



Published in final edited form as:

*Gynecol Oncol.* 2013 October ; 131(1): 21–26. doi:10.1016/j.ygyno.2013.07.100.

## Independent radiologic review of the Gynecologic Oncology Group Study 0218, a phase III trial of bevacizumab in the primary treatment of advanced epithelial ovarian, primary peritoneal, or fallopian tube cancer★

Robert A. Burger<sup>a,\*</sup>, Mark F. Brady<sup>b</sup>, Joon Rhee<sup>c</sup>, Mika A. Sovak<sup>c</sup>, George Kong<sup>c</sup>, Hoa P. Nguyen<sup>c</sup>, and Michael A. Bookman<sup>d</sup>

<sup>a</sup>Fox Chase Cancer Center, Philadelphia, PA

<sup>b</sup>Gynecologic Oncology Cohort Statistical and Data Center, Roswell Park Cancer Institute, Buffalo, NY

<sup>c</sup>Genentech, Inc., South San Francisco, CA

<sup>d</sup>The University of Arizona Cancer Center, Tucson, AZ

### Abstract

**Objectives**—Gynecologic Oncology Group Study 0218 (GOG-0218), a phase III, placebo-controlled trial in newly diagnosed stage III/IV ovarian cancer (OC), demonstrated a benefit in investigator (INV)-assessed progression-free survival (PFS) with bevacizumab (BEV) administered with and following carboplatin/paclitaxel (CP) for up to 15 months vs. CP alone. To determine the reliability of Response Evaluation Criteria in Solid Tumors (RECIST) in assessing disease progression (PD) in GOG-0218, an independent review of radiologic and clinical data (IRC) was conducted.

**Methods**—Blinded reviews followed RECIST 1.0 in accordance with the study protocol; PFS was analyzed in the intent-to-treat population.

**Results**—CP + BEV BEV achieved a significant PFS improvement in both assessments. Hazard ratios for PFS (IRC: 0.623; 95% confidence interval [CI]: 0.503–0.772;  $p < 0.0001$  vs. INV: 0.624; 95% CI: 0.520–0.749;  $p < 0.0001$ ) and the improvement in median PFS (IRC: 19.1 and 13.1 months vs. INV: 18.2 and 12 months) were similar between IRC and INV assessments. There was high concordance between IRC- and INV-determined PD status (77%) and date (73%). Subgroup

★Presented in part at the 47<sup>th</sup> Annual Meeting of the American Society of Clinical Oncology, Chicago, Illinois, June 3–7, 2011.

© 2013 Elsevier Inc. All rights reserved.

\*Corresponding author at: Fox Chase Cancer Center, 333 Cottman Ave, Philadelphia, PA 19111-2497, USA, Tel: 215-728-3150; Fax: 215-728-2773. robert.a.burger@fccc.edu (R. Burger).

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

### Conflict of interest statement

Robert A. Burger has served on scientific advisory boards for Genentech/Roche. Mark F. Brady declares no conflicts of interest. Joon Rhee, Mika A. Sovak, George Kong, and Hoa P. Nguyen are employees of Genentech, Inc. Michael A. Bookman has served on independent data monitoring committees (DMC) for phase 3 trials in ovarian cancer sponsored by Genentech/Roche (not directly related to the current study) with compensation for time and effort. In the last 5 years, he has also participated in ovarian cancer advisory boards sponsored by Genentech/Roche and received travel support and limited honoraria (when not serving on DMC).

analyses were consistent with the primary IRC findings. Early and late discontinuation discordance measures showed no evidence of INV bias.

**Conclusion**—IRC analysis confirmed a significant PFS improvement with CP + BEV + BEV vs. CP alone. Concordance was not influenced by extent of residual disease after cytoreductive surgery or initial stage. The IRC size, high participation rate, and strong concordance between IRC and INV assessments suggest that RECIST can be applied objectively in OC studies.

## Introduction

Over the last decade, regulatory authorities have recommended that phase III registration studies in oncology that use progression-free survival (PFS) as their primary endpoint incorporate centralized radiology review to determine reliability for efficacy assessments and to control for potential investigator (INV) bias.[1,2] A recent meta-analysis of more than 27 phase III independent review committee (IRC) trials (independent reviews of blinded radiologic and clinical data)[3] has shown high consistency between INV and IRC PFS assessments with little evidence of INV bias. The reliability of radiologic assessment of ovarian cancer has been a source of concern because ovarian cancers are characterized by diffuse peritoneal implantation and the prevalence of nontarget lesions, such as ascites.

The Gynecologic Oncology Group Study 0218 (GOG-0218) is a phase III, randomized, three-arm, double-blind, placebo-controlled trial in females with newly diagnosed, International Federation of Gynecology and Obstetrics stage III (with any gross [macroscopic or palpable] residual disease at the conclusion of initial surgery) or stage IV epithelial ovarian, primary peritoneal, or fallopian tube cancer.[4] This study was designed to test the clinical benefit of adding both short- and long-duration bevacizumab (BEV) to standard carboplatin and paclitaxel (CP) chemotherapy. In a prespecified PFS analysis censoring for progression based on cancer antigen 125 (CA-125), as directed by regulatory agencies, the median PFS per the Response Evaluation Criteria in Solid Tumors (RECIST) 1.0 guidelines, was 12.0 months for control vs. 18.2 months for the treatment cohort assigned to CP with BEV followed by BEV extended for a maximum of 16 additional cycles (hazard ratio [HR] 0.624; Table 1).[4] To address regulatory concerns about the reliability of imaging for tumor assessment in ovarian cancer, an independent and blinded review of radiology and oncology data was performed. Here we present results of the PFS analysis of GOG-0218 using the IRC-assessed data, as well as detailed metrics on concordance and discordance measures.

## Patients and methods

### Patient eligibility and trial design

The GOG-0218 study design (Fig. 1) and major inclusion criteria have been reported previously.[4] GOG-0218 was a double-blind, three-arm, placebo-controlled, phase III trial in which all cancer therapeutics were administered on day 1 of 21-day cycles, all participants were assigned to receive 6 cycles of CP, and BEV or placebo was initiated during the second cycle of CP to reduce the risk of wound-healing complications. Extended therapy consisted of BEV or placebo for 16 cycles. The control regimen (regimen 1) consisted of CP plus concurrent placebo, followed by placebo; regimen 2 consisted of CP plus concurrent BEV 15 mg/kg followed by placebo (CP + BEV); and regimen 3 consisted of CP plus concurrent BEV followed by continued BEV (CP + BEV + BEV). Treatment was discontinued upon disease progression, unacceptable toxicity, completion of 22 cycles or withdrawal—whichever came first. GOG-0218 was approved by institutional review boards at all participating sites, and informed consent was obtained from all participants (ClinicalTrials.gov, NCT00262847).

## Assessment of response and progression

The IRC used radiologic and clinical evidence (e.g. clinical symptoms) to determine disease progression in accordance with a protocol-specific IRC charter, following RECIST 1.0. Computed tomography or magnetic resonance imaging of at least the abdomen and pelvis was performed at baseline, after treatment cycles 3, 6, 10, 14, 18, and 22, and then every 3 months for 2 years, then every 6 months for 3 years and then annually as described in the study protocol. Imaging assessments were discontinued once disease progression was confirmed, except when disease progression was defined by rising serum CA-125 levels alone, according to Gynecologic Cancer InterGroup (GCIg) criteria.[5] When disease progression was defined by CA-125 alone, INVs were asked to obtain scans within 2 weeks of documented progression. As the IRC was implemented after the trial had been open to enrollment for 2 years, scan collection was both retrospective and prospective. Relevant radiographic images and clinical data until INV-determined disease progression were provided to the IRC for review. Data for patients who progressed solely on the basis of CA-125 criteria were censored at the last tumor assessment for which the patient was known to be progression-free, which was the last tumor assessment available for the IRC's review.

To participate in the IRC analysis, patients must have been on study for at least 9 weeks and must have had baseline images and at least 1 subsequent protocol-required set of images submitted to the IRC.

In response to regulatory concerns about the specificity of the CA-125 biomarker as a definitive indication of disease progression, progression events based on CA-125 criteria were not considered as events in the primary analysis of PFS. This differs from the PFS analysis as presented in the primary study publication,[4] in which CA-125– based progression events were counted. For the IRC analysis, CA-125 marker data were not sent to the IRC and were not a part of that review.

Imaging-based evaluation by the IRC was performed by 2 radiologists. In cases of disagreement, a third radiologist served as adjudicator to determine which of the 2 radiologic assessments would be used for analysis. An independent oncologist then reviewed specific clinical data in conjunction with the final radiologic evaluation to make a final determination of response and progression status. Clinical data provided to the independent oncologist included information on lesions found by physical examination, documented symptomatic deterioration, and ovarian cancer procedures as described in case report forms. All reviews were performed in a sequential blinded fashion.

## Analysis methods

PFS was assessed as an intent-to-treat analysis of all enrolled patients. In these analyses, data for patients who initiated nonprotocol antineoplastic therapy prior to disease progression were censored at the last tumor assessment prior to the initiation of therapy.

The primary analysis of PFS as determined by the IRC was a stratified log-rank test that compared each experimental cohort to the control cohort. Stratification factors were disease stage (macroscopic stage III with no residual tumor implants exceeding 1 cm in maximal diameter; stage III with any residual tumor implant exceeding 1 cm in any dimension; stage IV) and GOG performance status (0, 1, or 2). Kaplan-Meier methodology was used to estimate median times in the treatment cohorts. A stratified Cox regression model was used to estimate the stratified HRs.

The methodology for discordance measures was based on a meta-analysis of 27 randomized phase III trials in solid tumors with blinded IRC reviews conducted by the Pharmaceutical Research and Manufacturers Association (PhRMA)-sponsored PFS Independent Review

Working Group to evaluate the potential of INV bias in determining PFS.[3] This meta-analysis included a detailed analysis of discordance data from 12 of these trials and defined new discordance measures to evaluate potential INV bias. Unlike the standard metrics for discordance, in which the less efficacious regimen will have more discordance because of more progressive events, these new measures are not confounded by efficacy and are not dependent on data maturity. The meta-analysis showed that there was a high degree of association between INV and IRC estimates of the treatment effect ( $R=0.947$ ; 95% confidence interval [CI]: 0.88-0.97) for studies included in the analysis. Furthermore, discordance rates were similar among the studies between the treatment cohorts across all measures, with an average differential discordance close to 0.

Discordance measures of early discrepancy rate (EDR), late discrepancy rate (LDR), proportion of disagreement on censoring status, and proportion of disagreement on timing and occurrence of progressive disease were calculated as described in Amit et al.[3] The EDR represents the positive predictive value of INV assessment and quantifies the frequency with which the INV declares progression early relative to the IRC for each cohort as a proportion of the total number of INV-assessed disease progressions. The LDR quantifies the frequency that the INV declares progression later relative to the IRC as a proportion of the total number of discrepancies within the cohort. The EDR and LDR for each cohort can be compared. A negative differential discordance for EDR and/or a positive differential discordance for LDR suggest bias by the INV in favor of the experimental regimen. Moderate differences in differential discordance rate (i.e. the difference in discordance rate between treatment regimens) would reduce suspicion of systematic directional evaluation bias.

## Results

Demographic and tumor characteristics were balanced across all cohorts.[4] On the basis of IRC eligibility criteria, 91% of patients participated in the IRC analysis; 97.2% of these participants had all protocol-required scans submitted to the IRC.

### PFS analyses

The IRC determined that 558 PFS events occurred: 203 in cohort 1, 206 in cohort 2, and 149 in cohort 3. BEV added to and continued beyond chemotherapy (regimen 3) resulted in a statistically significant improvement in PFS compared with the control regimen (see Table 1; Fig. 2). Median PFS was 19.1 months for regimen 3 compared with 13.1 months for the control regimen (regimen 1), with a stratified HR of 0.623 (95% CI: 0.503-0.772). PFS for regimen 2 was not significantly different from PFS for the control regimen (Supplementary Fig. A1).

In the INV-determined PFS analysis censored for CA-125 progression, there were 742 PFS events: 277 events in cohort 1, 258 events in cohort 2 and 207 events in cohort 3. The stratified PFS HR for cohort 3 vs. cohort 1 was 0.624 (95% CI: 0.520–0.749) with median PFS of 18.2 months vs. 12 months, similar to the IRC analysis. As with the IRC analysis, PFS for regimen 2 was not significantly different from PFS for the control regimen (data not shown).

On the basis of the IRC analysis, the PFS benefit seen for regimen 3 was consistent across all subgroups (Supplementary Fig. A2) and was similar to the comparable subgroup analysis based on the GOG review.[4]

## Concordance

Concordance was observed between the IRC- and INV-determined disease progression status (77%) and progression date (73%) (Tables 2 and 3; Supplementary Table A1). Total concordance on progression status (agreement between INV and IRC on progression or no progression) was balanced across all 3 cohorts at 73.9% for cohort 1, 78.8% for cohort 2, and 77.9% for cohort 3. Disease progression status concordance among stage subgroups was also similar and balanced across cohorts (Supplementary Tables A2, A3, A4), suggesting that variability in radiologic assessment was not dependent on the extent of residual disease following cytoreductive surgery or the initial stage of disease. As shown in Table 3, there were a higher number of cases with exact matching progression date in cohort 3 (84.6% cohort 3 vs. 68% cohort 1 and 71.7% cohort 2), but overall concordance on progression date was consistent across all 3 cohorts. Importantly, in 86.6% of cases in which the progression date did not exactly match, the time differential for disease progression was within a 12-week period (or 1 scan interval).

To understand how PFS concordance from GOG-0218 compares with large phase III IRC trials in other solid tumor types and to explore further whether there was evidence of INV bias in PFS assessment in the GOG-0218 study, additional analyses were conducted using the methodology described in Amit et al.[3] The discordance measures for GOG-0218 are shown in Table 4. The small EDR and LDR differentials reflect the high level of concordance observed between the INV and the IRC assessments, with no evidence of INV bias.

## IRC radiologist reviewers and adjudication of scan interpretation

Fifteen radiologists, with a median of 9 years (range, 2-33) of experience at study start, participated in the IRC. Per the IRC charter, disagreement between readers on disease progression status required adjudication by a third radiologist. This occurred in 30.2% of the scans read, with similar rates observed in each treatment arm.

## Discussion

BEV is the first molecular-targeted and antiangiogenic agent to demonstrate benefit as frontline treatment of advanced ovarian (epithelial ovarian, primary peritoneal, or fallopian tube) cancer. GOG-0218 met its primary objective by showing that PFS for the extended BEV regimen (CP + BEV + BEV) was superior to that for the control regimen (CP alone). It also showed that the experimental regimen was generally well tolerated with a safety profile (including gastrointestinal perforation) similar to those of previous BEV studies. The IRC analyses presented here show striking similarity to the INV analysis and support the finding that BEV added to CP prolongs the PFS for women with newly diagnosed, advanced ovarian cancer.

Independent review of imaging data has become more common as regulatory authorities [1,2] have advocated for its use in trials for which imaging is used to define primary registration endpoints. While review of radiologic data independent of other clinically relevant historical, physical, and laboratory data may have its limitations in precision of disease assessment, the advantage of this method in evaluating participants in a trial in a blinded fashion is the elimination of investigator bias. Independent radiology review has been used successfully to confirm PFS or time to progression in phase III studies that supported the approval of new therapies in many solid tumors, including breast,[6] renal cell carcinoma,[7] and gastrointestinal stromal tumor.[8] Despite these successes, the reliability of radiologic imaging for tumor assessment in cancers with intraperitoneal spread patterns, such as ovarian cancer, has been questioned.

Owing to concerns about the reliability of image assessment in ovarian cancer, the US Food and Drug Administration (FDA) requested that an IRC be implemented for GOG-0218 and included as part of the registration application. Despite the difficulties associated with implementing an IRC for such a large study after the trial began enrollment, the participation rate was 91%, and overall results of the IRC analysis closely mirrored those from the primary and subgroup analyses, as determined by the GOG investigators. It is possible that if imaging data had been transferred for independent review in a prospective fashion at the time each imaging study was performed for the entire study population, the participation rate could have approximated 100%, and such methodology could be helpful in future studies. The most obvious limitation of the current study is the incorporation of serum CA-125 level in disease assessment within the clinical trial itself, which led to discontinuation of treatment and regularly scheduled imaging in participants meeting GCIG criteria for progression.

Recently the FDA held an Oncologic Drugs Advisory Committee meeting on the evaluation of radiologic review of PFS in nonhematologic malignancies [9] to discuss the need for IRCs for pivotal phase III studies. The FDA also endorsed the PhRMA methodology as a means for validating PFS results in future trials. Using this same methodology, discordance measures from the GOG-0218 study were compared with the PhRMA meta-analysis for which the proportion of disagreement on censoring and disease progression timing or occurrence was comparable to (and in some cases lower than) the studies in the meta-analysis; there was no evidence of INV bias. Importantly, the discordance metrics for GOG-0218 are consistent with the trials in the PhRMA meta-analysis and minimize concerns of systematic INV bias in ovarian cancer. In addition, the proportional disagreement due to censoring and proportional disagreement on timing or occurrence of disease progression for GOG-0218 was notably lower in the experimental cohorts than those seen in trials in the meta-analysis, showing that the overall level of discordance observed in GOG-0218 was actually lower than that observed in the trials from the meta-analysis. Overall, the discordance data suggest strongly that the PFS results, as determined by IRC and INV in the GOG-0218 trial, were as consistent as the results seen in other large phase III trials in other solid tumor indications.

GOG-0218 is by far the largest study in ovarian cancer to use IRC review for PFS analysis. Because of the size of the study, the high participation rate, and high concordance achieved in this study, the IRC analysis provides validation to the increased PFS seen with the addition of both extended bevacizumab to standard CP chemotherapy. In addition, the level of discordance between INVs and the IRC was consistent with that observed in other large phase III trials in other solid tumors and strongly suggests that radiologic assessment can be used reliably to assess progression events in ovarian cancers.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

Supported by the National Cancer Institute and Genentech, Inc. Support for third-party copyediting assistance for this manuscript, furnished by Melissa Hernandez-Warren, was provided by Genentech, Inc.

**Research support:** This study was supported by Genentech, Inc.

## References

1. United States Food and Drug Administration. United States Food and Drug Administration Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics.

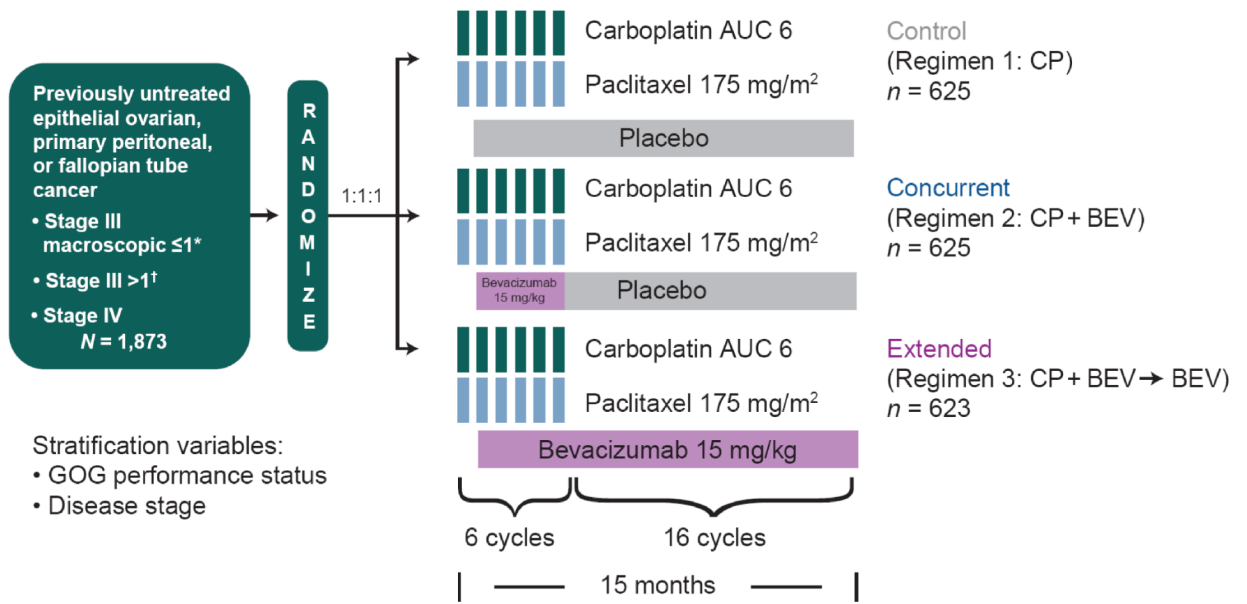
- US Department of Health and Human Services; Rockville, MD: 2007. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071590.pdf>
2. Committee for Medicinal Products for Human Use. European Medicines Agency Committee for Medicinal Products for Human Use (CHMP) Guideline for the Evaluation of Anticancer Medicinal Products in Man. London, UK: 2006. [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2009/12/WC500017748.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/12/WC500017748.pdf)
  3. Amit O, Mannino F, Stone AM, et al. Blinded independent central review of progression in cancer clinical trials: Results from a meta-analysis. *Eur J Cancer*. 2011; 47:1772–8. [PubMed: 21429737]
  4. Burger RA, Brady MF, Bookman MA, et al. Incorporation of bevacizumab in the primary treatment of ovarian cancer. *New Engl J Med*. 2011; 365:2473–83. [PubMed: 22204724]
  5. Stuart GC, Kitchener H, Bacon M, et al. 2010 Gynecologic Cancer InterGroup (GCIg) consensus statement on clinical trials in ovarian cancer: Report from the Fourth Ovarian Cancer Consensus Conference. *Int J Gynecol Cancer*. 2011; 21:750–5. [PubMed: 21543936]
  6. Geyer CE, Forster J, Lindquist D, et al. Lapatinib plus capecitabine for HER2-positive advanced breast cancer. *N Engl J Med*. 2006; 355:2733–43. [PubMed: 17192538]
  7. Kane RC, Farrell AT, Saber H, et al. Sorafenib for the treatment of advanced renal cell carcinoma. *Clin Cancer Res*. 2006; 12:7271–8. [PubMed: 17189398]
  8. Demetri GD, van Oosterom AT, Garrett CR, et al. Efficacy and safety of sunitinib in patients with advanced gastrointestinal stromal tumour after failure of imatinib: A randomised controlled trial. *Lancet*. 2006; 368:1329–38. [PubMed: 17046465]
  9. US Food and Drug Administration: FDA Briefing Document Oncologic Drugs Advisory Committee. Evaluation of Radiologic Review of Progression-Free Survival in Non-hematologic Malignancies. 2012. <http://www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/drugs/oncologicdrugsadvisorycommittee/ucm311141.pdf>

### Highlights

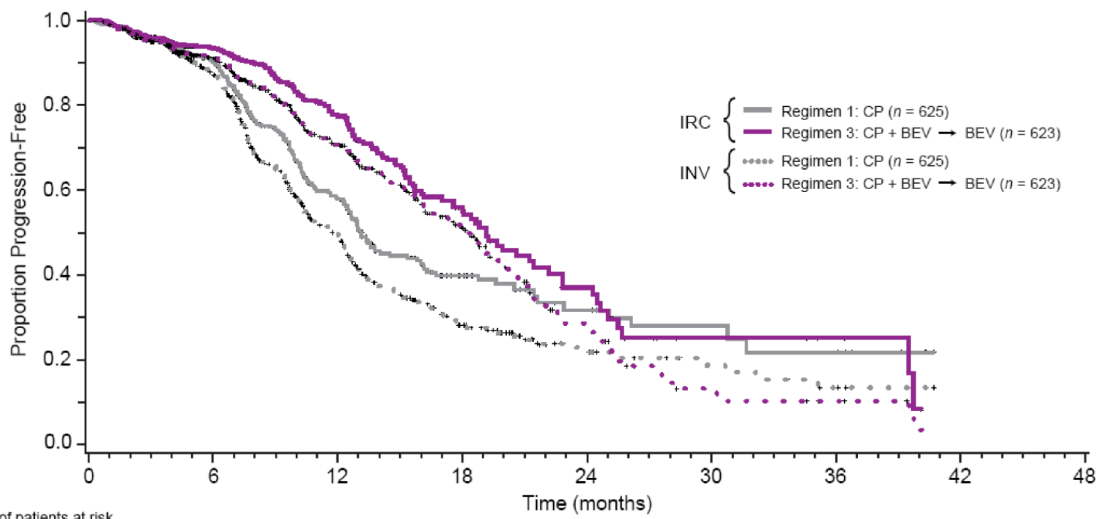
Burger RA et al. Independent radiologic review of the Gynecologic Oncology Group Study 0218, a phase III trial of bevacizumab in the primary treatment of advanced epithelial ovarian, primary peritoneal, or fallopian tube cancer

- Independent review (IRC) of PD in GOG218 confirmed a significant PFS improvement for CP+BEV vs. BEV vs. CP alone.
- High concordance of investigator and IRC assessments of PD status and PD date show no investigator bias.
- The IRC's size and concordance suggest that RECIST can be reliably used in OC trials.





**Fig. 1.** GOG-0218 study design. AUC, area under the curve; BEV, bevacizumab; CP, carboplatin plus paclitaxel; GOG, Gynecologic Oncology Group. \*No residual tumor implants exceeding 1 cm in maximal diameter. †Any residual tumor implant exceeding 1 cm in any dimension.



		Number of patients at risk							
		0	6	12	18	24	30	36	42
IRC	Regimen 1: CP	625	377	145	48	19	10	7	0
	Regimen 3: CP + BEV → BEV	623	401	180	62	22	7	5	0
INV	Regimen 1: CP	625	391	152	52	21	11	6	0
	Regimen 3: CP + BEV → BEV	623	419	193	74	27	9	5	0

**Fig. 2.** Kaplan-Meier estimates of PFS for regimen 3 vs. regimen 1 as determined by the IRC vs. INV. BEV, bevacizumab; CP, carboplatin plus paclitaxel; INV, investigator; IRC, independent review committee; PFS, progression-free survival.

**Table 1**

PFS as determined by the IRC and the INV\*

	IRC-Assessed PFS <sup>†</sup>			INV-Assessed PFS <sup>‡</sup>		
	Regimen 1 CP (n = 625)	Regimen 2 CP + BEV (n = 625)	Regimen 3 CP + BEV (n = 623)	Regimen 1 CP (n = 625)	Regimen 2 CP + BEV (n = 625)	Regimen 3 CP + BEV (n = 623)
Patients with events, no.	203	206	149	277	258	207
Median, months <sup>§</sup>	13.1	13.2	19.1	12	12.8	18.2
HR, stratified <sup>¶</sup>		0.927	0.623		0.812	0.624
95% CI		0.762-1.127	0.503-0.772		0.684-0.963	0.520-0.749
1-sided log-rank p value		0.2220	<0.0001		0.0084	<0.0001

BEV, bevacizumab; CA-125, cancer antigen 125; CI, confidence interval; CP, carboplatin plus paclitaxel; HR, hazard ratio; INV, investigator; IRC, independent review committee; PFS, progression-free survival.

\* PFS results presented in Burger et al.<sup>4</sup>

<sup>†</sup> Censored for nonprotocol therapy.

<sup>‡</sup> Censored for nonprotocol therapy and CA-125 progression.

<sup>§</sup> Kaplan-Meier estimates.

<sup>¶</sup> Relative to regimen 1 (CP).

**Table 2**

INV- vs. IRC-determined disease progression status concordance\*

	<b>Regimen 1 CP (n = 625)</b>	<b>Regimen 2 CP + BEV (n = 625)</b>	<b>Regimen 3 CP + BEV + BEV (n = 623)</b>	<b>All Patients (N = 1873)</b>
Total concordance, n (%)	462 (73.9)	492 (78.7)	485 (77.8)	1439 (76.8)

BEV, bevacizumab; CP, carboplatin plus paclitaxel; INV, investigator IRC, independent review committee.

\* Concordance is calculated by adding cases in which INV and IRC both agree on progression status of progression or no progression (see Table A1 for details on concordance of PD status for INV vs. IRC).

Table 3

INV - vs. IRC-determined disease progression date concordance

	Regimen 1 CP (n = 169)		Regimen 2 CP + BEV (n = 166)		Regimen 3 CP + BEV + BEV (n = 104)		All Patients (N = 439)	
	No.	%	No.	%	No.	%	No.	%
Directional difference, n (%)								
PDD <sub>IRC</sub> = PDD <sub>INV</sub>	115	68.0	119	71.7	88	84.6	322	73.3
INV declares PD later than IRC	42	24.9	34	20.5	11	10.6	87	19.8
INV declares PD earlier than IRC								
Absolute PDD <sub>IRC</sub> - PDD <sub>INV</sub> difference, n (%)	12	7.1	13	7.8	5	4.8	30	6.8
12 weeks	141	83.4	147	88.6	92	88.5	380	86.6
>12 weeks	28	16.6	19	11.4	12	11.5	59	13.4

BEV, bevacizumab; CP, carboplatin plus paclitaxel; INV, investigator; IRC, Independent Review Committee; PDD, progressive disease; PD, progressive disease date.

**Table 4**

Discordance and differentials for INV vs. IRC PFS analysis

	GOG-0218			PhRMA Meta-Analysis*	
	CP (n = 625)	CP + BEV (n = 625)	CP + BEV (n = 623)	Control (N = 12)	Experimental (N = 12)
Proportional disagreement on censoring status	0.36	0.29	0.27	0.36	0.37
Proportional disagreement on timing or occurrence of disease progression	0.35	0.29	0.25	0.51	0.52
Early discrepancy rate	0.45	0.41	0.52	0.41	0.45
Late discrepancy rate	0.40	0.40	0.31	0.33	0.33

BEV, bevacizumab; CP, carboplatin plus paclitaxel; INV, investigator; IRC, Independent Review Committee; PFS, progression-free survival; PhRMA, Pharmaceutical Research and Manufacturers Association.

\* Discordance data from 12 blinded studies with central independent review as described in Amit et al.<sup>3</sup>