Original Investigation | Oncology

Assessment of Progression-Free Survival as a Surrogate End Point of Overall Survival in First-Line Treatment of Ovarian Cancer
A Systematic Review and Meta-analysis

Xavier Paoletti, PhD; Liz-Anne Lewsley, MD, PhD; Gennaro Daniele, MD, PhD; Adrian Cook, MSc; Nozomu Yanaihara, MD, PhD; Anna Tinker, MD; Gunnar Kristensen, MD, PhD; Petronella B. Ottevanger, MD, PhD; Gerasimos Aravantinos, MD, PhD; Austin Miller, MD, PhD; Ingrid A. Boere, MD, PhD; Robert Fruscio, MD, PhD; Anna K. L. Reyners, MD, PhD; Eric Pujade-Lauraine, MD, PhD; Andrea Harkin, BA; Sandro Pignata, MD, PhD; Tatsuo Kagimura, MD, PhD; Stephen Welch, MD, PhD; James Paul, BSc; Eleni Karamouza, MSc; Rosalind M. Glasspool, MD, PhD; for the Gynecologic Cancer InterGroup (GCIG) Meta-analysis Committee

Abstract

IMPORTANCE The Gynecologic Cancer InterGroup (GCIG) recommended that progression-free survival (PFS) can serve as a primary end point instead of overall survival (OS) in advanced ovarian cancer. Evidence is lacking for the validity of PFS as a surrogate marker of OS in the modern era of different treatment types.

OBJECTIVE To evaluate whether PFS is a surrogate end point for OS in patients with advanced ovarian cancer.

DATA SOURCES In September 2016, a comprehensive search of publications in MEDLINE was conducted for randomized clinical trials of systematic treatment in patients with newly diagnosed ovarian, fallopian tube, or primary peritoneal cancer. The GCIG groups were also queried for potentially completed but unpublished trials.

STUDY SELECTION Studies with a minimum sample size of 60 patients published since 2001 with PFS and OS rates available were eligible. Investigational treatments considered included initial, maintenance, and intensification therapy consisting of agents delivered at a higher dose and/or frequency compared with that in the control arm.

DATA EXTRACTION AND SYNTHESIS Using the meta-analytic approach on randomized clinical trials published from January 1, 2001, through September 25, 2016, correlations between PFS and OS at the individual level were estimated using the Kendall τ model; between-treatment effects on PFS and OS at the trial level were estimated using the Plackett copula bivariate (R^2) model. Criteria for PFS surrogacy required $R^2 \geq 0.80$ at the trial level. Analysis was performed from January 7 through March 20, 2019.

MAIN OUTCOMES AND MEASURES Overall survival and PFS based on measurement of cancer antigen 125 levels confirmed by radiological examination results or by combined GCIG criteria.

RESULTS In this meta-analysis of 17 unique randomized trials of standard (n = 7), intensification (n = 5), and maintenance (n = 5) chemotherapies or targeted treatments with data from 11 029 unique patients (median age, 58 years [range, 18-88 years]), a high correlation was found between PFS and OS at the individual level (τ = 0.724; 95% CI, 0.717-0.732), but a low correlation was found at the trial level (R^2 = 0.24; 95% CI, 0-0.59). Subgroup analyses led to similar results. In the external

(continued)

Key Points

Question Is progression-free survival a validated surrogate end point for overall survival in first-line systemic treatment of ovarian cancer?

Findings In this systematic review and meta-analysis of 17 unique trials with individual data from 11 029 unique patients, a high correlation between progression-free and overall survival was found at the individual level, but a low correlation was found at the trial level.

Meaning These findings suggest that overall survival is the preferred end point in trials of first-line treatment or maintenance treatment, and progressive-free survival must be supported by additional end points if used as the primary end point.

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

Abstract (continued)

validation, 14 of the 16 hazard ratios for OS in the published reports fell within the 95% prediction interval from PFS.

CONCLUSIONS AND RELEVANCE This large meta-analysis of individual patient data did not establish PFS as a surrogate end point for OS in first-line treatment of advanced ovarian cancer, but the analysis was limited by the narrow range of treatment effects observed or by poststudy treatment. These results suggest that if PFS is chosen as a primary end point, OS must be measured as a secondary end point.

Introduction

In 2012, approximately 240 000 women worldwide were diagnosed with an advanced ovarian, epithelial, fallopian tube, or primary peritoneal cancer.[1] Approximately 75% of women have Fédération Internationale de Gynécologie et d'Obstétrique (FIGO) stage III or IV cancer at diagnosis. Initial management involves the combination of surgical cytoreduction and systemic chemotherapy. Carboplatin and paclitaxel constitute the universal standard regimen in the management of ovarian cancer, with a response rate of approximately 65%, median progression-free survival (PFS) ranging from 16 to 21 months, and median overall survival (OS) ranging from 32 to 57 months.[2] Currently, OS is the criterion standard for the evaluation of treatment, but both OS and PFS have led to drug approvals by regulatory agencies (the US Food and Drug Administration and European Medicines Agency). Progression-free survival gives an earlier assessment of antitumor activity, requires smaller sample sizes, and is not affected by postprogression therapy. The Gynecologic Cancer InterGroup (GCIG)[3] recommended that PFS can serve as a primary end point instead of OS, provided that secondary end points, such as quality of life, support the superiority of the investigated treatment. Evidence of the validity of PFS as a surrogate marker of OS in the modern era of different treatment types is lacking. In 2009, Buyse[4] showed that PFS was a good surrogate marker of OS in ovarian cancer, but that study was limited to 4 trials that investigated standard cytotoxic regimens (cyclophosphamide plus cisplatin vs cyclophosphamide plus doxorubicin hydrochloride [Adriamycin] plus cisplatin) and used the older World Health Organization definition of progression. A correlation at the individual level measured by a Kendall τ of 0.84 (95% CI, 0.83-0.85) and at the group level measured by a Pearson correlation of 0.95 (95% CI, 0.82-1.00) was found. In these trials, treatment effect on PFS was associated with treatment effect on OS.

Since then, novel targeted therapies have been introduced, many of which are used as maintenance therapy. Among the tools to evaluate progression and response to treatment, cancer antigen 125 (CA125) level is an important marker in epithelial ovarian cancer.[5] The GCIG integrated the elevation of CA125 levels into the radiological Response Evaluation Criteria in Solid Tumours (RECIST) to give a combined definition of progression.[6] These combined criteria have never, to our knowledge, been investigated as an OS surrogate using the meta-analytic approach. Trials use different methods of assessing progression, including clinical or CA125-triggered and regular computed tomographic (CT) scans. The effect of such different assessment methods on the surrogacy of PFS also has not been assessed. To formally assess PFS measured by RECIST and combined GCIG criteria as a potential surrogate end point of overall survival, the GCIG meta-analysis group launched a prospectively planned pooled analysis of data from 11 029 individual patients (individual patient data [IPD]) and 17 randomized clinical trials of first-line therapy (initial treatment, intensification treatment, or maintenance treatment) in advanced ovarian cancers.

Methods

This report follows the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA)–IPD guidelines for the registration of the protocol, trial identification, data collection and integrity, assessment of bias, and sensitivity analyses.[7] This meta-analysis was registered with PROSPERO (CRD42017068135). The Ethics Committee of Gustave Roussy Cancer Center, Villejuif, France, approved this study, and the French data protection authority waived the need for informed consent for the use of deidentified data.

Trial Selection

In September 2016, a comprehensive search in MEDLINE of publications on advanced ovarian cancer was conducted. The GCIG groups were queried for potentially completed but unpublished trials. Eligible trials were randomized clinical trials of systemic treatments in patients with previously untreated ovarian cancer (or investigating maintenance treatment after first-line systemic treatment) with a minimum sample size of 60 patients in total and published from January 1, 2001, through September 25, 2016, with both OS and PFS available. Investigational treatments considered were initial, maintenance, and intensification therapy that consisted of agents delivered at higher dose and/or frequency compared with that in the control arm. The investigators of all identified trials that met the eligibility criteria were contacted for IPD sharing.

Data and Outcomes

We requested data for all individual patients (whether or not they had been included in the primary analysis) enrolled in each trial. Overall survival was defined as the time from randomization to all-cause death or the date of the last follow-up used for censoring. Progression-free survival was defined as the time from randomization to progression or second cancer when this information was available, time to all-cause death, or the date of the last follow-up used for censoring, whichever came first. Detailed information on the type of progression was requested; this included the definition of progression, the radiological and/or clinical evaluation that documented progression, and serial measurements of CA125 levels. Assessment of progression was grouped into 3 main categories: (1) clinical examination and monitoring of 2 increases of CA125 levels to trigger CT scan confirmation of progression, (2) radiological monitoring based on RECIST, and (3) both CA125 levels and radiological assessment in line with the GCIG recommendations. Patients alive without documented disease progression were censored at the date of last follow-up. All data were centrally reanalyzed and checked for inconsistencies. In particular, diagnostic tools for randomization quality were systematically applied.[8,9] Analysis of surrogacy was performed January 7 through March 20, 2019.

Statistical Analysis

Forest plots were used to display the hazard ratios (HRs) overall and for individual trials, which were then used for the evaluation of surrogacy of PFS for OS. The HRs compared the hazard of an event in patients treated with an investigational regimen with the hazard in patients given the control treatment. A fixed-effect approach was implemented, and HRs were obtained from the expected and observed numbers of events. The pooled HR was then adjusted for the trial. The χ^2 heterogeneity test and I^2 statistic were used to investigate the overall heterogeneity between trials.[10] Survival curves were estimated with the actuarial-based approach of Peto et al[11] to account for the multiple trials. Evolution of the median survival time was assessed using a linear trend test at the trial level weighted by the number of events. Surrogacy can be evaluated at 2 different levels. At the individual level, correlation between PFS and OS means that patients with longer PFS are expected to have longer OS. However, this may only reflect the natural history of the disease, whatever the treatment is. For the assessment of the trial-level surrogacy, the treatment effect on PFS was correlated with the treatment effect on OS; in other words, we evaluated how much of the

treatment effect on OS could be predicted from (or explained by) the treatment effect on PFS. We used the Kendall τ (a rank-correlation coefficient) between PFS and OS to assess surrogacy at the individual level and the coefficient of determination (corresponding to the explained variation) between the natural logarithm of the HRs for PFS and OS to assess surrogacy at the trial level.[12-15] For both coefficients, 0 indicates absence of correlation, whereas 1.00 indicates perfect correlation. At the individual level, the association between the distribution of the true (OS) and surrogate (PFS) end points was evaluated using a bivariable model based on the Plackett copula combined with trial-specific Weibull models for PFS and OS.[13] The treatment effects on PFS and OS were obtained from the bivariate model. The linear association between the 2 treatment effects was estimated, which in turn provided the coefficient of determination R^2 for trial. Following the FLASH (Follicular Lymphoma Analysis of Surrogacy Hypothesis) initiative[16] and a report of childhood acute lymphoblastic leukemia,[17] a surrogate was considered to provide a reliable prediction of the treatment effect on OS from the PFS HR, when the trial-level correlation exceeded 0.8 and its 95% prediction interval excluded 0.6. This predefined threshold is arbitrary and served to limit post hoc biases (ie, choice of the threshold based on the data). Analyses were performed on an intention-to-treat basis (all patients analyzed in their allocated group irrespective of possible protocol deviations).

Sensitivity and Subpopulation Analysis

Leave-1-out cross-validation was implemented to assess the prediction performance of the regression model. The validation process was performed on all but 1 trial, and OS HR was predicted from the PFS HR for the left-out trial and compared with the observed value. The process was repeated for each of the 17 trials to identify potential influential trials and investigate the robustness of the results. Preplanned subgroup analyses investigated the surrogacy measures by definition of progression, by study design (initial, intensification, or maintenance treatment), and within trials that used paclitaxel and carboplatin as the control arm.

External Validation

To assess the external validity of our results, we used 16 trials for which we had not been able to receive IPD from the sponsors. Two of us (X.P. and E.K.) independently extracted the HRs and confidence intervals for PFS and OS from summary statistics published in these trials.[18] The HR on PFS reported in the publication served to predict HR on OS that we in turn compared with the published HR on OS. All analyses were done using SAS, version 9.4 (SAS Institute Inc), with macros developed by Tomasz Burzykowski, PhD, and R, version X, using R surrosurv package, version 1.1.25 (R Project for Statistical Computing).[19] Two-tailed $P < .05$ calculated using the test for heterogeneity was considered to signify statistical significance. Confidence and prediction intervals were computed at the 95% level.

Results

Trials' Descriptions

As illustrated in eFigure 1 in the Supplement, 37 trials were identified from the literature search and their investigators were contacted. Individual patient data were obtained on 11 029 unique patients from 17 unique eligible randomized clinical trials with documented OS and PFS.[2,20-36] **Table 1** lists the trial-level characteristics of the 17 studies; eTable 1 in the Supplement gives an assessment of the risk of bias. In 10 trials,[20,21,24-29,32,36] carboplatin and taxanes were the comparator. Seven studies[20,22,24,26,28,32,36] investigated initial treatment; 5 studies,[21,24,25,30,33] intensification treatment; and 5 studies,[23,27,31,34,35] maintenance treatment. Four trials tested molecularly targeted treatments.[23,27,31,34] A total of 10 trials[21-23,25,26,30,33,34,36] used CA125 levels to trigger follow-up CT scans after an initial increase in the biomarker. Six trials[24,27-29,31,35] used the GCIG criteria (1 multinational trial used both), and 2 trials[20,32] used CT scan only. Data on both end points were available for all 11 029 patients, of whom 7436 experienced progression and 5138 died during

follow-up. Detailed information about patients' characteristics by allocated treatment arm and median follow-up are provided in eTable 2 in the Supplement. Median patient age was 58 years (range, 18-88 years); 5990 (54.3%) had Eastern Cooperative Oncology Group performance status of

Table 1. Trial Characteristics

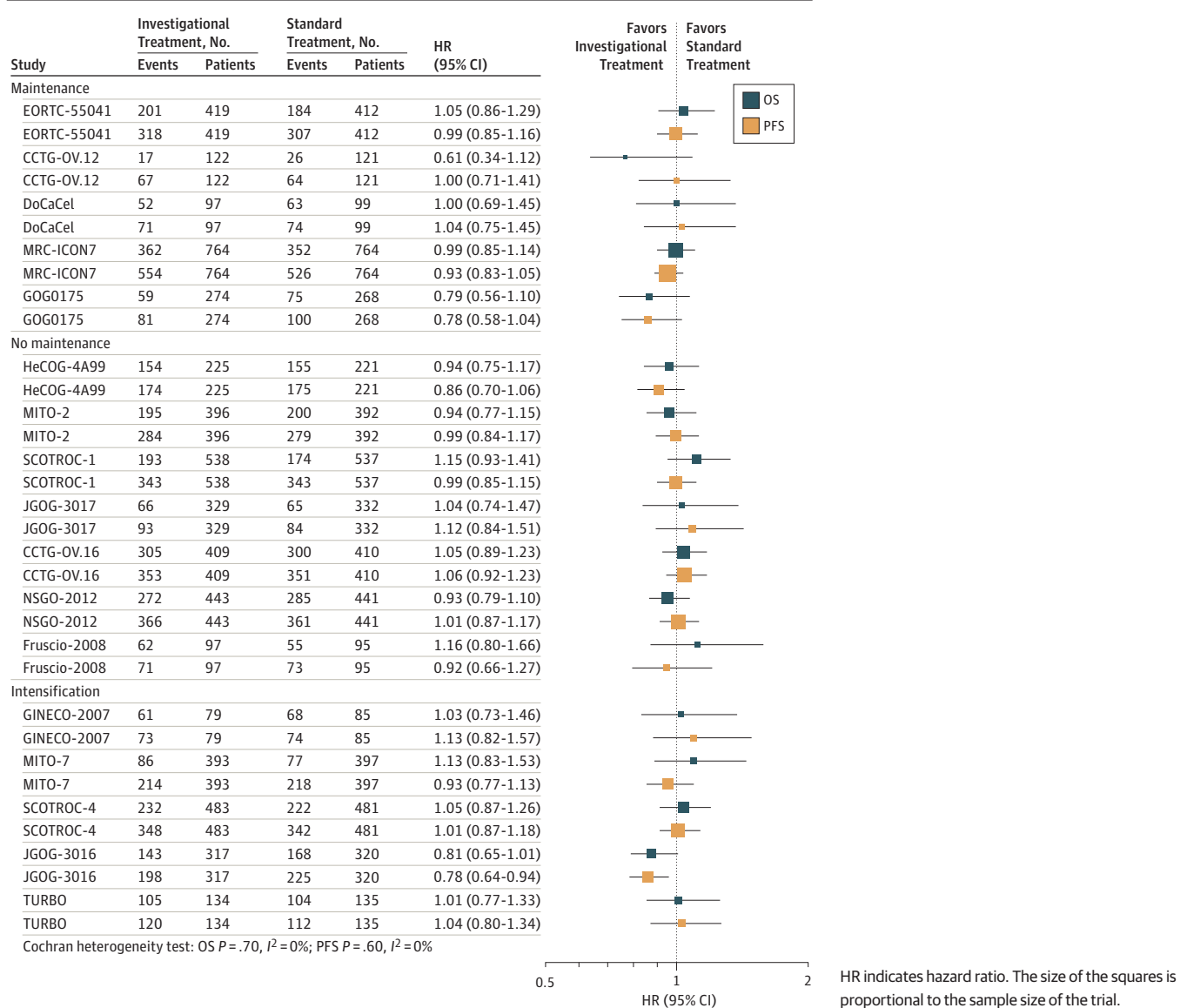| Source (Trial Name)	Investigational Regimen (No. of Patients)	Standard Regimen (No. of Patients)	Assessment of Progression[a]	Standard Arm, No. of Patients	Investigational Arm, No. of Patients	First Inclusion Date	Follow-up, Median (IQR), y	Median OS, y	Median PFS, y
Maintenance									
Vergote et al,[34] 2014 (EORTC-55041)	Erlotinib hydrochloride (420)	Observation (415)	Clinical CA125 level (confirmation with CT)	412	419	2005	4.3 (3.8-4.8)	4.6	1.0
Hirte et al,[23] 2006 (CCTG-OV.12)	Tanomastat (122)	Placebo (121)	Clinical CA125 level (confirmation with CT)	121	122	1998	0.9 (0.6-1.3)	NR	0.9
Reyners et al,[31] 2012 (DoCaCel)	Docetaxel, carboplatin, and celecoxib (97)	Docetaxel and carboplatin (99)	GCIG criteria	99	97	2003	4.1 (2.6-5.7)	2.9	1.2
Oza et al,[27] 2015 (MRC-ICON7)	Bevacizumab (764)	Standard chemotherapy (764)	GCIG criteria	764	764	2006	4.6 (4.2-5.1)	4.8	1.6
Mannel et al,[35] 2011 (GOG-0175)	Low-dose paclitaxel (274)	Observation (268)	GCIG criteria	268	274	1998	11.6 (8.5-13.7)	NR	NR
No Maintenance									
Aravantinos et al,[20] 2008 (HECOG-4A99)	Cisplatin, paclitaxel, and doxorubicin (236)	Paclitaxel and carboplatin (233)	CT scan	221	225	1999	13.7 (5.4-16.1)	3.2	1.3
Pignata et al,[28] 2011 (MITO-2)	Carboplatin and liposomal doxorubicin (410)	Carboplatin and paclitaxel (410)	Mixed[b]	392	396	2003	6.0 (5.0-7.1)	4.7	1.5
Vasey et al,[36] 2004 (SCOTROC-1)	Docetaxel and carboplatin (539)	Paclitaxel and carboplatin (538)	Clinical CA125 level (confirmation with CT)	537	538	1998	2.0 (1.6-2.4)	2.9	1.2
Sugiyama et al,[32] 2016 (JGOG-3017)	Irinotecan hydrochloride and cisplatin (332)	Carboplatin and paclitaxel (335)	CT scan	332	329	2009	3.7 (2.8-4.8)	NR	NR
Hoskins et al,[24] 2010 (CCTG-OV.16)	Cisplatin and topotecan followed by paclitaxel and carboplatin (409)	Paclitaxel and carboplatin (410)	GCIG criteria	410	409	2002	8.2 (7.5-8.9)	3.7	1.3
Lindemann et al,[26] 2012 (NSGO-2012)	Paclitaxel, carboplatin, and epirubicin hydrochloride (445)	Paclitaxel and carboplatin (442)	Clinical CA125 level (confirmation with CT)	441	443	1999	5.3 (4.3-5.9)	3.4	1.4
Fruscio et al,[22] 2008	Cisplatin, ifosfamide, and paclitaxel (106)	Cisplatin, epirubicin hydrochloride, and paclitaxel (103)	Clinical CA125 level (confirmation with CT)	95	97	1997	6.8 (6.2-7.3)	4.7	1.9
Intensification Therapy									
Ray-Coquard et al,[30] 2007 (GINECO-2007)	Cyclophosphamide, erubicin hydrochloride, cisplatin, and filgrastim (79)	Cyclophosphamide, erubicin hydrochloride, and cisplatin (85)	Clinical CA125 level (confirmation with CT)	85	79	1994	8.6 (6.2-9.9)	2.7	1.2
Pignata et al,[29] 2014 (MITO-7)	Weekly carboplatin and paclitaxel (406)	Every 3 wk carboplatin and paclitaxel (404)	GCIG criteria	397	393	2008	1.9 (1.4-2.6)	4.0	1.5
Banerjee et al,[21] 2013 (SCOTROC-4)	Carboplatin dose escalated (483)	Carboplatin flat dose (481)	Clinical CA125 level (confirmation with CT)	481	483	2005	2.7 (1.7-3.6)	2.7	1.0
Katsumata et al,[25] 2013 (JGOG-3016)	Dose-dense carboplatin (317)	Conventional carboplatin (320)	Clinical CA125 level (confirmation with CT)	320	317	2004	6.5 (5.9-7.2)	6.2	2.3
Van der Burg et al,[33] 2014 (TURBO)	Weekly paclitaxel and carboplatin (134)	3 Times per week paclitaxel and carboplatin (136)	Clinical CA125 level (confirmation with CT)	135	134	1998	9.4 (8.4-11.4)	3.6	1.5

Abbreviations: CA125, cancer antigen 125; CT, computed tomography; GCIG, Gynecologic Cancer InterGroup; IQR, interquartile range; NR, not reached; OS, overall survival; PFS, progression-free survival.

[a] "GCIG criteria" indicates that patients were followed up with both serial measurements of CA125 levels and radiological measurements.

[b] Progression of Groupe d'investigateurs national des Etudes des Cancers Ovariens (GINECO) patients was evaluated by CA125 level and confirmed by CT scan, whereas Multicenter Italian Trials in Ovarian Cancer and Gynecologic Malignancies (MITO) patients were evaluated following the GCIG guidelines.

0 at enrollment; and 8497 (77.0%) had FIGO stage III or IV disease. eFigure 2 in the Supplement shows the Peto survival curves for PFS and OS. No statistically significant time trends in the median OS and PFS according to the date of the first randomization were detected. The median OS ranged from 2.7 to 6.2 months and the median PFS ranged from 0.9 to 2.3 months (see Table 1 for median survival and eFigure 3 in the Supplement for their representation over time). No time trends according to the date of the first randomization were detected. **Figure 1** shows a forest plot of the treatment effects on OS and PFS for all trials (eFigure 4 in the Supplement gives forest plots grouped by progression assessment criteria). Overall and at the trial level, the effects of investigational chemotherapy on PFS and OS were almost null (HR for PFS, 0.97 [95% CI, 0.93-1.02]; HR for OS, 0.99 [95% CI, 0.94-1.05]). No heterogeneity across trials was detected for any of the end points (I^2 = 0% [P = .70] for OS and I^2 = 0% [P = .60] for PFS) (Figure 1).

Figure 1. Overall and Trial by Trial Treatment Effect on Overall Survival (OS) and Progression-Free Survival (PFS)

Study	Investigational Treatment, No.		Standard Treatment, No.		HR (95% CI)
	Events	Patients	Events	Patients	
Maintenance					
EORTC-55041	201	419	184	412	1.05 (0.86-1.29)
EORTC-55041	318	419	307	412	0.99 (0.85-1.16)
CCTG-OV.12	17	122	26	121	0.61 (0.34-1.12)
CCTG-OV.12	67	122	64	121	1.00 (0.71-1.41)
DoCaCel	52	97	63	99	1.00 (0.69-1.45)
DoCaCel	71	97	74	99	1.04 (0.75-1.45)
MRC-ICON7	362	764	352	764	0.99 (0.85-1.14)
MRC-ICON7	554	764	526	764	0.93 (0.83-1.05)
GOG0175	59	274	75	268	0.79 (0.56-1.10)
GOG0175	81	274	100	268	0.78 (0.58-1.04)
No maintenance					
HeCOG-4A99	154	225	155	221	0.94 (0.75-1.17)
HeCOG-4A99	174	225	175	221	0.86 (0.70-1.06)
MITO-2	195	396	200	392	0.94 (0.77-1.15)
MITO-2	284	396	279	392	0.99 (0.84-1.17)
SCOTROC-1	193	538	174	537	1.15 (0.93-1.41)
SCOTROC-1	343	538	343	537	0.99 (0.85-1.15)
JGOG-3017	66	329	65	332	1.04 (0.74-1.47)
JGOG-3017	93	329	84	332	1.12 (0.84-1.51)
CCTG-OV.16	305	409	300	410	1.05 (0.89-1.23)
CCTG-OV.16	353	409	351	410	1.06 (0.92-1.23)
NSGO-2012	272	443	285	441	0.93 (0.79-1.10)
NSGO-2012	366	443	361	441	1.01 (0.87-1.17)
Fruscio-2008	62	97	55	95	1.16 (0.80-1.66)
Fruscio-2008	71	97	73	95	0.92 (0.66-1.27)
Intensification					
GINECO-2007	61	79	68	85	1.03 (0.73-1.46)
GINECO-2007	73	79	74	85	1.13 (0.82-1.57)
MITO-7	86	393	77	397	1.13 (0.83-1.53)
MITO-7	214	393	218	397	0.93 (0.77-1.13)
SCOTROC-4	232	483	222	481	1.05 (0.87-1.26)
SCOTROC-4	348	483	342	481	1.01 (0.87-1.18)
JGOG-3016	143	317	168	320	0.81 (0.65-1.01)
JGOG-3016	198	317	225	320	0.78 (0.64-0.94)
TURBO	105	134	104	135	1.01 (0.77-1.33)
TURBO	120	134	112	135	1.04 (0.80-1.34)

Cochran heterogeneity test: OS P = .70, I^2 = 0%; PFS P = .60, I^2 = 0%

HR indicates hazard ratio. The size of the squares is proportional to the sample size of the trial.

Individual- and Trial-Level Associations

The individual-level association, as measured by the Spearman rank correlation coefficient, reached 0.885 (95% CI, 0.879-0.890). The Kendall τ estimate was 0.724 (95% CI, 0.717-0.732), indicating a good correlation between PFS and OS; that is, a patient who progresses later is more likely to survive longer than a patient who progresses earlier. On the contrary, a very low correlation was noted between ln(OS HR) and ln(PFS HR) (**Figure 2**), where ln denotes the natural log transformation of the HR for each end point. The coefficient of determination, R^2 for trial, for the estimated treatment effects was as low as 0.24 (95% CI, 0-0.59), indicating a low correlation between PFS and OS at the trial level. The linear regression model from the copula estimates was ln(OS HR) = 0.025 + [0.67 × ln(PFS HR)]. Standard errors were 0.03 and 0.31 for the intercept and slope, respectively. This is shown as a straight line in Figure 2, where the x-axis represents the treatment effect on PFS and the y-axis represents the treatment effect on OS. The shaded area corresponds to the 95% prediction limits that indicate the range of effect on OS that can be expected for a given effect on PFS, but owing to the very poor correlation, it remains largely theoretical. Despite large sample sizes, some trials with similar treatment effect on PFS had a different effect on OS, including PFS HR of greater than 1.00 together with OS HR of less than 1.00, translating into uncertainty in the prediction.

Sensitivity Analyses

Leave-1-out cross-validation demonstrated the robustness of the results, because we had consistency between observed and predicted OS treatment effects for each trial based on the PFS (eFigure 5 in the Supplement). Only 1 strongly influential trial was identified; the OV-12 trial[23] investigated tanomastat as maintenance therapy, which was interrupted by Bayer owing to negative results in other cancer types, and follow-up was stopped[23]; progression was assessed using CT scans after initial increase of CA125 levels. Excluding this trial increased the estimate of R^2 for trial to a moderate value of 0.66 (95% CI, 0.40-0.93), more in line with previous results.

Subgroup Analyses

Subpopulation analyses that separately focused on maintenance and nonmaintenance trials confirmed that treatment effect on PFS poorly predicted treatment effect on OS: trial-level surrogacy was low for maintenance trials (**Table 2**), with R^2 for trial estimates from 0.03 (95% CI, 0-0.35) for maintenance vs 0.67 (95% CI, 0.36-0.97) for nonmaintenance. The marked difference was mainly explained by the OV-12 trial in the maintenance subgroup, because the R^2 for the trials increased to

Figure 2. Association Between the Hazard Ratio (HR) for the Surrogate End Point Progression-Free Survival (PFS) and for the True End Point Overall Survival (OS) by Type of Trial
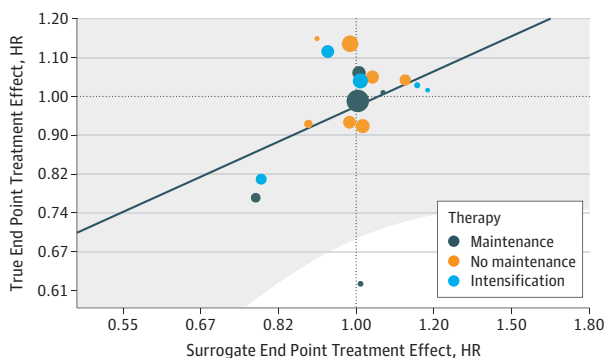
Each trial is represented by a bubble of a size proportional to the trial sample size. The solid straight line is the linear regression model from the copula estimates that relates the PFS HR to OS HR: ln(OS HR) = 0.025 + [0.67 × ln(PFS HR)]. The shaded area corresponds to the 95% prediction limits.

0.78 (95% CI, 0.40-1.00) after exclusion of this trial; the small number of trials in this subgroup strongly increased the results' instability. Trial-level correlation was also low (R^2 for trial = 0.15; 95% CI, 0-0.56) in trials that compared investigational treatment with carboplatin and taxanes. In the 6 trials[24,27-29,31,35] (4603 patients) that specified GCIG guidelines to assess progression, prediction of OS HR based on PFS HR was better (R^2 for trial = 0.43; 95% CI, 0.02-1.00) than that in trials that used CT scan after the initial increase of CA125 level; however, the OV-12 trial[23] again reduced the estimated association between the treatment effects in the trials that used other assessments of progression.

External Validation

Of the 20 trials in which we could not access the IPD (owing to refusal by the investigators,[37-56] no response to our request, or data declared no longer available), we could extract HRs for 16 of them (8 testing initial treatments and 8 testing maintenance treatments).[37-52] None of the trials demonstrated a statistically significant effect on OS, and only 3 studies[43,49,51] reported a statistically significant reduction of PFS. eTable 3 in the Supplement and **Figure 3** display observed OS HR and PFS HR with 95% CIs, and OS HR predicted from the model of Figure 2. Observed estimates for all except 2 trials fell within the 95% prediction intervals. However, the intervals are relatively large, reflecting the uncertainty around the prediction.

Discussion

This pooled analysis was performed on the IPD of 11 029 patients treated in 17 randomized trials of first-line treatment for advanced ovarian cancer initiated worldwide from 1995 through 2010. Although PFS was strongly associated with OS at the individual level, we did not find a strong correlation between the treatment effects on PFS and on OS (ie, HR on PFS did not predict the HR on OS at the trial level). Low correlation was observed in maintenance and nonmaintenance therapy trials. Overall, HRs on OS and PFS were close to 1.00, with little heterogeneity among trials whether maintenance or nonmaintenance treatments were explored. All trials required rigorous response assessment schedules, with clinical and physical examination, evaluation of CA125 levels, and CT imaging. At the trial level, PFS assessed by CT scans and CA125 levels following the GCIG guidelines was moderately correlated with OS in a subgroup of 6 trials.[24,27-29,31,35] Nevertheless, the role of CA125 measurements is controversial. No international standard has been established, leading to variability in calibration, assay design, and reagent specificities,[57] and CA125 level is not considered a stand-alone marker of progression.

One trial can be seen as an outlier; the tanomastat trial was interrupted by Bayer owing to negative results in pancreatic and small cell lung cancer trials, resulting in poor follow-up for OS[23]; exclusion of this trial led to moderate trial-level associations. Nevertheless, even after exclusion of this trial, the trial-level correlation was below the predefined threshold. In trials for which we could

Table 2. Overall and Subgroup Analyses of the Surrogacy of Progression-Free Survival for Overall Survival

Analysis	No. of Trials	No. of Patients	Individual-Level Correlation, Kendall τ (95% CI)[a]	Trial-Level Correlation, R^2 (95% CI)[b]
Overall	17	11 029	0.724 (0.717-0.732)	0.24 (0-0.59)
Design				
Maintenance	5	3340	0.72 (0.71-0.74)	0.03 (0-0.35)
Nonmaintenance	12	7689	0.72 (0.72-0.73)	0.67 (0.36-0.97)
Carboplatin and taxanes as control	10	7321	0.73 (0.72-0.74)	0.15 (0-0.56)
Progression assessment				
CA125 level confirmed by CT scan	10	5319	0.70 (0.69-0.71)	0.27 (0-0.74)
GCIG criteria	5	4603	0.74 (0.73-0.75)	0.43 (0.02-1.00)

Abbreviations: CA125, cancer antigen 125; CT, computed tomography; GCIG, Gynecologic Cancer InterGroup.

[a] Drawn from the joint Plackett copula model that quantifies the strength of the association between progression-free survival and overall survival for a given patient.

[b] Indicates the determination coefficient that quantifies the strength of the association between the treatment effects on progression-free survival (progression-free survival hazard ratio) and overall survival (overall survival hazard ratio).

not access IPD, no statistically significant treatment effect on OS had been reported. As external validation, we showed that for those 16 trials, the observed treatment effect (OS HR) fell within the interval predicted from PFS HR, but the interval was too large to draw accurate predictions on OS HR. These findings therefore do not support PFS as a substitute for OS in randomized clinical trials: demonstrating a reduction of PFS HR does not guarantee that a reduction in the hazard of death will be observed. If PFS is used, the GCIG criteria might be preferable as the means of assessment of progression.

Previous exploration of surrogacy in trials of first-line treatments in ovarian cancers by Buyse[4] found high correlation, but with 4 trials that were split into subunits to increase the number of treatment effect assessments. More recently, several authors[58,59] found moderate to high correlations at the trial level (R^2 range, 0.50-0.83) from summary statistics extracted from the literature. However, unlike IPD, literature-based meta-analysis does not enable consistent calculation of end points or the full use of survival-censored data after quality checks; in addition, estimation of joint model and hence accounting for the correlated PFS and OS measured in the same patient is insufficient, leading to potential biases.

The choice of the best measure to quantify the treatment effect is controversial. Although the HR is probably the most commonly used relative measure, its validity is limited by the requirement to have a proportional hazard (ie, that the HR is constant over time). However, in the clinical trial International Collaboration on Ovarian Neoplasm (ICON7),[27] this assumption did not hold for bevacizumab as maintenance treatment. The primary analysis was then based on an absolute measure, the difference in the restricted mean survival time between the 2 arms. The question of the surrogacy value of restricted mean survival times is to be explored in further analyses.

Limitations

The main limit of our approach is the lack of treatment effects as measured by HR in the collected trials. Indeed, the lack of heterogeneity in the treatment effects strongly limits our ability to detect an association between PFS HR and OS HR. The regression line in Figure 2 may have been more precisely estimated if HRs had been spread across a large range. However, as shown by the validation analysis, the trials that were not collected were also negative and followed the same association between PFS HR and OS HR; additional trials should not strongly modify the conclusions obtained from this large sample. The treatment of ovarian cancers is well standardized, probably thanks to the tradition of strong collaboration within the GCIG and European Society of Gynecological Oncological Trial groups (ENGOT). Most trials enrolled large numbers of patients, were multicentric (and many

Figure 3. Observed and Estimated Treatment Effect on Overall Survival in Validation Trials
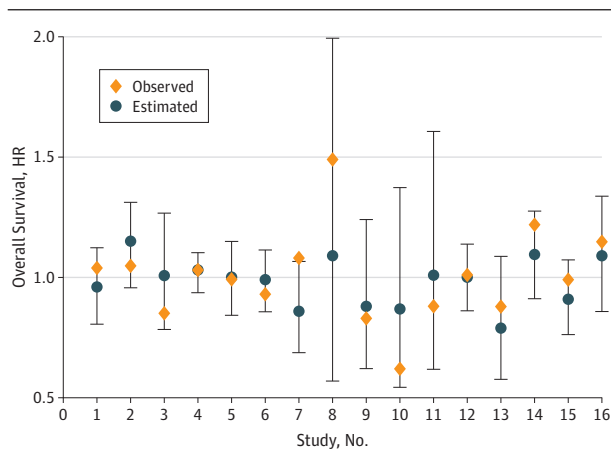
HR indicates hazard ratio; error bars, 95% CIs.

international), shared the same regimen as a control, and collected similar variables. This may explain the strong homogeneity in the trials' results; this also supports the generalizability of our findings.

A striking finding is the disappointing treatment effects measured on the PFS and the OS. This pooled analysis provides a useful benchmark for future trials. We hope that the recent improvements in PFS seen in a trial of poly–adenosine diphosphate ribose polymerase inhibitors[60] translate into improvements in OS. So far, the combination of carboplatin and paclitaxel remains the standard chemotherapy backbone for first-line treatment.

Conclusions

Progression-free survival cannot be validated as a strict surrogate of OS for assessing treatment effects in randomized clinical trials of first-line treatments of advanced ovarian cancers. Our findings support the GCIG Fifth Ovarian Cancer Consensus Conference statement that OS is the preferred primary end point for first-line clinical trials with or without a maintenance component,[3] but we recognize the practical challenges and the potential for confounding factors such as crossover and long postprogression survival. Progression-free survival is an alternative primary end point, but given that we have not been able to validate it as a surrogate of OS, following the US Food and Drug Administration and European Medicines Agency guidances,[61,62] it should represent a favorable risk-benefit association with a large magnitude of the effect or it should contribute to delaying administration of more toxic therapies as second-line treatments; therefore, if PFS is chosen, OS must be measured as a secondary end point and PFS must be supported by additional end points, such as predefined patient-reported outcomes, especially for maintenance therapy.

Corresponding Author: Xavier Paoletti, PhD, Department of Biostatistics, Institut Curie, 35 Rue Dailly, 92210 Saint-Cloud, France (xavier.paoletti@curie.fr).

Author Affiliations: Groupe d'investigateurs national des Etudes des Cancers Ovariens (GINECO), Paris, France (Paoletti); Gustave Roussy Cancer Center and Institut National de la Santé et de la Recherche Medicale Oncostat, Villejuif, France (Paoletti); Department of Biostatistics, University of Versailles St Quentin, Institut Curie, Saint-Cloud, France (Paoletti); Scottish Gynaecological Cancer Trials Group (SGCTG), Cancer Research United Kingdom Clinical Trial Unit, Institute of Cancer Sciences, University of Glasgow, Glasgow, United Kingdom (Lewsley, Harkin, Paul); Multicenter Italian Trials in Ovarian Cancer and Gynecologic Malignancies (MITO), Clinical Trials Unit, Istituto Nazionale Tumori– Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS), Fondazione G. Pascale, Napoli, Italia (Daniele); Medical Research Counsel Clinical Trials Unit, University College London, London, United Kingdom (Cook); Japanese Gynecologic Oncology Group (JGOG), Jikei University School of Medicine, Tokyo, Japan (Yanaihara); Canadian Cancer Trials Group (CCTG), University of British Columbia, Vancouver, British Columbia, Canada (Tinker); Nordic Society of Gynaecological Oncology, Norwegian Radium Hospital, Oslo, Norway (Kristensen); European Organisation for Research and Treatment of Cancer, Radboud University Medical Center, Nijmegen, the Netherlands (Ottevanger); Hellenic Cooperative Oncology Group, General Oncology Hospital of Kifissia, Nea Kifissia, Greece (Aravantinos); Gynecologic Oncology Group (GOG), Roswell Park Comprehensive Cancer Center, Buffalo, New York (Miller); Department of Medical Oncology, Erasmus Medical Center Cancer Institute, Rotterdam, the Netherlands (Boere); University of Milan Bicocca, San Gerardo Hospital, Monza, Italy (Fruscio); University of Groningen, University Medical Center Groningen, Groningen, the Netherlands (Reyners); Association de Recherche sur les Cancers dont Gynécologiques–GINECO, Université Paris Descartes, Assistance Publique–Hôpitaux de Paris, Paris, France (Pujade-Lauraine); MITO, Istituto Nazionale Tumori di Napoli IRCCS Fondazione G Pascale, Napoli, Italy (Pignata); JGOG, Foundation for Biomedical Research and Innovation at Kobe, Translational Research Center for Medical Innovation, Kobe, Japan (Kagimura); CCTG, London Health Sciences Centre, London, Ontario, Canada (Welch); Gustave Roussy Cancer Center, Villejuif, France (Karamouza); SGCTG,

Beatson West of Scotland Cancer Centre, NHS (National Health Service) Greater Glasgow and Clyde, Glasgow, United Kingdom (Glasspool).

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136(5):E359-E386. doi:10.1002/ijc.29210

2. Ozols RF, Bundy BN, Greer BE, et al; Gynecologic Oncology Group. 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(17):3194-3200. doi:10.1200/JCO.2003.02.153

3. Bookman MA, Okamoto A, Stuart G, et al; 5th Ovarian Cancer Consensus Conference. Harmonising clinical trials within the Gynecologic Cancer InterGroup: consensus and unmet needs from the Fifth Ovarian Cancer Consensus Conference. *Ann Oncol*. 2017;28(suppl 8):i30, i35. doi:10.1093/annonc/mdx449

4. Buyse M. Use of meta-analysis for the validation of surrogate endpoints and biomarkers in cancer trials. *Cancer J*. 2009;15(5):421-425. doi:10.1097/PPO.0b013e3181b9c602

5. Sölétormos G, Duffy MJ, Othman Abu Hassan S, et al. Clinical use of cancer biomarkers in epithelial ovarian cancer: updated guidelines from the European Group on Tumor Markers. *Int J Gynecol Cancer*. 2016;26(1):43-51. doi:10.1097/IGC.0000000000000586

6. Rustin GJ, Vergote I, Eisenhauer E, et al; Gynecological Cancer Intergroup. Definitions for response and progression in ovarian cancer clinical trials incorporating RECIST 1.1 and CA 125 agreed by the Gynecological Cancer Intergroup (GCIG). *Int J Gynecol Cancer*. 2011;21(2):419-423. doi:10.1097/IGC.0b013e3182070f17

7. Stewart LA, Clarke M, Rovers M, et al; PRISMA-IPD Development Group. Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. *JAMA*. 2015;313 (16):1657-1665. doi:10.1001/jama.2015.3656

8. Buyse M, George SL, Evans S, et al. The role of biostatistics in the prevention, detection and treatment of fraud in clinical trials. *Stat Med*. 1999;18(24):3435-3451. doi:10.1002/(SICI)1097-0258(19991230)18:24<3435::AID-SIM365>3.0.CO;2-O

9. Stewart LA, Clarke MJ; Cochrane Working Group. Practical methodology of meta-analyses (overviews) using updated individual patient data. *Stat Med*. 1995;14(19):2057-2079. doi:10.1002/sim.4780141902

10. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539-1558. doi:10.1002/sim.1186

11. Peto R, Davies C, Godwin J, et al; Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. *Lancet*. 2012;379(9814):432-444. doi:10.1016/S0140-6736(11)61625-5

12. 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. doi:10.1093/biostatistics/1.1.49

13. Burzykowski T, Molenberghs G, Buyse M, eds. *The Evaluation of Surrogate Endpoints*. New York, NY: Springer: 2005:163-194.

14. 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. doi:10.1177/0962280207081864

15. Mauguen A, Pignon JP, Burdett S, et al; Surrogate Lung Project Collaborative Group. Surrogate endpoints for overall survival in chemotherapy and radiotherapy trials in operable and locally advanced lung cancer: a re-analysis of meta-analyses of individual patients' data. *Lancet Oncol*. 2013;14(7):619-626. doi:10.1016/S1470-2045(13)70158-X

16. Shi Q, Flowers CR, Hiddemann W, et al. Thirty-month complete response as a surrogate end point in first-line follicular lymphoma therapy: an individual patient-level analysis of multiple randomized trials. *J Clin Oncol*. 2017; 35(5):552-560. doi:10.1200/JCO.2016.70.8651

17. Galimberti S, Devidas M, Lucenti A, et al. Validation of minimal residual disease as surrogate endpoint for event-free survival in childhood acute lymphoblastic leukemia. *J Natl Cancer Inst Cancer Spectr*. 2018;2(4): pky069. doi:10.1093/jncics/pky069

18. 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. doi:10.1002/(SICI)1097-0258(19981230)17:24<2815::AID-SIM110>3.0.CO;2-8

19. Rotolo F, Paoletti X, Michiels S. surrosurv: an R package for the evaluation of failure time surrogate endpoints in individual patient data meta-analyses of randomized clinical trials. *Comput Methods Programs Biomed*. 2018; 155:189-198. doi:10.1016/j.cmpb.2017.12.005

20. Aravantinos G, Fountzilas G, Bamias A, et al; Hellenic Cooperative Oncology Group study. 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(15):2169-2177. doi:10.1016/j.ejca.2008.06.035

21. Banerjee S, Rustin G, Paul J, et al. A multicenter, randomized trial of flat dosing versus intrapatient dose escalation of single-agent carboplatin as first-line chemotherapy for advanced ovarian cancer: an SGCTG (SCOTROC 4) and ANZGOG study on behalf of GCIG. *Ann Oncol*. 2013;24(3):679-687. doi:10.1093/annonc/mds494

22. Fruscio R, Colombo N, Lissoni AA, et al. A phase II randomised clinical trial comparing cisplatin, paclitaxel and ifosfamide with cisplatin, paclitaxel and epirubicin in newly diagnosed advanced epithelial ovarian cancer: long-term survival analysis. *Br J Cancer*. 2008;98(4):720-727. doi:10.1038/sj.bjc.6604231

23. Hirte H, Vergote IB, Jeffrey JR, et al. A phase III randomized trial of BAY 12-9566 (tanomastat) as maintenance therapy in patients with advanced ovarian cancer responsive to primary surgery and paclitaxel/platinum containing chemotherapy: a National Cancer Institute of Canada Clinical Trials Group Study. *Gynecol Oncol*. 2006;102(2):300-308. doi:10.1016/j.ygyno.2005.12.020

24. 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(20):1547-1556. doi:10.1093/jnci/djq362

25. Katsumata N, Yasuda M, Isonishi S, et al; Japanese Gynecologic Oncology Group. Long-term results of dose-dense paclitaxel and carboplatin versus conventional paclitaxel and carboplatin for treatment of advanced epithelial ovarian, fallopian tube, or primary peritoneal cancer (JGOG 3016): a randomised, controlled, open-label trial. *Lancet Oncol*. 2013;14(10):1020-1026. doi:10.1016/S1470-2045(13)70363-2

26. Lindemann K, Christensen RD, Vergote I, et al. First-line treatment of advanced ovarian cancer with paclitaxel/carboplatin with or without epirubicin (TEC versus TC): a gynecologic cancer intergroup study of the NSGO, EORTC GCG and NCIC CTG. *Ann Oncol*. 2012;23(10):2613-2619. doi:10.1093/annonc/mds060

27. Oza AM, Cook AD, Pfisterer J, et al; ICON7 trial investigators. Standard chemotherapy with or without bevacizumab for women with newly diagnosed ovarian cancer (ICON7): overall survival results of a phase 3 randomised trial. *Lancet Oncol*. 2015;16(8):928-936. doi:10.1016/S1470-2045(15)00086-8

28. 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(27):3628-3635. doi:10.1200/JCO.2010.33.8566

29. Pignata S, Scambia G, Katsaros D, et al; Multicentre Italian Trials in Ovarian Cancer (MITO-7); Groupe d'Investigateurs Nationaux pour l'Etude des Cancers Ovariens et du sein (GINECO); Mario Negri Gynecologic Oncology (MaNGO); European Network of Gynaecological Oncological Trial Groups (ENGOT-OV-10); Gynecologic Cancer InterGroup (GCIG) Investigators. Carboplatin plus paclitaxel once a week versus every 3 weeks in patients with advanced ovarian cancer (MITO-7): a randomised, multicentre, open-label, phase 3 trial. *Lancet Oncol*. 2014; 15(4):396-405. doi:10.1016/S1470-2045(14)70049-X

30. Ray-Coquard I, Paraiso D, Guastalla JP, et al. Intensified dose of cyclophosphamide with G-CSF support versus standard dose combined with platinum in first-line treatment of advanced ovarian cancer: a randomised study from the GINECO group. *Br J Cancer*. 2007;97(9):1200-1205. doi:10.1038/sj.bjc.6604026

31. Reyners AK, de Munck L, Erdkamp FL, et al; DoCaCel Study Group. A randomized phase II study investigating the addition of the specific COX-2 inhibitor celecoxib to docetaxel plus carboplatin as first-line chemotherapy for stage IC to IV epithelial ovarian cancer, fallopian tube or primary peritoneal carcinomas: the DoCaCel study. *Ann Oncol*. 2012;23(11):2896-2902. doi:10.1093/annonc/mds107

32. Sugiyama T, Okamoto A, Enomoto T, et al. Randomized phase III trial of irinotecan plus cisplatin compared with paclitaxel plus carboplatin as first-line chemotherapy for ovarian clear cell carcinoma: JGOG3017/GCIG trial. *J Clin Oncol*. 2016;34(24):2881-2887. doi:10.1200/JCO.2016.66.9010

33. van der Burg ME, Onstenk W, Boere IA, et al. Long-term results of a randomised phase III trial of weekly versus three-weekly paclitaxel/platinum induction therapy followed by standard or extended three-weekly paclitaxel/platinum in European patients with advanced epithelial ovarian cancer. *Eur J Cancer*. 2014;50(15):2592-2601. doi:10.1016/j.ejca.2014.07.015

34. Vergote IB, Jimeno A, Joly F, et al. Randomized phase III study of erlotinib versus observation in patients with no evidence of disease progression after first-line platin-based chemotherapy for ovarian carcinoma: a European Organisation for Research and Treatment of Cancer-Gynaecological Cancer Group, and Gynecologic Cancer Intergroup study. *J Clin Oncol*. 2014;32(4):320-326. doi:10.1200/JCO.2013.50.5669

35. Mannel RS, Brady MF, Kohn EC, et al. A randomized phase III trial of IV carboplatin and paclitaxel ×3 courses followed by observation versus weekly maintenance low-dose paclitaxel in patients with early-stage ovarian carcinoma: a Gynecologic Oncology Group Study. *Gynecol Oncol*. 2011;122(1):89-94. doi:10.1016/j.ygyno.2011.03.013

36. Vasey PA, Jayson GC, Gordon A, et al; Scottish Gynaecological Cancer Trials Group. Phase III randomized trial of docetaxel-carboplatin versus paclitaxel-carboplatin as first-line chemotherapy for ovarian carcinoma. *J Natl Cancer Inst*. 2004;96(22):1682-1691. doi:10.1093/jnci/djh323

37. Burger RA, Brady MF, Bookman MA, et al; Gynecologic Oncology Group. Incorporation of bevacizumab in the primary treatment of ovarian cancer. *N Engl J Med*. 2011;365(26):2473-2483. doi:10.1056/NEJMoa1104390

38. du Bois A, Herrstedt J, Hardy-Bessard AC, 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(27):4162-4169. doi:10.1200/JCO.2009.27.4696

39. Bolis G, Scarfone G, Raspagliesi F, et al. Paclitaxel/carboplatin versus topotecan/paclitaxel/carboplatin in patients with FIGO suboptimally resected stage III-IV epithelial ovarian cancer: a multicenter, randomized study. *Eur J Cancer*. 2010;46(16):2905-2912. doi:10.1016/j.ejca.2010.06.124

40. 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(9):1419-1425. doi:10.1200/JCO.2008.19.1684

41. Lhommé C, Joly F, Walker JL, et al; Phase III Study of Valspodar. Phase III study of valspodar (PSC 833) combined with paclitaxel and carboplatin compared with paclitaxel and carboplatin alone in patients with stage IV or suboptimally debulked stage III epithelial ovarian cancer or primary peritoneal cancer. *J Clin Oncol*. 2008;26(16):2674-2682. doi:10.1200/JCO.2007.14.9807

42. du Bois A, Weber B, Rochon J, et al; Arbeitsgemeinschaft Gynaekologische Onkologie; Ovarian Cancer Study Group; Groupe d'Investigateurs Nationaux pour l'Etude des Cancers Ovariens. 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(7):1127-1135. doi:10.1200/JCO.2005.03.2938

43. du Bois A, Floquet A, Kim JW, et al. Incorporation of pazopanib in maintenance therapy of ovarian cancer. *J Clin Oncol*. 2014;32(30):3374-3382. doi:10.1200/JCO.2014.55.7348

44. Herzog TJ, Scambia G, Kim BG, et al. A randomized phase II trial of maintenance therapy with sorafenib in front-line ovarian carcinoma. *Gynecol Oncol*. 2013;130(1):25-30. doi:10.1016/j.ygyno.2013.04.011

45. Vergote IB, Chekerov R, Amant F, et al. Randomized, phase II, placebo-controlled, double-blind study with and without enzastaurin in combination with paclitaxel and carboplatin as first-line treatment followed by maintenance treatment in advanced ovarian cancer. *J Clin Oncol*. 2013;31(25):3127-3132. doi:10.1200/JCO.2012.44.9116

46. Meier W, du Bois A, Rau J, et al. Randomized phase II trial of carboplatin and paclitaxel with or without lonafarnib in first-line treatment of epithelial ovarian cancer stage IIB-IV. *Gynecol Oncol*. 2012;126(2):236-240. doi:10.1016/j.ygyno.2012.04.050

47. Pecorelli S, Favalli G, Gadducci A, et al; After 6 Italian Cooperative Group. Phase III trial of observation versus six courses of paclitaxel in patients with advanced epithelial ovarian cancer in complete response after six courses of paclitaxel/platinum-based chemotherapy: final results of the After-6 protocol 1. *J Clin Oncol*. 2009;27(28):4642-4648. doi:10.1200/JCO.2009.21.9691

48. Pfisterer J, Weber B, Reuss A, et al; AGO-OVAR; GINECO. 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(15):1036-1045. doi:10.1093/jnci/djj296

49. Markman M, Liu PY, Moon J, et al. Impact on survival of 12 versus 3 monthly cycles of paclitaxel (175 mg/m2) administered to patients with advanced ovarian cancer who attained a complete response to primary platinum-paclitaxel: follow-up of a SWOG and GOG phase 3 trial. *Gynecol Oncol*. 2009;114(2):195-198. doi:10.1016/j.ygyno.2009.04.012

50. 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. *Gynecol Oncol*. 2011;123(3):479-485. doi:10.1016/j.ygyno.2011.08.018

Progression-Free Survival as a Surrogate End Point of Overall Survival in Ovarian Cancer

51. du Bois A, Kristensen G, Ray-Coquard I, et al; AGO Study Group led Gynecologic Cancer Intergroup/European Network of Gynaecologic Oncology Trials Groups Intergroup Consortium. Standard first-line chemotherapy with or without nintedanib for advanced ovarian cancer (AGO-OVAR 12): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet Oncol*. 2016;17(1):78-89. doi:10.1016/S1470-2045(15)00366-6

52. Sabbatini P, Harter P, Scambia G, et al. Abagovomab as maintenance therapy in patients with epithelial ovarian cancer: a phase III trial of the AGO OVAR, COGI, GINECO, and GEICO—the MIMOSA study. *J Clin Oncol*. 2013;31(12): 1554-1561. doi:10.1200/JCO.2012.46.4057

53. Mouratidou D, Gennatas C, Michalaki V, et al. A phase III randomized study comparing paclitaxel and cisplatin versus cyclophosphamide and cisplatin in patients with advanced ovarian cancer. *Anticancer Res*. 2007;27(1B): 681-685. http://ar.iiarjournals.org/content/27/1B/681.long. Accessed September 30, 2019.

54. Hainsworth JD, Thompson DS, Bismayer JA, et al. Paclitaxel/carboplatin with or without sorafenib in the first-line treatment of patients with stage III/IV epithelial ovarian cancer: a randomized phase II study of the Sarah Cannon Research Institute. *Cancer Med*. 2015;4(5):673-681. doi:10.1002/cam4.376

55. De Placido S, Scambia G, Di Vagno G, et al. Topotecan compared with no therapy after response to surgery and carboplatin/paclitaxel in patients with ovarian cancer: Multicenter Italian Trials in Ovarian Cancer (MITO-1) randomized study. *J Clin Oncol*. 2004;22(13):2635-2642. doi:10.1200/JCO.2004.09.088

56. Nicoletto MO, Tumolo S, Sorio R, et al; Goccne Group (Gruppo Oncologico Cooperativo Clinico Nord-est), Padua, Italy. Long-term survival in a randomized study of nonplatinum therapy versus platinum in advanced epithelial ovarian cancer. *Int J Gynecol Cancer*. 2007;17(5):986-992. doi:10.1111/j.1525-1438.2007.00862.x

57. Sturgeon CM, Duffy MJ, Stenman UH, et al; National Academy of Clinical Biochemistry. National Academy of Clinical Biochemistry laboratory medicine practice guidelines for use of tumor markers in testicular, prostate, colorectal, breast, and ovarian cancers. *Clin Chem*. 2008;54(12):e11-e79. doi:10.1373/clinchem.2008.105601

58. Shimokawa M, Ohki M, Kaku T. Correlation of progression-free and post-progression survival with overall survival in phase III trials of first-line chemotherapy for advanced epithelial ovarian cancer. *Eur J Gynaecol Oncol*. 2015;36(4):370-375. doi:10.12892/ejgo2643.2015

59. Colloca G, Venturino A. Trial-level analysis of progression-free survival and response rate as end points of trials of first-line chemotherapy in advanced ovarian cancer. *Med Oncol*. 2017;34(5):87. doi:10.1007/s12032-017-0939-9

60. González-Martín A, Pothuri B, Vergote I, et al; PRIMA/ENGOT-OV26/GOG-3012 Investigators. Niraparib in patients with newly diagnosed advanced ovarian cancer [published online September 28, 2019]. *N Engl J Med*. doi:10.1056/NEJMoa1910962

61. Food and Drug Administration. Clinical trial endpoints for the approval of cancer drugs and biologics guidance for industry. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-trial-endpoints-approval-cancer-drugs-and-biologics. Published December 2018. Accessed September 30, 2019.

62. European Medicines Agency. Guideline on the evaluation of anticancer medicinal products in man. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-evaluation-anticancer-medicinal-products-man-revision-5_en.pdf. Published September 22, 2017. Accessed September 30, 2019.

SUPPLEMENT.
eFigure 1. PRISMA Flow Diagram
eFigure 2. Overall Survival (A) and Progression Free Survival (B)
eFigure 3. Overall Survival (OS) in Each Trial According to the Year of Trial's Initiation
eFigure 4. Overall and Trial by Trial Treatment Effect (HR) on Overall and Progression-Free Survival
eFigure 5. Re-estimating the Relationship Between the Hazard Ratio (HR) on OS and HR(PFS) by Leaving One Trial Out at a Time
eTable 1. Risk of Bias Summary: Authors' Judgments About Each Risk of Bias Item for Each Included Study
eTable 2. Patients' Characteristics
eTable 3. Observed and Predicted Treatment Effect on Overall Survival (OS HR), Based on the Observed Treatment Effect on Progression-Free Survival (PFS HR)