Progression-free survival as a surrogate endpoint for overall survival in modern ovarian cancer trials: a meta-analysis

Katrin M. Sjoquist, Sarah J. Lord, Michael L. Friedlander, Robert John Simes, Ian C. Marschner and Chee Khoon Lee

Abstract

Background: Progression-free survival (PFS) has been adopted as the primary endpoint in many randomized controlled trials, and can be determined much earlier than overall survival (OS). We investigated whether PFS is a good surrogate endpoint for OS in trials of first-line treatment for epithelial ovarian cancer (EOC), and whether this relationship has changed with the introduction of new treatment types.

Methods: In a meta-analysis, we identified summary data [hazard ratio (HR) and median time] from published randomized controlled trials. Linear regression was used to assess the association between treatment effects on PFS and OS overall, and for subgroups defined by treatment type, postprogression survival (PPS) and established prognostic factors.

Results: Correlation between HRs for PFS and OS, in 26 trials with 30 treatment comparisons comprising 24,870 patients, was modest ($r^2 = 0.52$, weighted by trial sample size). The correlation diminished with recency: preplatinum/paclitaxel era, $r^2 = 0.66$; platinum/paclitaxel, $r^2 = 0.44$; triplet combinations, $r^2 = 0.22$; biologicals, $r^2 = 0.30$. The median PPS increased over time for the experimental ($P_{trend} = 0.03$) and control arms ($P_{trend} = 0.003$). The difference in median PPS between treatment arms strongly correlated with the difference in median OS ($r^2 = 0.83$). In trials where the control therapy had median PPS of less than 18 months, correlation between PFS and OS was stronger ($r^2 = 0.64$) than where the median PPS was longer ($r^2 = 0.48$).

Conclusions: In EOC, correlation in the relative treatment effect between PFS and OS in first-line platinum-based chemotherapy randomized controlled trials is moderate and has weakened with increasing availability of effective salvage therapies.

Keywords: chemotherapy, clinical trials, ovarian cancer, therapy

Correspondence to:
Katrin M. Sjoquist
National Health and
Medical Research Council
(NHMRC) Clinical Trials
Centre, Australia New
Zealand Gynaecological
Oncology Group, University
of Sydney, Locked Bag 77,
Camperdown NSW 1450,
Australia
**Katrin.Sjoquist@ctc.usyd.
edu.au**

Sarah J. Lord
NHMRC Clinical Trials
Centre, University of
Sydney, Camperdown,
Australia
School of Medicine,
University of Notre Dame,
Sydney, Australia

Michael L. Friedlander
NHMRC Clinical Trials
Centre, Australia New
Zealand, Gynaecological
Oncology Group, University
of Sydney, Camperdown,
Australia
Department of Medical
Oncology, Prince of Wales
Hospital, Randwick,
Australia

Robert John Simes
NHMRC Clinical Trials
Centre, University of
Sydney, Camperdown,
Australia

Ian C. Marschner
NHMRC Clinical Trials
Centre, University of
Sydney, Camperdown,
Australia
Department of Statistics,
Macquarie University,
Sydney, Australia

Chee Khoon Lee
NHMRC Clinical Trials
Centre, Australia New
Zealand Gynaecological
Oncology Group, University
of Sydney, Camperdown,
Australia

Introduction

Epithelial ovarian cancer (EOC) remains a highly lethal disease, despite improvements in treatment over the last three decades that have increased the median survival but not the proportion of women cured.[1] Most patients with stage III disease relapse within 2 years after debulking surgery and platinum-based chemotherapy, and more than half die within 5 years.[2] There is an urgent need to accelerate development of active new treatments.

Overall survival (OS) has traditionally been regarded as the gold standard primary endpoint for phase III randomized controlled trials evaluating the efficacy of new treatments for EOC.[3,4] Demonstrating an improvement in OS requires trials to be larger, with longer follow up, and hence more cost. Most patients now receive multiple postprogression treatments, including chemotherapy, biological-targeted therapies and surgery, which can significantly confound and dilute the effects of the investigational therapy on

the OS endpoint,[5] and could impede the development of new potentially active therapies. Progression-free survival (PFS) can be determined earlier than OS and has potential both as an independent, valid endpoint and a potential surrogate for OS in certain circumstances. PFS is unaffected by postprogression therapies and may provide earlier evidence of efficacy of new treatments, which can expedite regulatory approval. The consensus of the Gynaecological Cancer InterGroup (GCIG), which includes 29 academic international trials groups, was that while OS remains the gold standard for demonstrating benefit in first-line trials, PFS assessed using validated assessment tools is a valid primary endpoint for phase III trials of first-line therapies for ovarian cancer.[6] Furthermore, the GCIG statement recognizes that differences in OS may be increasingly difficult to demonstrate in first-line trials given the availability of active therapies following progression.

In patients with recurrent ovarian cancer, other goals of treatment, including time to treatment failure, improvement in cancer-related symptoms and delaying time to subsequent therapy, are important in considering the benefit of new therapies, apart from improvement in survival outcomes.

While recognizing that PFS may be a valid endpoint in its own right, it is of interest to consider to what extent improving PFS would be expected to translate to a benefit in OS at a trial level. Furthermore, for future first-line trials in EOC, the value of these survival endpoints for determining the benefit of new therapeutics remains important. Evaluation of the surrogacy relationship between PFS and OS at a trial level will continue to have value in guiding future trial design.

Since previous work evaluating the relationship between PFS and OS in first-line trials of EOC,[7,8] multiple new trials have been conducted with active agents subsequently available in clinical practice. We therefore performed a new literature-based meta-analysis with the primary objective of quantifying the strength of the relationship between the relative treatment effects on PFS and OS in phase III randomized controlled trials of first-line treatments for EOC. We further evaluated, as secondary objectives, the potential impact of the increased availability and number of salvage therapies over time, the duration of PPS, and the impact of known prognostic factors on the relationship between PFS and OS.

Methods

Search strategy

We searched MEDLINE, EMBASE and the Central Registry of Controlled Trials of the Cochrane Library (1 January 1996–30 June 2012) using search terms 'ovarian neoplasms' or 'ovarian cancer/carcinoma', 'chemotherapy' and 'clinical trials' (supplemental file S1). The search strategy was limited to studies in humans and in the English language. Conference proceedings, references of relevant review articles, citations of included studies, and trial cooperative-group websites were hand searched.

Study selection

All randomized phase III trials of first-line therapy in patients with stages IC–IV EOC in which the treatment and intervention arms contained a platinum chemotherapy backbone were eligible for inclusion. Trials that included planned interval debulking were allowed. Trials were required to report relative treatment effects for both PFS and OS. If these data were incomplete, trials were still included if sufficient information could be retrieved from published Kaplan–Meier curves. Trials of maintenance therapies or high-dose chemotherapy with stem cell rescue were ineligible.

Availability of anticancer agents for recurrent EOC over time

The timing of the availability of anticancer treatments for recurrent EOC was recorded as the year of approval by the United States Food and Drug Administration (US FDA) for any clinical indication, as recorded on its website.[9] Data were collected only for treatments with demonstrated activity in EOC[10] that could potentially be used for treatment of recurrent disease.

Data extraction

For each included trial, we extracted the trial name, year of publication or conference presentation, summary statistics of clinicopathologic characteristics (stage, performance status, extent of debulking), type, and median duration of chemotherapy per treatment arm. We also recorded the number of patients who were randomized and who progressed and died, for each treatment arm. We extracted data for hazard ratios (HRs) and 95% confidence intervals (CIs), and median OS
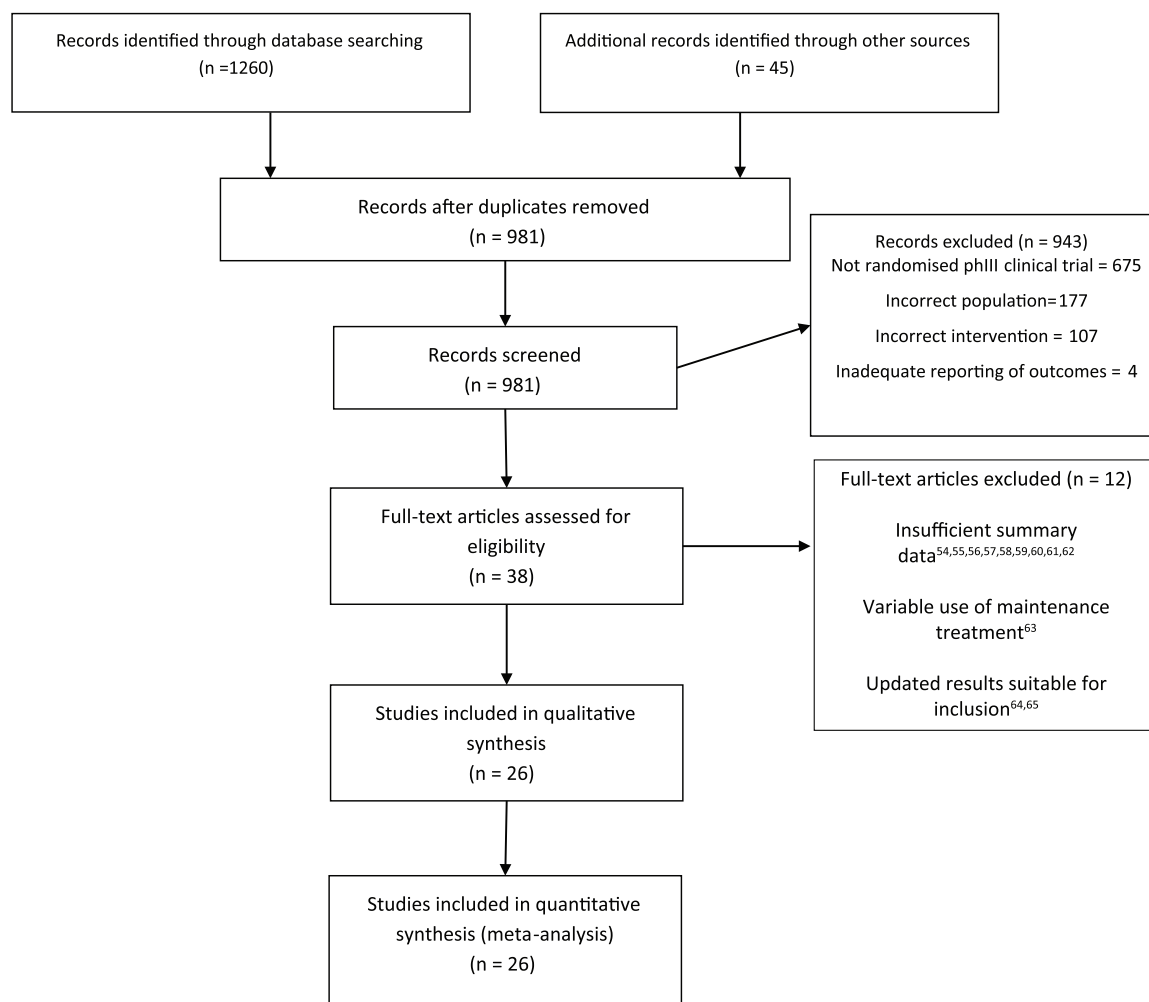
Figure 1. PRISMA diagram/flow chart.
PRISMA, preferred reporting items for systematic reviews and meta-analyses.

and PFS durations. In some trials, where cases of death from causes other than ovarian cancer were censored observations, time to progression was used as the surrogate endpoint instead of PFS. In this review, we considered time to progression and PFS as interchangeable endpoints, given that most patients with advanced ovarian cancer survive beyond the first relapse.

Data on adjusted HRs were used in preference to unadjusted HRs whenever both results were available. In cases where multiple publications of the same trial were available, the results with maximum follow up were used. In trials where there were more than two treatment arms, we obtained the HRs and 95% CIs from the pairwise comparison between the experimental treatments against a common control therapy, and we treated each comparison independently. If HRs and CIs were

not reported, they were estimated using the methods described by Parmar and colleagues.[11]

Data were extracted independently by two authors (KS, SL), and discrepancies were resolved by consensus. Preferred reporting items for systematic reviews and meta-analyses (PRISMA) reporting guidelines were followed for applicable items and the study selection process was summarized in a flow diagram (Figure 1).[12] Publication bias is not a major consideration for this analysis and was not assessed.

Statistical analysis
Because a larger difference in treatment effect for PFS (surrogate endpoint) is assumed biologically to translate into a larger difference in OS (true endpoint), a linear model was fitted by the use of

ordinary least-squares regression. We inspected residual *versus* predicted plots and performed diagnostic tests for normality and heteroscedasticity (nonconstant error variance) to assess consistency with the assumptions of linear regression. All analyses were performed unweighted and then weighted by trial size.

We reported r^2, the trial-level correlation coefficient, between PFS and OS, both unweighted and weighted by trial size, as derived from the regression models. Any r^2 value of 0.72 or greater was considered a strong correlation, and r^2 from 0.49 to less than 0.72 was considered modest correlation.[7] The 95% CIs of r^2 values were obtained by the bootstrap method with 1000 replications.

Subgroup analyses were also carried out for trials that examined different treatment paradigms from different eras: before the use of platinum or paclitaxel as control therapies, when platinum and taxanes were used as control therapies, and in the trials exploring triplet therapies and biological therapies. We also classified these trials into subgroups on the basis of the median distributions according to the proportion of patients with different prognostic characteristics (stage, performance status, and extent of debulking). Sensitivity analyses were performed to evaluate the extent to which the relationship changed with differing proportions of established baseline prognostic factors.

We also tested for the correlation between the difference in median PPS of experimental *versus* control treatment arms and the difference in median OS. The median PPS of a treatment arm was defined as the difference between the median OS and the median PFS. The difference in the median PPS between the treatment arms for trials conducted at different times was examined by classifying trials by the year of the first patient accrual, or if this was not available, the year of the first trial publication. Differences in associations between the HR for PFS and the HR for OS were also evaluated for PPS at the cutoff point of 18 months for the control therapy. This cutoff point was chosen on the basis of a prior study of simulated data,[13] which reported a strong correlation for PPS less than 18 months and a moderate to weak correlation for PPS of 18 months or longer.

We performed sensitivity analyses to examine the impact on the overall results of excluding trials of: (1) intraperitoneal treatment, given that participants in these trials were likely to have complete surgical debulking and hence an overall better prognosis; and (2) biological therapies, as many trials in other advanced cancers had shown a significant relative PFS advantage but no OS difference.

Analyses used STATA, version 14 (StataCorp: College Station, TX, USA)

Results

In total, 26 trials with 30 treatment comparisons and comprising 24,870 patients were included (Figure 1 and Table 1). Most of the patients in these studies had advanced EOC (median of rates for all treatment arms 72.5% stage III, 17% stage IV). Overall, two studies[14,15] contained multiple comparisons among different experimental therapies and a common control arm. There were two trials[15,16] of biological therapies and another two trials[17,18] of intraperitoneal therapy. In total, seven comparisons reported an improvement in PFS (upper limit of the 95% CI for HR <1.00 or reported $p < 0.05$) and four comparisons reported an improvement in OS (Table 1). In five trials, at least one HR was not reported and had to be calculated.[19–23] One trial used time to tumour progression in place of PFS.[19]

Figure 2 is a plot of the HR for PFS *versus* the HR for OS. The notable outlier was a trial comparing cisplatin-paclitaxel with cisplatin-cyclophosphamide, the first to compare two platinum combinations and to include a platinum-taxane combination. Both PFS and OS were significantly better in the experimental arm.[24] Another outlier trial compared cisplatin-paclitaxel with carboplatin-paclitaxel, and reported a nonsignificant difference between the treatment arms for both PFS and OS.[28] When all trials were included the correlation between HRs for PFS and OS was moderate (unweighted r^2, 0.53, 95% CI 0.23–0.72; r^2 weighted by sample size, 0.52, 95% CI 0.30–0.67).

Data on PPS available from 22 treatment comparisons showed a trend to an increasing median PPS over time for both the experimental ($P_{trend} = 0.03$) and control arms ($P_{trend} = 0.003$) [Figure 3(a)]. The difference in median PPS between treatment arms strongly correlated with the difference in median OS [unweighted r^2, 0.75; 95% CI 0.36–0.92; r^2 weighted by sample size, 0.83, 95% CI 0.58–0.92; Figure 3(b)]. Details of post progression therapy were reported for five trials.[24,26,27,29,34,40]

Table 1. 26 trials and 30 comparisons included in the analysis (including biologics).

Trial	Year[a]	Treatment	n (per arm)	HR for PFS (95% CI)	HR for OS (95% CI)	Median PFS (months)	Median OS (months)
GOG 111[24]	1996	cis, cyclo	202	0.7[b]	0.6[b]	13.0	24.0
		cis, tax	184	(0.5–0.8)	(0.5–0.8)	18.0	38.0
North Thames Ovary Group[19]	1997	cis or carbo (5 cycles)	118	0.91[c d]	1.02	13.0	24.0
		cis or carbo (8 cycles)	115	(0.71–1.16)[c d]	(0.76–1.35)[c]	14.0	25.0
GOCA[25]	1997	cis, cyclo	77	1.2	1	26.0	37.0
		carbo, cyclo	81	~[e]	~[e]	19.0	35.0
ICON2[26]	1998	carbo	760	0.92	1	15.5	33.0
		CAP	766	(0.81–1.04)	(0.86–1.16)	17.0	33.0
GOG 132[27]	2000	cisplatin	200	1.06[b]	0.99[b]	16.4	30.2
		cis, tax	201	(0.86–1.3)	(0.80–1.23)	14	26.6
Danish collaboration[28]	2000	cis, tax	108	1.07	0.85	~[e]	30.0
		carbo, tax	100	(0.78–1.48)	(0.59–1.24)	~[e]	32.0
GOG114/ SWOG[17]	2001	IV (cis, tax)	227	0.78	0.81	22.2	52.2
		IP (IP cis, IV carbo, tax)[f]	235	(0.66–0.94)	(0.65–1.00)	27.9	63.2
OV10 (updated)[20]	2003	cis, cyclo	338	0.81[c]	0.75	11.0	25.8
		cis, tax	342	(0.68–0.95)	(0.63–0.9)	15.0	35.6
ICON3[29]	2002	carbo or CAP	1364	0.93	0.98	16.1	35.4
		carbo, tax	710	(0.84–1.03)	(0.87–1.1)	17.3	36.1
AGO[30]	2003	cis, tax (PT)	386	1.05	1.045	19.1	44.1
		carbo, tax (TC)	397	(0.89–1.23)	(0.87–1.26)	17.2	43.3
GOG 158[31]	2003	cis, tax	400	0.88[b]	0.84[b]	19.4	48.7
		carbo, tax	392	(0.75–1.03)	(0.70–1.02)	20.7	57.4
SCOTROC[32]	2004	carbo, tax	538	0.97[b]	1.13[b]	14.8	36.0
		carbo, docetaxel	539	(0.83–1.13)	(0.92–1.39)	15	35.0
HeCOG[21]	2005	carbo, tax	121	1.01[c]	1.04[c]	38.0	40.6
		tax, carbo, alt, cis	126	(~[e])	(~[e])	39.0	38.6
GOG172[18]	2006	IV (cis, tax)	210	0.80	0.75	18.3	49.7
		IP (IP cis, IV cis, tax)[f]	205	(0.64–1)	(0.58–0.97)	23.8	65.6

(Continued)

Table 1. (Continued)

Trial	Year[a]	Treatment	*n* (per arm)	HR for PFS (95% CI)	HR for OS (95% CI)	Median PFS (months)	Median OS (months)
AGO OVAR[33]	2006	carbo, tax	635	0.95[b]	0.93[b]	17.9	51.5
		TEC (carbo, tax, epirubicin)	647	(0.83–1.07)	(0.81–1.08)	18.4	49.5
GINECO/AGO/OVAR-5[34]	2006	carbo, tax	650	0.97[b]	1.01[b]	18.5	44.5
		carbo, tax, topotecan	658	(0.85–1.1)	(0.86–1.18)	18.2	43.1
GOG 162[35]	2007	carbo, tax (24 h)	140	1[b]	1.17[b]	12.36	29.88
		carbo, tax (96 h)	140	(0.78–1.28)	(0.90–1.52)	12.6	30.48
HeCOG[22]	2008	carbo, tax	223	0.75	0.92[c]	13.25	37.97
		cis, tax, doxorubicin	228	(0.6–0.93)	~[e]	18.13	44.33
GOG 182[14]	2009	carbo, tax C1–8	864			16	44.1
		carbo, tax, gemcitabine C1–8	864	1.03[b]	1.01[b]	16.3	44.1
				(0.92–1.14)	(0.89–1.14)		
		carbo, tax ×8 plus PLD C 1,3,5,7	862	0.98[b]	0.95[b]	16.4	44.2
				(0.88–1.10)	(0.84–1.09)		
		carbo, topotecan C1–4 carbo, tax C5–8	861	1.07[b]	1.05[b]	15.4	40.2
				(0.96–1.19)	(0.93–1.19)		
		carbo, gemcitabine C1–4 carbo, tax C5–8	861	1.04[b]	1.11[b]	15.4	39.6
				(0.93–1.15)	(0.98–1.26)		
AGO OVAR/GINECO/NSGO[36]	2010	carbo, tax (TC)	882	1.18	1.05	19.3	51.5
		carbo, tax, gemcitabine	860	(1.06–1.32)	(0.091–1.2)	17.8	49.5
NCIC/EORTC/GEICO[23]	2010	cis, topotecan C1–4 carbo, tax C5–8	409	1.10[c]	1.08[c]	16.2	~[e]
		carbo, tax C1–8	410	(0.94–1.28)	(0.93–1.27)	14.6	~[e]
MITO-2[37]	2011	carbo, tax	410	0.95	0.89	16.8	53.2
		carbo, PLD	410	(0.81–1.13)	(0.72–1.12)	19.0	61.6
HCOG[38]	2012	carbo, tax ×8 (CP8)	192	1.37[b]	1.21[b]	21.9	52.3
		carbo ×8, tax ×4 (C8P4)	190	(1.05–1.8)	(0.93–1.56)	16.5	46.7
JGOG (updated)[39]	2012	carbo, 3-weekly tax	319	0.75	0.79	17.5	~[e]

Table 1. (Continued)

| Trial	Year[a]	Treatment	*n* (per arm)	HR for PFS (95% CI)	HR for OS (95% CI)	Median PFS (months)	Median OS (months)
		carbo, weekly tax	312	(0.62–0.91)	(0.63–0.99)	28.1	NR
Biological agents							
GOG 218[15]	2011	carbo, tax + bev (short)	625	0.91		11.2	38.7
				(0.80–1.04)			
		carbo, tax + bev	623	0.72		14.1	39.7
				(0.63–0.82)			
ICON 7[16]	2012	carbo, tax	764	0.87		17.4	NR
		carbo, tax + bev	764	(0.77–0.99)		19.8	NR

[a]Year of publication.
[b]Adjusted hazard ratio reported.
[c]Hazard ratio extrapolated from available information.
[d]Time to tumour progression reported.
[e]Result not given or able to be extracted.
[f]Trials of intraperitoneal therapies. All treatments were given intravenously except where indicated.
alt, alternating; bev, bevacizumab; CAP, cyclophosphamide, adriamycin, and cisplatin; carbo, carboplatin; CI, confidence interval; cis, cisplatin; cyclo, cyclophosphamide; HR, hazard ratio; IP, intraperitoneal; IV, intravenous; OS, overall survival; PFS, progression-free survival; PLD, pegylated liposomal doxorubicin; PT, cisplatin/taxol; NR, not reached; tax, paclitaxel; TC, taxol/carbo; TEC, taxol/epirubicin/carbo.
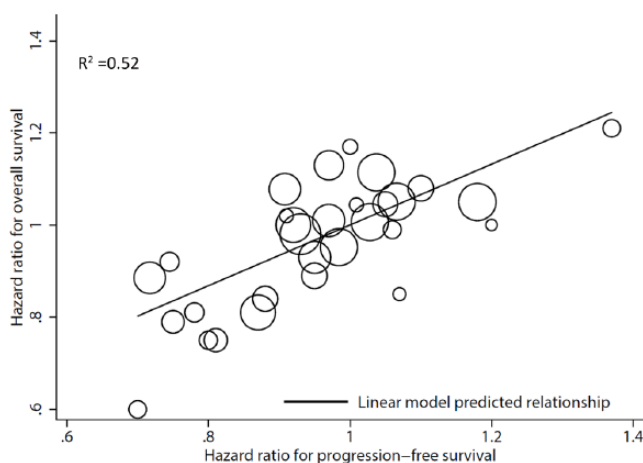
Figure 2. Correlation between hazard ratios for progression-free and overall survival (all trials). The linear regression line is shown. The circles indicate the weighting according to trial size.

Correlations between HRs for PFS and OS varied for different treatment eras (Figure 4): preplatinum/taxane ($n = 8$; unweighted r^2, 0.61, 95% CI 0.01–0.90; r^2 weighted by sample size, 0.66, 95% CI 0.02–0.96), platinum/paclitaxel ($n = 11$; unweighted r^2, 0.44, 95% CI 0.01–0.77; r^2 weighted by sample size, 0.44, 95% CI 0.01–0.77), triplet combination therapies ($n = 7$; unweighted r^2, 0.25, 95% CI 0.00–0.66; r^2 weighted by sample size, 0.22, 95% CI 0.00–0.66), and novel therapies ($n = 4$; unweighted r^2, 0.21, 95% CI 0.00–1.00; r^2 weighted by sample size, 0.30, 95% CI 0.00–0.56) Correlations between HRs for PFS and OS also varied according to PPS.
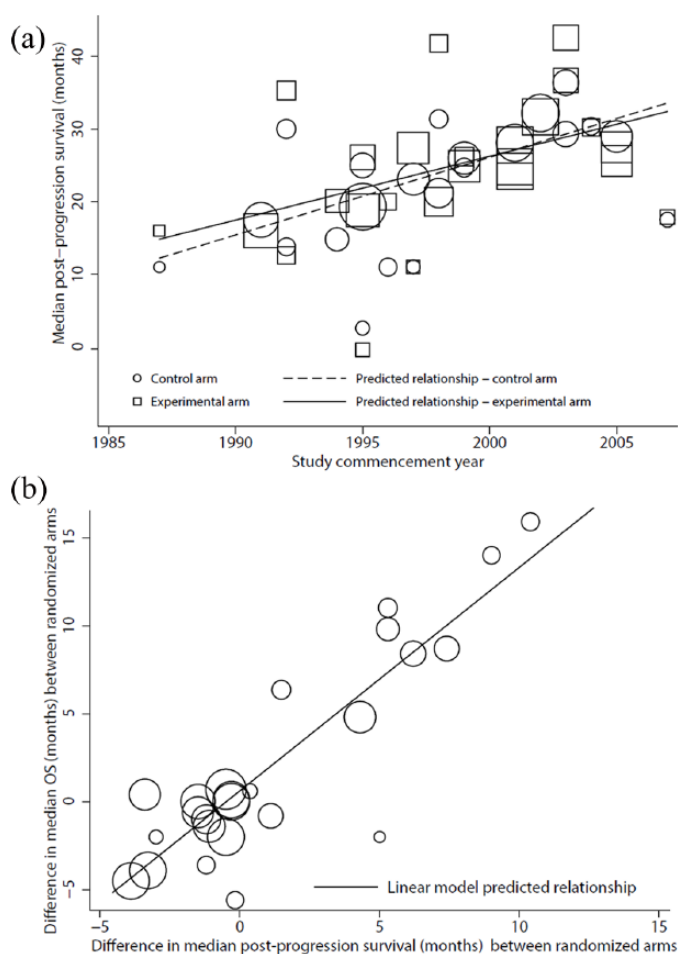
Figure 3. (a). Median postprogression survival by treatment arm over time. The lines show predicted relationships in the experimental arm (solid line) and the control arm (dashed line). The weights according to trial size are shown by squares in the experimental arm and circles in the control arm.
(b). Differences in median progression-free survival and median overall survival (months) between intervention and control arms. The line shows the linear regression line and the circles show the weight according to trial size.

In trials ($n = 8$) where the median PPS was less than 18 months with control therapy, the correlation was higher (unweighted r^2, 0.55, 95% CI 0.01–0.98; r^2 weighted by sample size, 0.64, 95% CI 0.00–0.98) than those trials ($n = 18$) in which the median PPS was at least 18 months (unweighted r^2, 0.59, 95% CI 0.32–0.85; r^2 weighted by sample size, 0.48, 95% CI 0.14–0.71; Figure 5).

In subgroup analyses, trials that included 10% or more patients (median distribution of trial populations) with Eastern Cooperative Oncology Group (ECOG) performance status ⩾2 had stronger correlation between PFS and OS ($n = 9$; unweighted r^2, 0.79, 95% CI 0.12–0.76; r^2 weighted by sample size, 0.76, 95% CI 0.14–0.74)

than trials with less than 10% performance status ⩾2 patients ($n = 18$; unweighted r^2, 0.52, 95% CI 0.04–0.94; r^2 weighted by sample size, 0.53, 95% CI 0.04–0.94; Figure 6). When trials with more patients with stage IV disease (18% or greater of trial populations; median distribution of trial populations) were compared with those with fewer patients (less than 18% of trial population with stage IV disease), the correlations were similar (r^2 weighted by sample size, 0.49 *versus* 0.48)

Table 2 lists the year of US FDA approval of anticancer agents with clinical activity in EOC. Since paclitaxel was approved in 1992, the number of active agents has almost doubled, expanding the options for subsequent lines of therapies beyond the initial trial therapy.
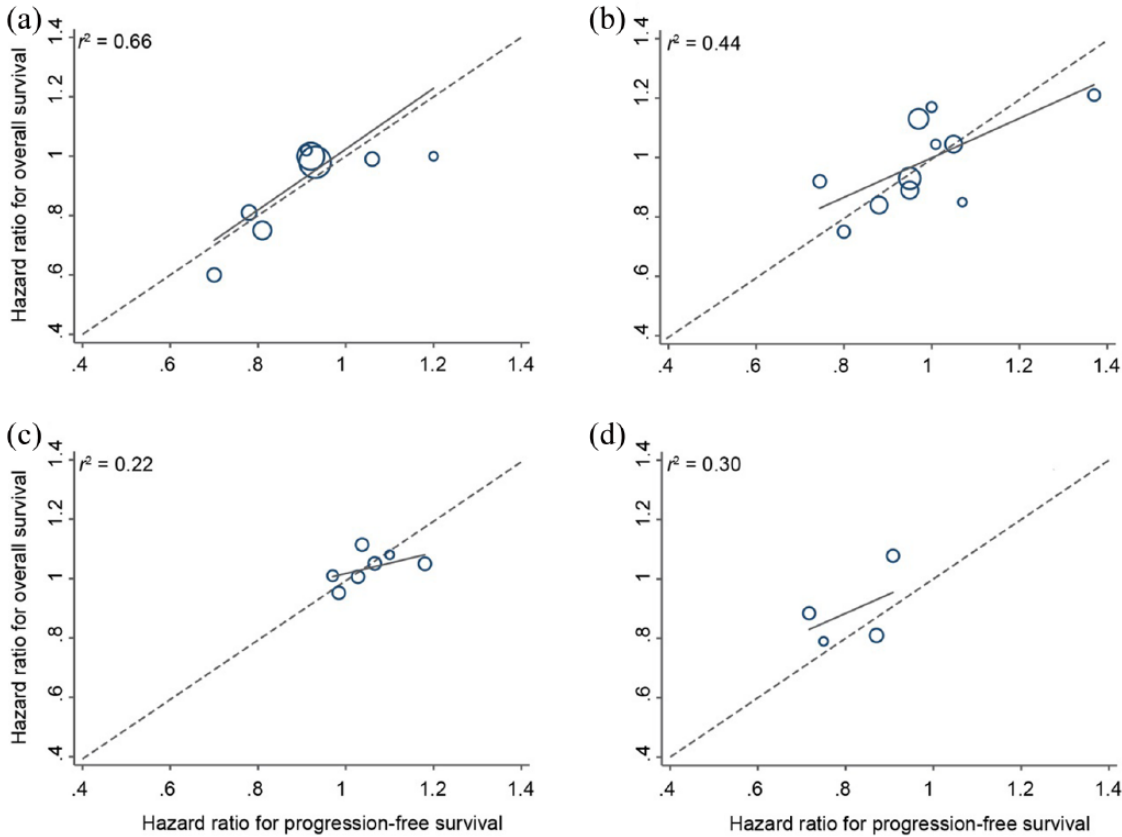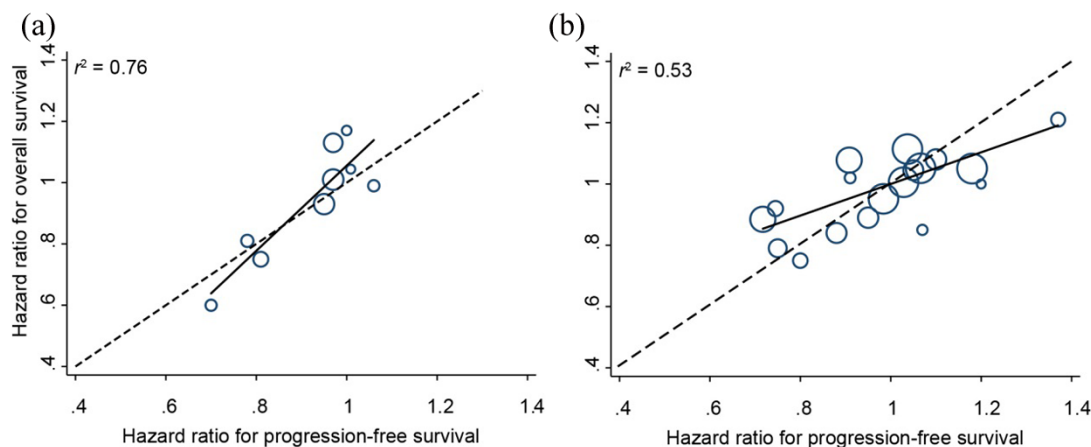
Figure 4. Correlation between hazard ratios for progression-free and overall survival by treatment regimen in different eras, weighted by sample size: (a) preplatinum/paclitaxel; (b) platinum/paclitaxel; (c) triplet combinations; (d) biological and other novel therapies.
—— predicted linear relationship ------ ideal relationship.
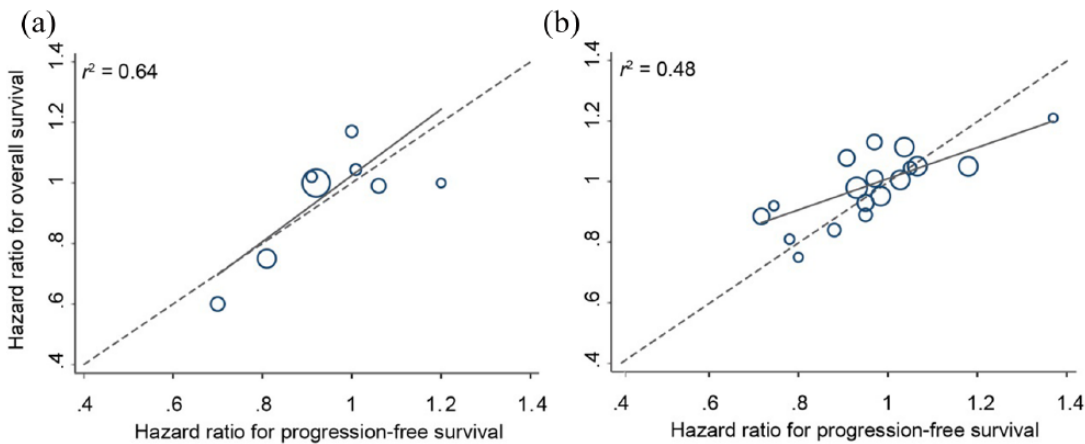
Figure 5. Correlations between hazard ratios for progression-free and overall survival according to postprogression survival. (a) Median postprogression survival less than 18 months. (b) Median postprogression survival at least 18 months.
—— predicted linear relationship ------- ideal relationship ○ weights according to trial size.

In sensitivity analyses, excluding trials of intraperitoneal treatment (unweighted r^2, 0.49, 95% CI 0.18–0.69; r^2 weighted by sample size, 0.49, 95% CI 0.26–0.66), and trials of biological

Figure 6. Correlations between hazard ratios for progression-free and overall survival according to the proportion of patients with poor performance status. (a) Fewer than 10% of patients with Eastern Cooperative Oncology Group performance status $\geq$2; (b) 10% or more patients with Eastern Cooperative Oncology Group performance status $\geq$2.
——— predicted linear relationship ----- ideal relationship ○ weights according to trial size.

therapies (unweighted r^2, 0.58, 95% CI 0.27–0.77; r^2 weighted by sample size, 0.58, 95% CI 0.32–0.76), did not change the overall results significantly.

Discussion

For PFS to be useful as a surrogate endpoint at trial level, a strong correlation between the relative treatment effects on PFS and OS is required.[41] Correlations between PFS and OS have been stronger in studies examining a limited number of EOC trials that included contemporary standard platinum-based therapies, (r^2 ranges from 0.85[42] to 0.94[7]) but not more recent trials, particularly those including biological-targeted and other novel therapies. Moreover, in two different trials conducted almost 10 years apart, the median PPS in EOC almost doubled in cohorts of patients treated with the same therapy of carboplatin-gemcitabine.[43,44] We sought to address this question given its important implications for future trial design, selection of endpoints, drug approvals by regulatory bodies, and healthcare funding.[41]

In clinical trials of advanced EOC, there was only a moderate correlation ($r^2 = 0.52$) between the treatment effects on PFS and OS. When the correlations were examined for different treatment paradigms based on clinical trials conducted in different eras, the strength of the relationship between the HRs for PFS and OS was less for more recent regimens. Our finding of a significant

trend to an increase in the median PPS over time and a strong correlation ($r^2 = 0.83$ (weighted)) between the relative effects of treatment on PPS and OS supports the hypothesis that postprogression therapy can dilute the relationship between PFS and OS. This analysis is limited by the inability to adjust for baseline characteristics in the absence of individual patient data. It is therefore best considered hypothesis generating, with the aim of encouraging further research.

The results of this study differ from the findings of earlier studies, which reported strong correlations in relative treatment effect between PFS and OS.[7,42] One possible explanation for this difference might be changes in the definition of PFS over time. Before 2000, World Health Organization criteria[45] or clinical progression criteria were used to define disease progression in clinical trials. In some of the earlier trials, a second-look laparotomy was planned,[25,27] or was reported to have occurred,[28] and the extent to which the laparotomy findings influenced assessment of progression is unclear from published information. Since then, new guidelines to evaluate the response to treatment and to define progression using both imaging and CA125 levels have been introduced and widely adopted in EOC trials.[46,47]

It is more likely that the impact and greater availability of more effective salvage therapies explain the dilution of the previously observed relationship between the relative effects of treatment on

Table 2. Available salvage therapies for recurrent ovarian cancer.

Name	Year US FDA first approved
Carboplatin	1989 (ovarian) – as paraplatin
Cisplatin	1978 (prior to 1984)
Paclitaxel	1992 (December)
Docetaxel	1996 (for breast cancer)
Gemcitabine	1998 (for lung cancer)
Liposomal doxorubicin	1995
Etoposide	1983
Topotecan	1996
Altretamine	1990
Capecitabine	1998 (breast, 2001 colorectal)
Cyclophosphamide	Prior to 1984
Ifosfamide	1988
Irinotecan	1998 (full, accelerated 1996)
Melphalan	Prior to 1984 (oral form)
Oxaliplatin	2002
Nab-paclitaxel	2012 (for non-small cell lung cancer)
Pemetrexed	2004
Vinorelbine	1988
Bevacizumab	2012 (ovarian)/ 2004 (non-small cell lung cancer)
Anastrozole	1996 (breast)
Letrozole	1998 (breast)
Leuprolide acetate	1985 (prostate)
Megestrol acetate	Prior to 1984
Tamoxifen	Prior to 1984 (1977)
US FDA, United States Food and Drug Administration	

PFS and OS. Few of the trials included in this study provided any details of postprogression therapies or the proportion of patients who crossed over to receive the active experimental therapy at progression. Of all the included trials, only a single study[24] of the six published before 2000 showed a statistically significant benefit of the experimental treatment over control for PFS. In contrast, 6 trials or comparisons[15,17,22,39,48] of 18 published after 2000 reported a statistically significant benefit in favour of the experimental treatment.

The duration of PPS affects the relationship between the relative treatment effects for PFS and OS. Broglio and Berry[13] used simulated data to demonstrate that the probability of a statistical significant difference in OS between treatment arms lessens with increasing duration of PPS, despite a statistical difference for PFS. Our results

in EOC trials support the findings of Broglio and Berry (Figure 3), although our results are limited by reliance on events occurring following randomization, and should therefore be considered exploratory.

It is possible that improved imaging modalities and the increasing use of CA125 to define progression could result in earlier detection of disease recurrence and hence inflate PPS in the more recently conducted trials. However, we do not believe that these factors alone would account for all the improvement in PPS. Availability of effective salvage therapies remains the most likely explanation for the increased PPS over time. This is supported by the differing results of two second-line studies conducted almost a decade apart, the Oceans trial[43] and the AGO-OVAR2.25 trial.[44] Both had a carboplatin-gemcitabine arm. The PFS with carboplatin-gemcitabine in the AGO trial was 8.4 months and in the Oceans trial it was 8.6 months, but the median OS was respectively 18.0 and 32.9 months. The eligibility criteria were very similar, but in the Oceans trial patients had a median of 5 (range 1–14) lines of subsequent treatment, which almost certainly accounted for the significantly longer PPS after second-line therapy.

Our hypothesis of the influence of salvage therapies diluting the relationship between relative treatment effects on PFS and OS is further supported by sensitivity analyses of trials that included a greater proportion of patients with an ECOG performance status of 2. In trials with 10% or more patients with performance status ⩾2, PFS and OS correlated more strongly than in those with less than 10%. We speculate that patients with a poor performance status were less likely to receive second-line salvage therapies, and therefore the relationship between the relative treatment effects on PFS and OS was not compromised.

Our work has a number of limitations. Published summary data, instead of individual patient data, means analyses could not be adjusted for baseline prognostic factors that affect OS or for the number and type of salvage therapies used after initial disease progression. We were also unable to examine the individual patient-level correlations between PFS and OS, which would require individual patient data. Our work is limited to clinical trials of platinum-based chemotherapies because these treatments are considered optimal and

standard first-line therapy for advanced EOC.[10] The result of this study might not be applicable to trials of nonplatinum regimens.

This study has evaluated the relationship between PFS and OS in first-line trials of EOC in the modern era and has demonstrated that the correlation between treatment effects for PFS and OS has weakened. We expect that this relationship will continue to decline with the increasing availability of treatment options, including crossover to the active experimental treatment following disease progression. Therefore, it is increasingly unlikely future trials will demonstrate a relative improvement in treatment effect for OS with first-line therapy. Using OS as primary endpoint will require larger, longer trials in order for first-line treatments to demonstrate an OS benefit. The financial and opportunity costs of such trials make this approach largely infeasible. Other approaches include designing trials so that crossover is not allowed but recognizing that access to other salvage therapies will still occur outside trials. Trials could also be designed with standardized postprogression treatments[49] and meta-analyses of trials with similar class of agents could also be planned prospectively. Furthermore, novel statistical approaches, such as penalized Cox regression[49] that incorporate external estimates of the impact of salvage therapies in order to adjust and preserve the randomized comparisons between different treatment groups could be considered. Finally, a measure of net clinical benefit, such as quality-adjusted PFS,[50] could be considered for treatment recommendations, which would be appropriate even if a relative advantage of OS has not been demonstrated.

Our findings support the fifth GCIG consensus statement,[6] ENREF_5, which advocates the use of PFS as the primary trial endpoint in first-line trials of advanced EOC, but this approach does have limitations. Unlike OS, PFS is more prone to bias, and consequently strict definitions of progression and mandated intervals between imaging studies in trials are essential.[1,4] The value of PFS as the primary endpoint continues to be an issue of ongoing debate, and PFS should be supported and underpinned by additional endpoints, such as patient-reported outcomes, time to second disease progression (PFS2), and time to first and second subsequent treatments.[1,41,51,52] Alternatively endpoints such as quality-adjusted PFS,[50,53] which represent a measure of net clinical benefit, could be used as primary endpoints and for clinical decision making and regulatory approval. It is also

important to demonstrate no OS detriment if PFS is used as the primary endpoint.

In conclusion, the relative treatment effects for PFS and OS are moderately correlated in first-line trials using platinum-based chemotherapy for advanced EOC. This relationship has weakened with time and increasing availability of effective salvage therapies.

Conflict of interest statement
The authors declare that there is no conflict of interest.

Supplemental Material
Supplemental material for this article is available online.

References
1. Herzog TJ, Armstrong DK, Brady MF, *et al.* Ovarian cancer clinical trial endpoints: Society of Gynecologic Oncology white paper. *Gynecol Oncol* 2014; 132: 8–17.

2. DeSantis CE, Lin CC, Mariotto AB, *et al.* Cancer treatment and survivorship statistics, 2014. *CA Cancer J Clin* 2014; 64: 252–271.

3. Oza A, Castonguay V, Tsoref D, *et al.* Progression-free survival in advanced ovarian cancer: a Canadian review and expert panel perspective. *Curr Oncol* 2011; 18: S20–S27.

4. Matulonis UA, Oza AM, Ho TW, *et al.* Intermediate clinical endpoints: a bridge between progression-free survival and overall survival in ovarian cancer trials. *Cancer* 2015; 121: 1737–1746.

5. Saad ED and Buyse M. Overall survival: patient outcome, therapeutic objective, clinical trial end point, or public health measure? *J Clin Oncol* 2012; 30: 1750–1754.

6. Karam A, Ledermann JA, Kim JW, *et al.* Fifth ovarian cancer consensus conference of the gynecologic cancer intergroup: first-line interventions. *Ann Oncol* 2017; 28: 711–717.

7. Burzykowski T, Molenberghs G, Buyse M, *et al.* Validation of surrogate end points in multiple randomized clinical trials with failure time end points. *J R Stat Soc Ser C (Applied Statistics)* 2001; 50: 405–422.

8. Wilkerson J and Fojo T. Progression-free survival is simply a measure of a drug's effect while administered and is not a surrogate for overall survival. *Cancer J* 2009; 15: 379–385.

9. US Food and Drug Administration. Drugs @ FDA: FDA approved drug products, http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm (2015, accessed 7 August 2015).

10. Morgan RJ, Alvarez RD, Armstrong DK, *et al.* Ovarian cancer including fallopian tube cancer prim peritoneal cancer. Version 2 .2013. NCCN clinical practice guidelines in oncology (NCCN guidelines) 2, http://www.nccn.org/professionals/physician_gls/pdf/ovarian.pdf (2013, accessed 15 September 2014).

11. Parmar MK, Torri V and Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Stat Med* 1998; 17: 2815–2834.

12. Moher D, Liberati A, Tetzlaff J, *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; 6: e1000097. DOI: 10.1371/journal.pmed.1000097.

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

14. Bookman MA, Brady MF, McGuire WP, *et al.* Evaluation of new platinum-based treatment regimens in advanced-stage ovarian cancer: a phase III trial of the gynecologic cancer intergroup. *J Clin Oncol* 2009; 27: 1419–1425.

15. Burger RA, Brady MF, Bookman MA, *et al.* Incorporation of bevacizumab in the primary treatment of ovarian cancer. *N Engl J Med* 2011; 365: 2473–2483.

16. Kristensen G, Perren T, Qian W, *et al.* Result of interim analysis of overall survival in the GCIG ICON7 phase III randomized trial of bevacizumab in women with newly diagnosed ovarian cancer. *J Clin Oncol* 2011; 29: Abstract LBA 5006.

17. Markman M, Bundy BN, Alberts DS, *et al.* Phase III trial of standard-dose intravenous cisplatin plus paclitaxel versus moderately high-dose carboplatin followed by intravenous paclitaxel and intraperitoneal cisplatin in small-volume

stage III ovarian carcinoma: an intergroup study of the Gynecologic Oncology Group, Southwestern Oncology Group, and Eastern Cooperative Oncology Group. *J Clin Oncol* 2001; 19: 1001–1007.

18. Armstrong DK, Bundy B, Wenzel L, *et al.* Intraperitoneal cisplatin and paclitaxel in ovarian cancer. *N Engl J Med* 2006; 354: 34–43.

19. Lambert HE, Rustin GJ, Gregory WM, *et al.* A randomized trial of five versus eight courses of cisplatin or carboplatin in advanced epithelial ovarian carcinoma. A North Thames Ovary Group study. *Ann Oncol* 1997; 8: 327–333.

20. Piccart MJ, Bertelsen K, Stuart G, *et al.* Long-term follow-up confirms a survival advantage of the paclitaxel-cisplatin regimen over the cyclophosphamide-cisplatin combination in advanced ovarian cancer. *Int J Gynecol Cancer* 2003; 13(Suppl. 2): 144–148.

21. Aravantinos G, Fountzilas G, Kosmidis P, *et al.* Paclitaxel plus carboplatin versus paclitaxel plus alternating carboplatin and cisplatin for initial treatment of advanced ovarian cancer: long-term efficacy results: a Hellenic Cooperative Oncology Group (HeCOG) study. *Ann Oncol* 2005; 16: 1116–1122.

22. Aravantinos G, Fountzilas G, Bamias A, *et al.* Carboplatin and paclitaxel versus cisplatin, paclitaxel and doxorubicin for first-line chemotherapy of advanced ovarian cancer: a Hellenic Cooperative Oncology Group (HeCOG) study. *Eur J Cancer* 2008; 44: 2169–2177.

23. Hoskins P, Vergote I, Cervantes A, *et al.* Advanced ovarian cancer: phase III randomized study of sequential cisplatin-topotecan and carboplatin-paclitaxel vs carboplatin-paclitaxel. *J Natl Cancer Inst* 2010; 102: 1547–1556.

24. McGuire WP, Hoskins WJ, Brady MF, *et al.* Cyclophosphamide and cisplatin compared with paclitaxel and cisplatin in patients with stage III and stage IV ovarian cancer. *N Engl J Med* 1996; 334: 1–6.

25. Meerpohl HG, Sauerbrei W, Kuhnle H, *et al.* Randomized study comparing carboplatin/cyclophosphamide and cisplatin/cyclophosphamide as first-line treatment in patients with stage III/IV epithelial ovarian cancer and small volume disease. German Ovarian Cancer Study Group (GOCA). *Gynecol Oncol* 1997; 66: 75–84.

26. The ICON Collaborators. ICON2: randomised trial of single-agent carboplatin against three-drug combination of CAP (cyclophosphamide, doxorubicin, and cisplatin) in women with ovarian cancer. *Lancet* 1998; 352: 1571–1576.

27. Muggia FM, Braly PS, Brady MF, *et al.* Phase III randomized study of cisplatin versus paclitaxel versus cisplatin and paclitaxel in patients with suboptimal stage III or IV ovarian cancer: a gynecologic oncology group study. *J Clin Oncol* 2000; 18: 106.

28. Neijt JP, Engelholm SA, Tuxen MK, *et al.* Exploratory phase III study of paclitaxel and cisplatin versus paclitaxel and carboplatin in advanced ovarian cancer. *J Clin Oncol* 2000; 18: 3084–3092.

29. The ICON Collaborators. Paclitaxel plus carboplatin versus standard chemotherapy with either single-agent carboplatin or cyclophosphamide, doxorubicin, and cisplatin in women with ovarian cancer: the ICON3 randomised trial. *Lancet* 2002; 360: 505–515.

30. du Bois A, Lück H-J, Meier W, *et al.* A randomized clinical trial of cisplatin/paclitaxel versus carboplatin/paclitaxel as first-line treatment of ovarian cancer. *J Natl Cancer Inst* 2003; 95: 1320–1329.

31. Ozols RF, Bundy BN, Greer BE, *et al.* Phase III trial of carboplatin and paclitaxel compared with cisplatin and paclitaxel in patients with optimally resected stage iii ovarian cancer: a Gynecologic Oncology Group study. *J Clin Oncol* 2003; 21: 3194–3200.

32. Vasey PA, Jayson GC, Gordon A, *et al.* Phase III randomized trial of docetaxel–carboplatin versus paclitaxel–carboplatin as first-line chemotherapy for ovarian carcinoma. *J Natl Cancer Inst* 2004; 96: 1682–1691.

33. du Bois A, Weber B, Rochon J, *et al.* Addition of epirubicin as a third drug to carboplatin-paclitaxel in first-line treatment of advanced ovarian cancer: a prospectively randomized gynecologic cancer intergroup trial by the Arbeitsgemeinschaft Gynaekologische Onkologie Ovarian Cancer Study Group and the Groupe d'Investigateurs Nationaux pour l'Etude des Cancers Ovariens. *J Clin Oncol* 2006; 24: 1127–1135.

34. Pfisterer J, Weber B, Reuss A, *et al.* Randomized phase III trial of topotecan following carboplatin and paclitaxel in first-line treatment of advanced ovarian cancer: a gynecologic cancer intergroup trial of the AGO-OVAR and GINECO. *J Natl Cancer Inst* 2006; 98: 1036–1045.

35. Spriggs DR, Brady MF, Vaccarello L, *et al.* Phase III randomized trial of intravenous cisplatin plus a 24- or 96-hour infusion of paclitaxel in epithelial ovarian cancer: a Gynecologic Oncology Group Study. *J Clin Oncol* 2007; 25: 4466–4471.

36. du Bois A, Herrstedt J, Hardy-Bessard A-C, *et al.* Phase III trial of carboplatin plus paclitaxel with

or without gemcitabine in first-line treatment of epithelial ovarian cancer. *J Clin Oncol* 2010; 28: 4162–4169.

37. Pignata S, Scambia G, Ferrandina G, *et al.* Carboplatin plus paclitaxel versus carboplatin plus pegylated liposomal doxorubicin as first-line treatment for patients with ovarian cancer: the MITO-2 randomized phase III trial. *J Clin Oncol* 2011; 29: 3628–3635.

38. Bamias A, Timotheadou E, Aravantinos G, *et al.* Randomized, phase III study of carboplatin plus paclitaxel for 8 cycles versus carboplatin x 8 cycles plus paclitaxel x 4 cycles in advanced ovarian, fallopian, or primary peritoneal carcinoma. *J Clin Oncol* 2012; 30: Abstract 5033.

39. Katsumata N, Yasuda M, Isonishi S, *et al.* Long-term follow-up of a randomized trial comparing conventional paclitaxel and carboplatin with dose-dense weekly paclitaxel and carboplatin in women with advanced epithelial ovarian, fallopian tube, or primary peritoneal cancer: JGOG 3016 trial. *J Clin Oncol* 2012; 30: Abstract 5003.

40. Piccart MJ, Bertelsen K, Stuart G, *et al.* Long-term follow-up confirms a survival advantage of the paclitaxel–cisplatin regimen over the cyclophosphamide–cisplatin combination in advanced ovarian cancer. *Int J Gynecol Cancer* 2003; 13: 144–148.

41. Booth CM and Eisenhauer EA. Progression-free survival: meaningful or simply measurable? *J Clin Oncol* 2012; 30: 1030–1033.

42. Wilkerson J and Fojo T. Progression-free survival is simply a measure of a drug's effect while administered and is not a surrogate for overall survival. *Cancer J* 2009; 15: 379–385.

43. Aghajanian C, Blank SV, Goff BA, *et al.* OCEANS: A randomized, double-blind, placebo-controlled phase III trial of chemotherapy with or without bevacizumab in patients with platinum-sensitive recurrent epithelial ovarian, primary peritoneal, or fallopian tube cancer. *J Clin Oncol* 2012; 30: 2039–2045.

44. Pfisterer J, Plante M, Vergote I, *et al.* Gemcitabine plus carboplatin compared with carboplatin in patients with platinum-sensitive recurrent ovarian cancer: an intergroup trial of the AGO-OVAR, the NCIC CTG, and the EORTC GCG. *J Clin Oncol* 2006; 24: 4699–4707.

45. Miller AB, Hoogstraten B, Staquet M, *et al.* Reporting results of cancer treatment. *Cancer* 1981; 47: 207–214.

46. Therasse P, Arbuck SG, Eisenhauer EA, *et al.* New guidelines to evaluate the response to treatment in solid tumors. *J Natl Cancer Inst* 2000; 92: 205–216.

47. Vergote I, Rustin GJS, Eisenhauer EA, *et al.* Re: New guidelines to evaluate the response to treatment in solid tumors [ovarian cancer]. *J Natl Cancer Inst* 2000; 92: 1534–1535.

48. Piccart MJ, Bertelsen K, James K, *et al.* Randomized intergroup trial of cisplatin–paclitaxel versus cisplatin–cyclophosphamide in women with advanced epithelial ovarian cancer: three-year results. *J Natl Cancer Inst* 2000; 92: 699–708.

49. Simes J, Voysey M, O'Connell R, *et al.* A novel method to adjust efficacy estimates for uptake of other active treatments in long-term clinical trials. *PLoS One* 2010; 5: e8580.

50. Stockler MR, Harvey VJ, Francis PA, *et al.* Capecitabine versus classical cyclophosphamide, methotrexate, and fluorouracil as first-line chemotherapy for advanced breast cancer. *J Clin Oncol* 2011; 29: 4498–4504.

51. Bast RC, Thigpen JT, Arbuck SG, *et al.* Clinical trial endpoints in ovarian cancer: report of an FDA/ASCO/AACR public workshop. *Gynecol Oncol* 2007; 107: 173–176.

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

53. Glasziou PP, Simes RJ and Gelber RD. Quality adjusted survival analysis. *Stat Med* 1990; 9: 1259–1276.

54. Alberts DS, Liu P, Hannigan EV, *et al.* Intraperitoneal cisplatin plus intravenous cyclophosphamide versus intravenous cisplatin plus intravenous cyclophosphamide for stage III ovarian cancer. *New England Journal of Medicine* 1996; 335: 1950–1955.

55. Skarlos DV, Aravantinos G, Kosmidis P, *et al.* Carboplatin alone compared with its combination with epirubicin and cyclophosphamide in untreated advanced epithelial ovarian cancer: a hellenic co-operative oncology group study. *European Journal of Cancer* 1996; 32: 421–428. DOI: 10.1016/0959-8049(95)00537-4.

56. Wrigley E, Weaver A, Joyson G, *et al.* A randomised trial investigating the dose intensity of primary chemotherapy in patients with ovarian carcinoma: a comparison of chemotherapy given every four weeks with the same chemotherapy given at three week intervals. *Annals of oncology* 1996; 7: 705–711.

57. Marth C, Trope C, Vergote I, *et al.* Ten-year results of a randomised trial comparing cisplatin with cisplatin and cyclophosphamide in advanced, suboptimally debulked ovarian cancer. *European Journal of Cancer* 1998; 34: 1175–1180.

58. Cocconi G, Bella M, Lottici R, *et al.* Mature results of a prospective randomized trial comparing a three-weekly with an accelerated weekly schedule of cisplatin in advanced ovarian carcinoma. *American journal of clinical oncology* 1999; 22: 559.

59. Polyzos A, Tsavaris N, Kosmas C, *et al.* A comparative study of intraperitoneal carboplatin versus intravenous carboplatin with intravenous cyclophosphamide in both arms as initial chemotherapy for stage III ovarian cancer. *Oncology* 1999; 56: 291–296.

60. Gadducci A, Carnino F, Chiara S, *et al.* Intraperitoneal versus intravenous cisplatin in combination with intravenous cyclophosphamide and epidoxorubicin in optimally cytoreduced advanced epithelial ovarian cancer: a randomized trial of the Gruppo Oncologico Nord-Ovest. *Gynecologic oncology* 2000; 76: 157–162.

61. Merkle E, Ackermann S, Beck EP, *et al.* High-dose versus low-dose cisplatin chemotherapy plus treosulfan in epithelial ovarian carcinoma FIGO II-IV: Results of a prospective randomized trial. *Onkologie* 2000; 23: 232–238.

62. Yen MS, Juang CM, Lai CR, *et al.* Intraperitoneal cisplatin-based chemotherapy vs. intravenous cisplatin-based chemotherapy for stage III optimally cytoreduced epithelial ovarian cancer. *International journal of gynecology & obstetrics* 2001; 72: 55–60.

63. Gordon AN, Teneriello M, Janicek MF, *et al.* Phase III trial of induction gemcitabine or paclitaxel plus carboplatin followed by paclitaxel consolidation in ovarian cancer. *Gynecologic Oncology* 2011; 123: 479–485.

64. Piccart MJ, Bertelsen K, James K, *et al.* Randomized Intergroup Trial of Cisplatin–Paclitaxel Versus Cisplatin–Cyclophosphamide in Women With Advanced Epithelial Ovarian Cancer: Three-Year Results. *Journal of the National Cancer Institute* 2000; 92: 699–708. DOI: 10.1093/jnci/92.9.699.

65. Katsumata N, Yasuda M, Takahashi F, *et al.* Dose-dense paclitaxel once a week in combination with carboplatin every 3 weeks for advanced ovarian cancer: a phase 3, open-label, randomised controlled trial. *The Lancet* 2009; 374: 1331–1338.