenis A, Doi T, Kojima T, et al. Pembrolizumab plus chemotherapy versus chemotherapy alone for first-line treatment of advanced oesophageal cancer (KEYNOTE-590): a randomised, placebo-controlled, phase 3 study. Lancet. 2021;398(10302):759-771.
19. Paoletti X, Oba K, Bang YJ, Bleiberg H, Boku N, Bouche O, Catalano P, et al. Progression-free survival as a surrogate for overall survival in advanced/recurrent gastric cancer trials: a meta-analysis. J Natl Cancer Inst. 2013;105(21):1667-1670.
20. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. Nature. 2017;541(7637):321-330.
21. Beaver JA, Hazarika M, Mulkey F, Mushti S, Chen H, He K, Sridhara R, et al. Patients with melanoma treated with an anti-PD-1 antibody beyond RECIST progression: a US Food and Drug Administration pooled analysis. Lancet Oncol. 2018;19(2):229-239.
22. Ferrara R, Mezquita L, Texier M, Lahmar J, Audigier-Valette C, Tessonnier L, Mazieres J, et al. Hyperprogressive disease in patients with advanced non-small cell lung cancer treated with PD-1/PD-L1 inhibitors or with single-agent chemotherapy. JAMA Oncol. 2018;4(11):1543-1552.
23. Mushti SL, Mulkey F, Sridhara R. Evaluation of overall response rate and progression-free survival as potential surrogate endpoints for overall survival in immunotherapy trials. Clin Cancer Res. 2018;24(10):2268-2275.
24. Nie RC, Chen FP, Yuan SQ, Luo YS, Chen S, Chen YM, Chen XJ, et al. Evaluation of objective response, disease control and progression-free survival as surrogate endpoints for overall survival in anti-programmed death-1 and anti-programmed death ligand 1 trials. Eur J Cancer. 2019;106:1-11.
25. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med. 1998;17(24):2815-2834.
26. Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analyses of randomized experiments. Biostatistics. 2000;1(1):49-67.
27. Buyse M, Burzykowski T, Michiels S, Carroll K. Individual- and trial-level surrogacy in colorectal cancer. Stat Methods Med Res. 2008;17(5):467-475.
28. A'Hern RP, Ebbs SR, Baum MB. Does chemotherapy improve survival in advanced breast cancer? A statistical overview. Br J Cancer. 1988;57(6):615-618.
29. Validity of surrogate endpoints in oncology Executive summary of rapid report A10-05, Version 1.1. In: Institute for Quality and Efficiency in Health Care: Executive Summaries. Cologne, Germany, 2005.
30. Julious SA, Campbell MJ, Walters SJ. Predicting where future means will lie based on the results of the current

trial. Contemp Clin Trials. 2007;28(4):352-357.
31. Ajani JA, Kato K, Doki Y et al. CheckMate 648: A randomized phase 3 study of nivolumab plus ipilimumab or nivolumab combined with fluorouracil plus cisplatin versus fluorouracil plus cisplatin in patients with unresectable advanced, recurrent, or metastatic previously untreated esophageal squamous cell carcinoma. Journal of Clinical Oncology. 2018;36:TPS193-TPS193.
32. Bang YJ, Ruiz EY, Van Cutsem E, Lee KW, Wyrwicz L, Schenker M, Alsina M, et al. Phase III, randomised trial of avelumab versus physician's choice of chemotherapy as third-line treatment of patients with advanced gastric or gastro-oesophageal junction cancer: primary analysis of JAVELIN Gastric 300. Ann Oncol. 2018;29(10):2052-2060.
33. Shitara K, Ozguroglu M, Bang YJ, Di Bartolomeo M, Mandala M, Ryu MH, Fornaro L, et al. Pembrolizumab versus paclitaxel for previously treated, advanced gastric or gastro-oesophageal junction cancer (KEYNOTE-061): a randomised, open-label, controlled, phase 3 trial. Lancet. 2018;392(10142):123-133.
34. Boku N, Ryu MH, Oh DY, et al. Nivolumab plus chemotherapy versus chemotherapy alone in patients with previously untreated advanced or recurrent gastric/gastroesophageal junction (G/GEJ) cancer: ATTRACTION-4 (ONO-4538-37) study. Annals of Oncology. 2020;31:S1192.
35. Shitara K, Van Cutsem E, Bang YJ, Fuchs C, Wyrwicz L, Lee KW, Kudaba I, et al. Efficacy and safety of pembrolizumab or pembrolizumab plus chemotherapy vs chemotherapy alone for patients with first-line, advanced gastric cancer: the KEYNOTE-062 phase 3 randomized clinical trial. JAMA Oncol. 2020;6(10):1571-1580.
36. Janjigian YY, Ajani JA, Moehler M, et al. Nivolumab (NIVO) plus chemotherapy (Chemo) or ipilimumab (IPI) vs chemo as first-line (1L) treatment for advanced gastric cancer/gastroesophageal junction cancer/esophageal adenocarcinoma (GC/GEJC/EAC): CheckMate 649 study. Annals of Oncology. 2021;32(suppl_5):S1283-S1346.
37. Moehler M, Dvorkin M, Boku N, Ozguroglu M, Ryu MH, Muntean AS, Lonardi S, et al. Phase III trial of avelumab maintenance after first-line induction chemotherapy versus continuation of chemotherapy in patients with gastric cancers: results from JAVELIN Gastric 100. J Clin Oncol. 2021;39(9):966-977.
38. Shen L, Kato K, Kim S-B, et al. RATIONALE 302: Randomized, phase 3 study of tislelizumab versus chemotherapy as second-line treatment for advanced unresectable/metastatic esophageal squamous cell carcinoma. Journal of Clinical Oncology. 2021;39:4012-4012.
39. Shen L, Lu Z, Wang J, et al. Sintilimab plus chemotherapy versus chemotherapy as first-line therapy in patients with advanced or metastatic esophageal squamous cell cancer: First Results of the Phase 3 ORIENT-15 study. Annals of Oncology. 2021;(suppl 5):S1040-S1075.
40. Xu J, Jiang H, Pan Y, et al. Sintilimab plus chemotherapy (chemo) versus chemo as first-line treatment for advanced gastric or gastroesophageal junction (G/GEJ) adenocarcinoma (ORIENT-16): first results of a randomized, double-blind, phase 3 study. Annals of Oncology. 2021;32(suppl_5):S1283-S1346.
41. Xu R, Wang F, Cui C, et al. A randomized, double-blind, phase III study of toripalimab versus placebo in combination with first-line chemotherapy for treatment naive advanced or metastatic esophageal squamous cell carcinoma (ESCC). Annals of Oncology. 2021;32(suppl_5):S1040-S1075.
42. Janjigian YY, Bendell J, Calvo E, Kim JW, Ascierto PA, Sharma P, Ott PA, et al. CheckMate-032 study: efficacy and safety of nivolumab and nivolumab plus ipilimumab in patients with metastatic esophagogastric cancer. J Clin Oncol. 2018;36(28):2836-2844.
43. Stein A, Paschold L, Tintelnot J, et al. Ipilimumab or FOLFOX in combination with nivolumab and trastuzumab in previously untreated HER2 positive locally advanced or metastastic esophagogastric adenocarcinoma (EGA): Results of the randomized phase II INTEGA trial (AIO STO 0217). Annals of Oncology. 2021;32(suppl_5):S1283-S1346.
44. Tang J, Shalabi A, Hubbard-Lucey VM. Comprehensive analysis of the clinical immuno-oncology landscape. Ann Oncol. 2018;29(1):84-91.
45. Dentro SC, Leshchiner I, Haase K, Tarabichi M, Wintersinger J, Deshwar AG, Yu K, et al. Characterizing genetic intra-tumor heterogeneity across 2,658 human cancer genomes. Cell. 2021;184(8):2239-2254.e2239.
46. Nie RC, Yuan SQ, Wang Y, Zou XB, Chen S, Li SM, Duan JL, et al. Surrogate endpoints for overall survival in anti-programmed death-1 and anti-programmed death ligand 1 trials of advanced melanoma. Ther Adv Med Oncol. 2020;12:1758835920929583.
47. Kok PS, Cho D, Yoon WH, Ritchie G, Marschner I, Lord S, Friedlander M, et al. Validation of progression-free survival rate at 6 months and objective response for estimating overall survival in immune checkpoint inhibitor trials: a systematic review and meta-analysis. JAMA Netw Open. 2020;3(9):e2011809.
48. Wolchok JD, Hoos A, O'Day S, Weber JS, Hamid O, Lebbe C, Maio M, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res. 2009;15(23):7412-7420.
49. Nishino M, Giobbie-Hurder A, Gargano M, Suda M, Ramaiya NH, Hodi FS. Developing a common language for tumor response to immunotherapy: immune-related response criteria using unidimensional measurements. Clin Cancer Res. 2013;19(14):3936-3943.
50. Seymour L, Bogaerts J, Perrone A, Ford R, Schwartz LH, Mandrekar S, Lin NU, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. Lancet Oncol. 2017;18(3):e143-e152.