the systemic therapy as well as surgery will be allocated with equal probability within blocks of treatments. The treatment assignment will remained concealed until after the patient has been successfully registered onto the study. All interim and final reports will include an accounting of all patients registered, regardless of compliance to the assigned treatment or eligibility to the study.**(08/29/11)**

11.2 Measures of Efficacy and safety

The principle observations for evaluating the therapeutic effects of treatment are:

11.21 Primary efficacy endpoint: Overall survival

11.22 Secondary efficacy endpoint: Progression-free survival (PFS)

11.23 Safety endpoints: frequency and severity of adverse events (Common Terminology Criteria for Adverse Events (CTCAE) – version 3.0).

11.3 Treatment efficacy

Overall type I error: This study includes two primary objectives. The first objective is to determine whether the addition of bevacizumab increases overall survival relative to carboplatin and paclitaxel alone. The second objective is to determine whether surgical cytoreduction increases overall survival. The study design will allocate 2.5% (one-tail) type I error to *each* of these two objectives accounting for interim analyses.

Expected median survival on the standard treatment: Previous studies indicate that the expected death rate for platinum-sensitive patients treated with a platinum-taxane regimen who do not undergo debulking surgery is approximately 0.378 $year^{-1}$ (median survival time = 22 months).

Accrual target for evaluating the efficacy of systemic therapy
The targeted accrual for this component of the study is 660 patients. It is anticipated that 240 eligible patients per year can be enrolled from GOG treatment centers. Therefore, the expected time to accrue the targeted sample size is 2.75 years. An additional 1.5 year post-accrual follow-up period is anticipated.

Statistical power for evaluating the efficacy of biologic therapy: The first objective of this study is to determine whether bevacizumab (CTB) reduces the overall death rate when compared to the standard treatment (CT). The null hypotheses: H_{01}: $\Delta_{01} = \lambda_{CTB} / \lambda_{CT} \geq 1$will be assessed, where λ.is the death rate for the indicated treatment. The treatment regimens will be compared with a logrank procedure which includes ***all*** of the patients categorized by their randomly assigned treatment. This comparison will not include those patients who were enrolled after August 28, 2011 and hence selected their systemic treatment. The

type I error for this comparison will be limited to 2.5% (one-tail) accounting for the planned interim analyses. The logrank test will be stratified by the secondary surgical debulking status (randomized to undergo cytoreduction, vs randomized to not undergo secondary cytoreduction vs not a candidate or did not consent to secondary surgical cytoreduction) and the duration of treatment free-interval prior to enrolling onto this study (6-12 months vs > 12 months).**(08/29/11)(12/19/11)**

If the bevacizumab-containing regimen reduces the overall death rate 25% relative to the control regimen, then this is considered clinically significant. Assuming proportional hazards, this effect size is comparable to increasing the expected proportion surviving at least 22 months (median) 9.5% (50% vs 59.5%). In order to provide an 81% chance of detecting this effect size, the study will be considered sufficiently mature to permit a final analysis of the systemic regimens when there are at least 214 deaths (214/330=0.65) reported among those patients assigned to the standard regimen (CT). If the alternative hypothesis is true then the expected total number of deaths at the time of the final analyses is 394. The power curve for comparing the biologic-containing regimens to the control regimen is displayed in figure 1.

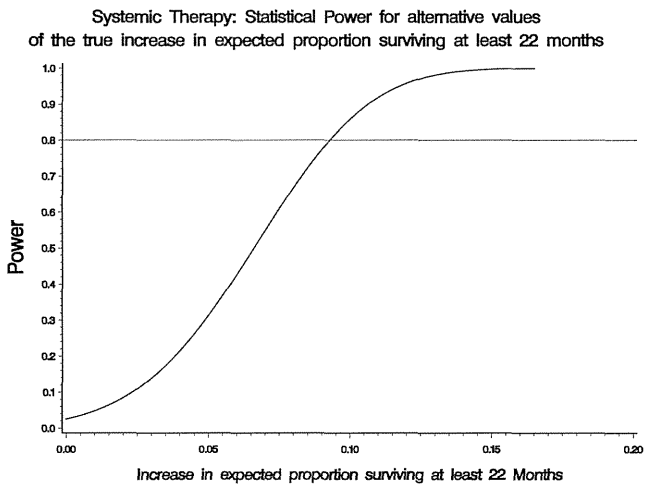

Figure 1.

Statistical Power- evaluating the efficacy of surgical cytoreduction: In order to assess the hypothesis that cytoreductive surgery does not improve overall survival (H_{02}: $\Delta_{02} = \lambda_{surgery} / \lambda_{No\ surgery} \geq 1$), only those patients who were considered candidates for surgery and consented to have their surgical treatment determined by randomization will be included in this analysis. In order to evaluate the efficacy of surgical cytoreduction, patients will be grouped by their randomly assigned surgical treatment regardless of compliance or the degree of actual tumor debulking. This hypothesis will be assessed with a logrank test stratified by their chemotherapeutic/biologic treatment (CT vs CTB vs. CG vs CGB) and the duration of the treatment-free interval prior to enrolling onto this study (6-12

months vs > 12 months). The type I error will be limited to 2.5% for a one-tail test including the error spent due to interim analyses. **(08/29/11)(10/01/12)**

This study will be considered sufficiently mature for an analysis of the surgical cytoreduction hypothesis, H_{02}, when there are at least 250 deaths reported among those enrolled (and randomized) onto the surgical component of this study. This target size provides 80% power, if surgical cytoreduction truly decreases the death rate 30%. This treatment effect size is comparable to increasing the percent surviving 22 months or longer 11.5% (50% to 61.5%). The power curve for this study objective is summarized in Figure 2. The anticipated total accrual for this component of the study is 360 patients.**(08/29/11)**

An addendum to the statistical considerations following the amendment to extend recruitment to the surgical component of the study. ***(12/19/2011)***

As of Nov-1-2011, there were 114 patients enrolled onto the surgical component of this study and 17 of these patients had died. The planned total number of patients to be enrolled onto the surgical component of this study is 360 patients. There were 35 patient enrolled during 2010, and 34 patients are projected to be enrolled during 2011. Therefore, assuming the future accrual is 35 patients per year, this study is expected to complete its targeted accrual in 7 years (Nov-2018). Patients enrolled after Aug-28-2011 will have their cytoreductive surgery determined by randomization. All of these patients will receive a standard carboplatin-paclitaxel regimen and permitted to choose whether to have bevacizumab supplement their treatment.

The data currently available from GOG-0213, indicates that the marginal hazard of death is approximately 0.021 $month^{-1}$. Assuming a constant hazard, the expected number of deaths when the accrual is completed among the 114 patients, who are already enrolled onto the surgical component of this study is 97.4. Assuming a similar death rate for all future patients, Simpson's rule (Schoenfeld, Biometrics 1983) can be used to estimate the number of patients among those who will be enrolled over the next 7 years and will have died when the accrual has been completed (130.1 deaths). Hence the expected total number deaths at the time when the target accrual is completed is 97.4+130.1 = 227.5. Likewise, the expected number of deaths reported 9 months after the targeted accrual is complete is 100.2+150.0=250.2, which is approximately the number required for the final analysis.

Therefore, the targeted date for completing the accrual to the surgical component of this study is Fall-2018. The required number of deaths required for the final analysis (250 deaths) is expected to occur nearly 1 year after the targeted accrual has been completed. In the event that the required number of deaths for the final analysis is observed before the targeted accrual is completed, then accrual will be stopped prior to attaining the targeted accrual. The planned analyses and the

power calculations provided above are unchanged by this revised recruitment plan.**(12/19/2011)**

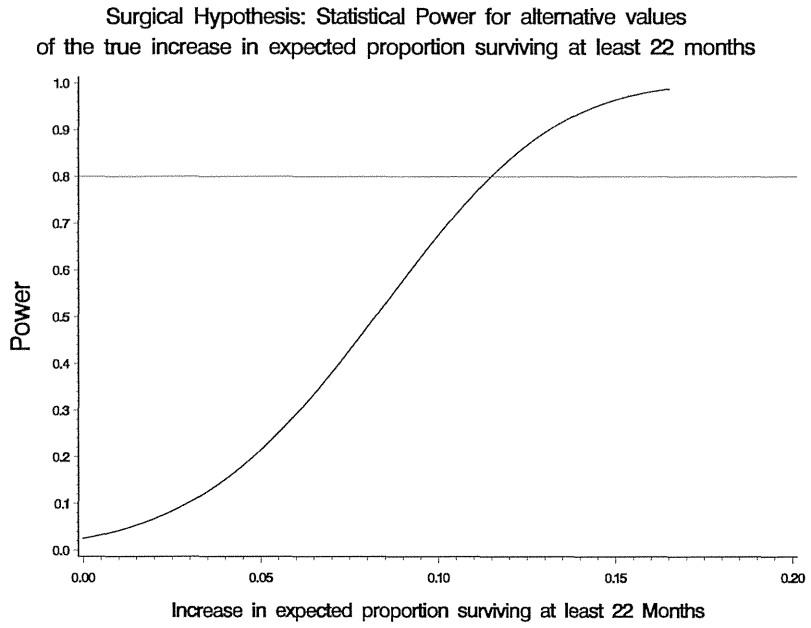

Figure 2.

The number of patients to be enrolled onto the surgical component of this study depends on the proportion of patients who are candidates for surgery and willing to have their surgical treatment determined through randomization. In the event that the targeted accrual for objective 1 (660 patients) is attained and there are too few patients enrolled who are either not candidates for surgery or do not consent to surgery (objective 2) then consideration will be given to continuing randomization to the surgical cytoreduction factor only. That is, accrual will be continued but randomization to the systemic therapies will be stopped, and only randomization to the surgical intervention will continue. Adjuvant chemotherapy will be determined at that time. In the event that the chemotherapy objectives are known, the choice of regimen will follow that finding the "winning arm". In the event that this is unknown, adjuvant therapy will be the control arm, paclitaxel and carboplatin with or without bevacizumab. It is anticipated that at least 360 patients will need to be enrolled onto the surgical component of this study.**(08/29/11)**

Interim Analyses: Interim analyses are planned when there are at least 110 deaths reported among all those patients randomly allocated (prior to August 29, 2011) to the CT regimen (approximately 50% of the full information of the systemic therapy component of this study), The actual time for the interim analyses will coincide with the nearest scheduled Data Monitoring Committee (DMC) meeting for which the required number of events has occurred. The semi-annual DMC meetings coincide with the GOG business meetings which are held in January and

July each year and the precise date of these meetings is set without confidential knowledge of the study results. **(10/01/12)**

The interim analyses will include an assessment of treatment efficacy. An alpha-spending function proposed by Lan and DeMets[82], which mimics the O'Brien and Fleming[83] group sequential boundary, will be used to calculate the critical values used for the interim analyses. The proportion of the total information available at the interim analysis will be calculated as the fraction: number of observed deaths among those randomly allocated to the CT regimen to the planned total number of deaths required for final analysis. For example, if the interim analysis occurs at 55% of the information time, H_{01} will be assessed using the previously described stratified logrank test and the critical p-value set to 0.0082 for the interim analysis and 0.0475 for final analysis. **(08/29/11)**

H_{02} will also be assessed at this interim analysis with a similar error spending function, but the critical values for this assessment will be based on the proportion of the total information calculated as the number of reported deaths among those enrolled into this component of the study relative to the total number of deaths required for the final analyses. Finally, a second interim analysis of H_{02} will occur when at least 50% of the planned number of deaths has been reported. The critical values for this assessment will be based on the error spending function, the type I error spent on the previously mentioned interim analysis, and the actual proportion of deaths reported at the time of this interim analysis.

The interim analysis that will occur at approximately the 50% of the total information time for H_{01} (or H_{02}) will also include futility analyses. Since the purpose of the study is to identify interventions that increase overall survival duration, consideration will be given to stopping randomization to the experimental interventions (CTB or cytoreductive surgery) if it exhibits poorer survival relative to its control treatment (CT or no surgical intervention, respectively) indicated by an adjusted hazard ratio, $\Delta_{01} > 1.0$ (or $\Delta_{02} > 1.0$) at the time of the interim analyses. This interim decision rule decreases the statistical power for each pair-wise comparison by less than 1%. The results of the interim analyses are reviewed by the GOG Data Monitoring Committee (DMC). The decision to terminate randomization to any particular regimen includes consideration of toxicities, treatment compliance, progression-free survival, and results from external studies.

Final analysis: The study will be considered sufficiently mature to permit a final assessment of H_{01} when there are at least 214 deaths reported among those patients assigned to the standard regimen (CT). The study will be considered sufficiently mature to permit an assessment of H_{02} when there at least 250 deaths reported among those enrolled (and randomized) onto the surgical component of this study. The previously described logrank test will be performed and the corresponding treatment hazard ratio will be estimated. The critical values for rejecting the null hypotheses will be adjusted for interim analyses, using the

O'Brien and Fleming-like type I error spending function proposed by Lan and DeMets (1986).

Supplemental final analyses: A proportional hazards model will be fitted to the survival data. Apart from the randomized therapy, other factors such as: prior exposure to bevacizumab, histologic cell type, initial performance status, and age will be included in the model as potential confounders because there exists evidence that these factors may have an effect on survival in these patients. Race and ethnicity will also be assessed, but there is no specific hypothesis concerning an interaction between these factors and treatment proposed.

Due to the lack of knowledge concerning interactions between treatments and these confounders, prior to assessing the main effects of the treatment, the homogeneity of the treatment effects will be assessed by testing the null hypothesis of no interactions. The likelihood ratio test will be performed by comparing models with and without interaction terms. Rejecting the null hypothesis of homogeneity at the 5% significance level will be considered sufficient evidence to warrant reporting the relative treatment effects within each factor and a cautious interpretation of the pooled estimates.

Safety Analyses: The GOG Data Safety and Monitoring Board (DSMB) reviews accumulating summaries of toxicities, serious adverse event (SAE) reports and deaths in which study treatment may have been a contributing cause. This committee does not review efficacy results. The DSMB may recommend study amendments pertaining to patient safety.

Grading and classification of adverse events will follow the CTCAE v. 3.0 toxicity criteria. The primary analysis will consist of comparing the relative odds of grade 3 or worse toxicities occurring during and following study treatment that are reported to be at least possibly related to study treatment. Specific attention will be given to the frequency of neutropenia, anemia, thrombocytopenia, hypertension, proteinuria, rash and gastrointestinal toxicities, which are unintended but not infrequent side effects from the study treatments. The safety analysis will focus on those patients who at least initiated study therapy. A logistic model will permit an estimate of the relative odds of grade 3 or worse adverse treatment effects for both randomized factors while adjusting for potential confounders like age. Deaths considered to be at least partially attributable to treatment will be reported and summarized. Reasons for stopping study treatment (e.g. patient refusal, toxicity, progression or death) will be reported.

11.4 Quality of Life

There are primarily three quality of life issues of interest:

11.41 Patients undergoing secondary cytoreduction may initially experience a decrease in QOL after the surgery.

11.42 Patients undergoing secondary cytoreduction may have better QOL compared to patients without surgery, after surgical healing.

11.43 Patients receiving carboplatin and paclitaxel only may experience a better QOL relative to those who receive these agents combined with bevacizumab.

The primary measures used in this study to assess the quality of life (QOL) are the self-administered FACT-O TOI for ovarian cancer patients and the Physical Functioning (PF) subscale of the Rand-SF 36. Each patient will be asked to complete these questionnaires at the following time points during their participation in the study:

i. Prior to surgery (only those undergoing surgical cytoreduction),
ii. Prior to cycle 1
iii. Prior to cycle 3 (6 weeks after staring systemic therapy),
iv. Prior to cycle 6 (12 weeks after starting systemic therapy),
v. 6 months after starting systemic therapy,
vi. 12 months after starting systemic therapy.

The times in parentheses indicate the assessment points for those patients who do not complete the entire study regimen.

Construct and content
The Functional Assessment of Cancer Therapy scale developed for ovarian cancer (FACT-O TOI) is a tool that provides a general QOL score. It consists of 3 subscales: physical well being (7 items), functional well being (7 items) and the Ovarian Cancer subscale (12 items)[43]. The Physical Functioning (PF) subscale of the Rand-SF 36 is the 10-item subscale of this general quality of life questionnaire[44].

Descriptive analyses of baseline QOL scores
Descriptive statistics from the baseline QOL data will be calculated. These will include descriptions of the distribution of QOL scores (mean, standard deviation, median, etc.). For all patients the baseline scores will be calculated using the questionnaire completed prior to treatment cycle 1, except for those undergoing cytoreductive surgery. In that case, the baseline score is calculated using the questionnaire completed prior to surgery. Therefore, comparisons involving the patients who were allocated to surgery, the effects of time are confounded with effects of surgery.

Differences in FACT-TOI scores between patients receiving CT and CTB: Data available from GOG-172 provides some estimates that can be used for planning the current study. In that study women with advanced ovarian cancer received 6 cycles of a platinum-taxane based treatment. The mean and variance of the baseline FACT-TOI scores were 67.2 and 15.9, respectively. The correlation

between the baseline FACT-TOI score and the same score reported 3 to 6 weeks after the sixth cycle of treatment was 0.36. The target sample size for this study is based on study objective 1 and is 660 patients (330 patients in each arm). It is anticipated that 90% of the patients will report FACT-TOI scores prior to initiating treatment and prior to treatment cycle 6. If bevacizumab truly changes patients' FACT-TOI scores 4.0 units after 6 cycles of treatment, the targeted sample size has about 91% power for an analysis of covariance, when the type I error is limited to 5% for a two tail test. However, a linear mixed model will be used for the final analysis of this data since this approach is more efficient, accommodates missing data, and accounts for correlations due to repeated measurements from the same individual. The actual power for this analysis, however, depends on the unknown pattern of missing values.

Difference in Rand SF-36 Physical Functioning (PF) subscale after surgery: This analysis will include those patients who were candidates for surgery and consented to have their surgical intervention determined through randomization. A paired t-test will be used to test the null hypothesis that there is no difference between baseline PF scores and the PF scores prior to cycle 1 for those patients randomized to cytoreductive surgery. The paired t-test is generally robust for moderate sample sizes when the distribution of PF scores is not normal. Data from GOG-2222 can be used for planning purposes. In that study, patients with newly diagnosed endometrial cancer completed the SF-36 PF subscale prior to initiating study treatment. When the scores are rescaled (0-100), the mean and standard deviation were 75.9 and 27.8, respectively. Assuming that at least 360 patients will be eligible and consent to participate in this component of the study, this sample size has 87% power for detecting a true difference of 9 units, when type I error is limited to 5% for a two-tail test.

Differences in FACT-TOI scores between patients undergoing cytoreductive surgery and patients not undergoing cytoreductive surgery. This analysis will include those patients who were candidates for surgery and consented to have their surgical intervention determined through randomization. There will be up to 5 or 6 time points for patients to report their FACT-TOI scores, depending on whether the patient was randomized to the secondary cytoreductive surgery arm of the study. There are no specific hypotheses being posited for how the treatment groups will differ in their mean QOL scores over time. Therefore, a linear mixed model will be used to model the difference in mean QOL scores over time. That is, the mean QOL scores will be modeled in order to compare those patients randomized to secondary cytoreductive surgery vs no surgery, as well as, the differences in mean QOL scores among the three types of systemic therapy. The model will assess whether there is evidence of a treatment-time interaction as well as whether differences in mean QOL scores between treatment groups varies as a linear or possibly a quadratic function of time.

11.5 Translational Research Statistical Considerations

Overview
The overall goal for the translational research component of this study is to discover molecular and/or biochemical profiles that may be useful for determining which patients from this patient population are likely to respond or experience longer survival. There are no specific up-front hypotheses proposed. The primary challenges related to this component of the study are the practicality of finding useful biomarker profiles from among potentially tens of thousands of biomarkers, as in the case of a gene microarray experiment. This challenge is further exasperated due to the limited number of available biologic specimens. In order to address these challenges, this study will utilize a training dataset to develop a prognostic index from the biomarker measurements and a separate and distinct validation dataset to assess the predicative value of the index. The steps to be used in this study for developing a molecular/biochemical profile are:

a) Identifying those individuals to be included in the training data set.
b) Developing an index from the molecular marker data and outcome data contained in the training set.
c) Assess reliability of the putative prognostic index.
d) Validate the putative prognostic index.

Anticipated sample size for translational objectives
The targeted accrual for the randomized systemic treatment component of this study is 660 patients. It is anticipated that 30% - 50% (approximately 198-330 patients) of these individuals will be candidates for and consent to secondary surgical cytoreduction. Only half of these individuals (99-165 patients) will be randomized to cytoreductive surgery. It is expected that viable tissue collected during cytoreductive surgery will be available in most of these cases. Also, a serum sample, which is drawn prior to surgery, will be available. The ratio of the size of the training dataset to the size of the validation dataset will range from 1:1 to 3:1.

Therefore, assuming that a biologic specimen is available from 130 eligible and evaluable patients who were treated with a randomized systemic treatment and undergoing surgery, the size of the training set is expected to range from 65-98 patients and the size of the validation dataset is expected to range from 32-65 patients.

Training and validation set
In order to establish a training dataset for the primary translational research objectives a sample of sequentially enrolled eligible and evaluable patients will be established in which at least 50 deaths have been reported. This requirement may need to be relaxed for follow-up studies since samples will eventually become depleted. A validation cohort will be derived in a similar fashion as the training cohort. That is, the training and validation cohorts will consist of sequentially enrolled eligible and evaluable patients. Individuals will not be permitted to be members of both the training and the validation cohorts.

Identifying biologic/molecular profiles
Investigators will not be restricted to utilizing a particular technique for building the classifier. In fact, several classifiers may be identified. However, prior to the validation phase a single classifier corresponding to the primary study objective will be selected and deemed the 'final' classifier. Data from the validation dataset will not be used to select the 'final' classifier.

Reliability
The classifier should provide similar results for the same experimental unit. That is, a biologic specimen with a high prognostic index score should exhibit a high prognostic index score when it is re-evaluated. An index score which cannot be replicated lacks test-retest reliability. This occurs when there are other sources of between-specimen variation that are uncontrolled in the experiment.

In order to assess reliability some specimens will be selected from the training set for repeat assessment. While randomly selecting specimens from the training set for replication is preferable, it may be necessary to randomly select from a subset of the training set due to the availability of adequate biologic material. When possible, the samples will be identified in such a way that the laboratory investigator will be unable to identify which specimens are replicates. These samples will have their biomarkers (i.e. gene expression, serum marker) assessed twice. Since replication can be expensive, depending of the laboratory procedure, the number of samples selected for replication will vary from a dozen to a few dozen, depending on practical considerations like cost and feasibility. The data from the replicated samples will be used to assess reliability of the putative index before proceeding to the validation phase. Reproducibility is a prerequisite for a clinically useful classifier.

Validation
Prior to initiating the validation phase, the 'final' classifier will be completely documented (i.e. computer program or pseudo-code). This documentation will be reviewed by individuals in the GOG Statistical and Data Center (SDC) who are not participating in the analyses. The purpose of this review will be to determine whether the final classifier has been unambiguously defined.

The c-index will be used to measure the classifier's predictive ability. This index assesses the strength of the rank correlation between the predicted outcome and the actual outcome. If the classifier produces a continuous prognostic score and response is dichotomous, then the c-index is comparable to the Wilcoxon two-sample rank score. It can be calculated by taking all possible pairs of individuals in which one individual responded and the other did not respond. In this case, the c-index is the proportion of these pairs in which the responder has a higher predicted probability of responding. A c-index value of 0.5 indicates a useless classifier, and a value of 1.0 indicates perfect prediction. The c-index is Somer's rank correlation index when it is rescaled to vary linearly from 0 to 1. The c-

index can be used when the outcome is partially censored survival time. In this case it measures the proportion of all pairs of individuals in the data set in which the individual with the expected lower risk of failure is known to survive longer.

Other descriptive summaries of predictive ability will also be considered including: Kaplan-Meier curves when the outcome is a time-to-failure or a ROC curve when the outcome is dichotomous.

The publication which describes the results for the primary objective of this study will include a description of the accuracy of the final classifier. While other classifiers may also be described, the final classifier will be clearly distinguished from the other classifiers. The documentation describing the final classifier will be available to other investigators from the SDC upon request.

After the study objectives have been completed, the GOG may elect to make some or all of the validation data set available to other investigators, since the specimens in the training set may become exhausted. Any classifiers developed subsequently will not be permitted to claim that they were independently validated without additional supporting external evidence.

11.6 The anticipated distribution of patients' race and ethnicity for the systemic therapy
portion of this trial is (all are female):

White (not Hispanic)	584
Black (not Hispanic)	39
Hispanic	14
Asian	17
American Indian or Alaskan Native	3
Native Hawaiian or other Pacific Islander	3

12.0 BIBLIOGRAPHY

1. Vaccarello L, Rubin SC, Vlamis V, Wong G, Jones WB, Lewis JL, et al. Cytoreductive surgery in ovarian carcinoma patients with a documented previously complete surgical response. Gynecol Oncol 1995; 57(1):61-5.

2. Bristow RE, Lagasse LD, Karlan BY. Secondary surgical cytoreduction for advanced epithelial ovarian cancer. Patient selection and review of the literature. Cancer 1996; 78(10):2049-62.

3. Berek JS, Hacker NF, Lagasse LD, Nieberg RK, Elashoff RM. Survival of patients following secondary cytoreductive surgery in ovarian cancer. Obstet Gynecol 1983; 61(2):189-93.

4. Janicke F, Holscher M, Kuhn W, von Hugo R, Pache L, Siewert JR, et al. Radical surgical procedure improves survival time in patients with recurrent ovarian cancer. Cancer 1992; 70(8):2129-36.

5. Eisenkop SM, Friedman RL, Spirtos NM. The role of secondary cytoreductive surgery in the treatment of patients with recurrent epithelial ovarian carcinoma. Cancer 2000; 88(1):144-53.

6. Munkarah AR, Coleman RL. Critical evaluation of secondary cytoreduction in recurrent ovarian cancer. Gynecol Oncol 2004; 95(2):273-80.

7. Morris M, Gershenson DM, Wharton JT, Copeland LJ, Edwards CL, Stringer CA. Secondary cytoreductive surgery for recurrent epithelial ovarian cancer. Gynecol Oncol 1989; 34(3):334-8.

8. Gordon AN, Tonda M, Sun S, Rackoff W. Long-term survival advantage for women treated with pegylated liposomal doxorubicin compared with topotecan in a phase 3 randomized study of recurrent and refractory epithelial ovarian cancer. Gynecol Oncol 2004; 95(1):1-8.

9. Dizon DS, Hensley ML, Poynor EA, Sabbatini P, Aghajanian C, Hummer A, et al. Retrospective analysis of carboplatin and paclitaxel as initial second-line therapy for recurrent epithelial ovarian carcinoma: application toward a dynamic disease state model of ovarian cancer. J Clin Oncol 2002; 20(5):1238-47.

10. Goldberg JM, Piver MS, Hempling RE, Recio FO. Paclitaxel and cisplatin combination chemotherapy in recurrent epithelial ovarian cancer. Gynecol Oncol 1996; 63(3):312-7.

11. Parmar MK, Ledermann JA, Colombo N, du Bois A, Delaloye JF, Kristensen GB, et al. Paclitaxel plus platinum-based chemotherapy versus conventional platinum-based chemotherapy in women with relapsed ovarian cancer: the ICON4/AGO-OVAR-2.2 trial. Lancet 2003;361(9375):2099-106.

12. Gonzalez-Martin AJ, Calvo E, Bover I, Rubio MJ, Arcusa A, Casado A, et al. Randomized phase II trial of carboplatin versus paclitaxel and carboplatin in platinum-sensitive recurrent advanced ovarian carcinoma: a GEICO (Grupo Espanol de Investigacion en Cancer de Ovario) study. Ann Oncol 2005; 16(5):749-55.

13. Pfisterer J, Plante M, Vergote I, du Bois A, Wagner U, Hirte H, et al. Gemcitabine/carboplatin (GC) vs. carboplatin (C) in platinum sensitive recurrent ovarian cancer (OVCA). Results of a Gynecologic Cancer Intergroup randomized phase III trial of the AGO OVAR, the NCIC CTG and the EORTC GCG. Proc Am Soc Clin Oncol 2004; 22(14S): Abstr#5005.

14. Gueritte-Voegelein F, G.D., Lavell F, et al., Relationships between the structure of Taxol analogues and their antimitotic activity. J Med Chem, 1991. **34**: p. 992-998.

15. Riou, J.F., A. Naudin, and F. Lavelle, Effects of Taxotere on murine and human tumor cell lines. Biochem Biophys Res Commun, 1992. **187**(1): p. 164-70.

16. Ringel, I. and S.B. Horwitz, Studies with RP 56976 (taxotere): a semisynthetic analogue of taxol. J Natl Cancer Inst, 1991. **83**(4): p. 288-91.

17. Kelland, L.R. and G. Abel, Comparative in vitro cytotoxicity of taxol and Taxotere against cisplatin-sensitive and -resistant human ovarian carcinoma cell lines. Cancer Chemother Pharmacol, 1992. **30**(6): p. 444-50.

18. Aapro M, B.B., Dietel M, Hill B, Kelland L, Lelieveld P, Silvestrini R, Zoli W., Superior activity of Taxotere (Ter) over Taxol (tol) in vitro. Proc Am Assoc Cancer Res, 1992. **33**: p. 3086.

19. Hanauske, A.R., et al., Effects of Taxotere and taxol on in vitro colony formation of freshly explanted human tumor cells. Anticancer Drugs, 1992. **3**(2): p. 121-4.

20. Alberts DS, G.D., Fanta P, Liu D, Roe D, Salmon SE., Comparative cytotoxicities of Taxol and Taxotere in vitro against fresh human ovarian cancers. Proc Am Soc Clin Oncol, 1992. **aa**: p. 719.

21. Hennequin, C., N. Giocanti, and V. Favaudon, S-phase specificity of cell killing by docetaxel (Taxotere) in synchronized HeLa cells. Br J Cancer, 1995. **71**(6): p. 1194-8.

22. Diaz, J.F. and J.M. Andreu, Assembly of purified GDP-tubulin into microtubules induced by taxol and taxotere: reversibility, ligand stoichiometry, and competition. Biochemistry, 1993. **32**(11): p. 2747-55.

23. Hill, B.T., et al., Differential cytotoxic effects of docetaxel in a range of mammalian tumor cell lines and certain drug resistant sublines in vitro. Invest New Drugs, 1994. **12**(3): p. 169-82.

24. Garcia, P., et al., Comparative effects of taxol and Taxotere on two different human carcinoma cell lines. Cancer Chemother Pharmacol, 1994. **34**(4): p. 335-43.

25. Bissery, M.C., et al., Experimental antitumor activity of taxotere (RP 56976, NSC 628503), a taxol analogue. Cancer Res, 1991. **51**(18): p. 4845-52.

26. Francis, P., et al., Phase II trial of docetaxel in patients with platinum-refractory advanced ovarian cancer. J Clin Oncol, 1994. **12**(11): p. 2301-8.

27. Kavanagh, J.J., et al., Phase II study of docetaxel in patients with epithelial ovarian carcinoma refractory to platinum. Clin Cancer Res, 1996. **2**(5): p. 837-42.

28. Piccart, M.J., et al., Docetaxel: an active new drug for treatment of advanced epithelial ovarian cancer. J Natl Cancer Inst, 1995. **87**(9): p. 676-81.

29. Valero, V. Docetaxel as single-agent therapy in metastatic breast cancer: clinical efficacy. Semin Oncol. 1997 Aug; 24 (4 Suppl 13): S13-11-S13-18. Review.

30. Verschraegen C, K.A., Steger M, Edwards C, Kavanagh J. Randomized phase II study of two dose levels of docetaxel in patients with advanced epithelial ovarian cancer who have failed paclitaxel chemotherapy (Meeting abstract). in 1997 ASCO Annual Meeting. 1997.

31. Oulid-Aissa D, B.R., Lebecq A, Zukiwski A, Sheiner L, Riva A. Taxotere safety in patients with impaired liver function (LF) (Meeting abstract). in 1996 ASCO Annual Meeting. 1996.

32. Rose PG, Blessing JA, Ball HG, Hoffman J, Warshal D, DeGeest K, Moore DH. A phase II study of docetaxel in paclitaxel-resistant ovarian and peritoneal carcinoma: a Gynecologic Oncology Group Study. Gynecol Oncol. 2003 Feb; 88 (2): 130-5.

33. Markman, M Zanotti K, Webster K, Peterson G, Kulp B, Belinson J. Phase 2 trial of single agent docetaxel in platinum and paclitaxel-refractory ovarian cancer, fallopian tube cancer, and primary carcinoma of the peritoneum. Gynecol Oncol. 2003 Dec; 91 (3): 573-6.

34. Vasey PA, Jayson GC, Gordon A, Gabra H, Coleman R, Atkinson R, Parkin D, Paul J, Hay A, Kaye SB; Scottish Gynecological 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 Nov 17; 96 (22): 1682-91.

35. Alvarez AA, Krigman HR, Whitaker RS, Dodge RK, Rodriguez GC. The prognostic significance of angiogenesis in epithelial ovarian carcinoma. Clin Cancer Res 1999;5(3):58791.

36. Gasparini G, Bonoldi E, Viale G, Verderio P, Boracchi P, Panizzoni GA, et al. Prognostic and predictive value of tumour angiogenesis in ovarian carcinomas. Int J Cancer 1996; 69(3):205-11.

37. Nakanishi Y, Kodama J, Yoshinouchi M, Tokumo K, Kamimura S, Okuda H, et al. The expression of vascular endothelial growth factor and transforming growth factor-beta associates with angiogenesis in epithelial ovarian cancer. Int J Gynecol Pathol 1997; 16(3):256-62.

38. Yoneda J, Kuniyasu H, Crispens MA, Price JE, Bucana CD, Fidler IJ. Expression of angiogenesis-related genes and progression of human ovarian carcinomas in nude mice. J Natl Cancer Inst 1998; 90(6):447-54.

39. Folkman J. Tumor angiogenesis: therapeutic implications. N Engl J Med 1971; 285(21):1182-6.

40. Malonne H, Langer I, Kiss R, Atassi G. Mechanisms of tumor angiogenesis and therapeutic implications: angiogenesis inhibitors. Clin Exp Metastasis 1999; 17(1):1-14.

41. Kim KJ, Li B, Winer J, Armanini M, Gillett N, Phillips HS, et al. Inhibition of vascular endothelial growth factor-induced angiogenesis suppresses tumour growth in vivo. Nature 1993; 362(6423):841-4.

42. Presta LG, Chen H, O'Connor SJ, Chisholm V, Meng YG, Krummen L, et al. Humanization of an anti-vascular endothelial growth factor monoclonal antibody for the therapy of solid tumors and other disorders. Cancer Res 1997; 57(20):4593-9.

43. Burger RA, Sill M, Monk BJ, Greer BE, Sorosky J. Phase II trial of bevacizumab in persistent or recurrent epithelial cancer or peritoneal primary cancer: a Gynecologic Oncology Group study. Proc Am Soc Clin Oncol 2005; 23(14S): Abst#5009.

44. Zanotti KM, Rybicki LA, Kennedy AW, Belinson JL, Webster KD, Kulp B, Peterson G, Markman M: Carboplatin skin testing: a skin-testing protocol for predicting hypersensitivity to carboplatin chemotherapy. J Clin Oncol 2001; 19:3126-3129.

45. Rose PG, Fusco N, Smrekar M, Mossbruger K, Rodriguez M: Successful administration of carboplatin in patients with clinically documented carboplatin hypersensitivity. Gynecol Oncol 2003; 89:429-433.

46. Robinson JB, Singh D, Bodurka-Bevers DC, Wharton JT, Gershenson DM, Wolf JK: Hypersensitivity reactions and the utility of oral and intravenous desensitization in patients with gynecologic malignancies. Gynecol Oncol 2001; 82:550-558.

47. Markman M: Hypersensitivity reactions to carboplatin. Gynecol Oncol 2002; 84:353-354.

48. Lee CW, Matulonis UA, Castells MC: Rapid inpatient/outpatient desensitization for chemotherapy hypersensitivity: Standard protocol effective in 57 patients for 255 courses. Gynecol Oncol 2005;

49. Markman M, Kennedy A, Webster K, Elson P, Peterson G, Kulp B, Belinson J: Clinical features of hypersensitivity reactions to carboplatin. J Clin Oncol 1999; 17:1141.

50. Shepherd FA, Pereir JR, Ciuleanu T, Tan EH, Hirsh V, Thongpraser S, Campos D, Maoleekoonpiroj S, Smylie M, Martins R, van Kooten M, Dediu M, Findlay B, Tu D, Johnston D, Bezjak A, Clark G, Santabarbara P, Seymour L for the National Cancer Institute of Canada Clinical Trials Group. Erlotinib in previously treated non-small-cell lung cancer. The New England Journal of Medicine, 353: 123-132, 2005.

51. Miller KD, Chap LI, Holmes FA, Cobleigh MA, Marcom K, Fehrenbacher L, Dickler M, Overmoyer BA, Reimann JD, Sing AP, Langmuir V, Rugo HS. Randomized Phase III trial of capecitabine compared with bevacizumab plus capecitabine in patients with previously treated metastatic breast cancer. Journal of Clinical Oncology 23: 792-799, 2005.

52. Hurwitz H, Fehrenbacher L, Novotny W, Cartwright T, Hainsworth J, Heim W, Berlin J, Baron A, Griffing S, Holmgren E, Ferrara N, Fyfe G, Rogers B, Ross R, Kabbinavar F. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. The New England Journal of Medicine, 350: 2335-2342, 2004.

53. Basen-Engquist K, Bodurka-Bevers D, Fitzgerald MA, Webster K, Cella D, Hu S, Gershenson DM. Reliability and validity of the Functional Assessment of Cancer Therapy – Ovarian (FACT-O). J Clin Onc 19(6): 1809-1817, 2001.

54. Stewart AL, Ware JE, Jr. Measuring functioning and well-being. Durham, NC: Duke University Press, 1992.

55. Ware JE, Sherbourne CD. The MOS 36-item short form health survey (SF36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Medical Care 31:247-263, 1993.

56. Hays, RD. Rand-36 Health Status Inventory. San Antonio: The Psychological Corporation, Harcourt Brace & Company, 1998.

57. McHorney CA, Ware JE, Raczek AE. The MOS 36-item short form health survey (SF36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Medical Care 30:247-263, 1993.

58. Mangione CM, Goldman L, Orav J, Marcantonio ER, Pedan A, Ludwig LE, Donaldson MC, Sugarbaker DJ, Poss R, Lee TH. Health-related quality of life after elective surgery. Journal of General Internal Medicine 12:686-697, 1997.

59. Nguyen NT, Goldman C, Rosenquist CJ, Arango A, Cole CJ, Lee SJ, Wolfe BM.

Laparoscopic versus open gastric bypass: A randomized study of outcomes, quality of life, and costs. Annals of Surgery 234:279-291, 2001.

60. Velanovich V. Comparison of symptomatic and quality of life outcomes of laparoscopic versus open antireflux surgery. Surgery 126:782-789, 1999.

61. Cella D. Manual of the Functional Assessment of Chronic Illness Therapy (FACIT) Scales. Evanston, IL: Center on Outcomes, Research and Education (CORE), Northwestern Healthcare and Northwestern University, 1997.

62. Golub TR, Slonim DK, Tamyo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286(5439): p.531-7, 1999.

63. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403(6769): p.503-11, 2000.

64. Boulikas T, Vougiouka M. Cisplatin and platinum drugs at the molecular level (Review). OncolRep 10(6): p.1663-82, 2003.

65. Orr GA, Verdier-Pinard P, McDaid H, Horwitz SB. Mechanisms of Taxol resistance related to microtubules (Review). Oncogene 22(47): p.7280-95, 2003.

66. Crum CP, Drapkin R, Kindelberger D, et al: Lessons from BRCA: The Tubal Fimbria Emerges as an Origin for Pelvic Serous Cancer. Clin Med Res 5:35-44, 2007

67. Longacre TA, Oliva E, Soslow RA: Recommendations for the reporting of fallopian tube neoplasms. Hum Pathol, 2007

68. Pectasides D, Pectasides E, Economopoulos T: Fallopiantube carcinoma: a review. Oncologist 11:902-12, 2006

69. Pocock, SJ and Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics 31:103-115 1975.

70. Dunnett, CW. A multiple comparison procedure for comparing several treatments with a control. JASA 50(272): 1096-1121 1955.

71. Lan KKG and DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika 70:659-663 1983.

72. O'Brien PC and Fleming TR. A multiple testing procedure for clinical trials. Biometrics 35:549-556 1979.

73. 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; Apr 23, ePub ahead of print

APPENDIX I

FIGO STAGE GROUPING FOR PRIMARY CARCINOMA OF THE OVARY

(1985)

These categories are based on findings at clinical examination and/or surgical exploration. The histologic characteristics are to be considered in the staging, as are results of cytologic testing as far as effusions are concerned. It is desirable that a biopsy be performed on suspicious areas outside the pelvis.

Stage I	Growth limited to the ovaries.
Stage IA	Growth limited to one ovary; no ascites. No tumor on the external surface; capsule intact.
Stage IB	Growth limited to both ovaries; no ascites. No tumor on the external surfaces; capsules intact.
Stage IC*	Tumor either Stage IA or IB but with tumor on the surface of one or both ovaries; or with capsule ruptured; or with ascites present containing malignant cells or with positive peritoneal washings.
Stage II	Growth involving one or both ovaries with pelvic extension.
Stage IIA	Extension and/or metastases to the uterus and/or tubes.
Stage IIB	Extension to other pelvic tissues.
Stage IIC*	Tumor either Stage IIA or IIB but with tumor on the surface of one or both ovaries; or with capsule(s) ruptured; or with ascites present containing malignant cells or with positive peritoneal washings.
Stage III	Tumor involving one or both ovaries with peritoneal implants outside the pelvis and/or positive retroperitoneal or inguinal nodes. Superficial liver metastasis equals Stage III. Tumor is limited to the true pelvis but with histologically verified malignant extensions to small bowel or omentum.
Stage IIIA	Tumor grossly limited to the true pelvis with negative nodes but with histologically confirmed microscopic seeding of abdominal peritoneal surfaces.
Stage IIIB	Tumor of one or both ovaries with histologically confirmed implants of abdominal peritoneal surfaces, none exceeding 2 cm in diameter. Nodes negative.
Stage IIIC	Abdominal implants >2 cm in diameter and/or positive retroperitoneal or inguinal nodes.
Stage IV	Growth involving one or both ovaries with distant metastasis. If pleural effusion is present there must be positive cytologic test results to allot a case to Stage IV. Parenchymal liver metastasis equals Stage IV.

* In order to evaluate the impact on prognosis of the different criteria for allotting cases to Stage IC or IIC, it would be of value to know if rupture of the capsule was (1) spontaneous or (2) caused by the surgeon and if the source of malignant cells detected was (1) peritoneal washings or (2) ascites.

APPENDIX II

SECONDARY CYTOREDUCTIVE SURGICAL PROCEDURE

Purpose : Maximum resection of recurrent ovarian cancer.

Timing: Surgical exploration should be undertaken within 4 weeks of study entry.

Content of Procedure:

1.0 The abdominal incision must be adequate to explore the entire abdominal cavity and allow safe cytoreductive surgery. A vertical incision is recommended but not required.
2.0 All peritoneal surfaces including the undersurface of both diaphragms and the serosa and mesentery of the entire gastrointestinal tract will be visualized and palpated for evidence of metastatic disease.
3.0 Visible metastatic abdominal and pelvic disease should be resected or ablated completely, if possible.
4.0 Diaphragmatic recurrent disease should be resected. Ablation of disease with electrocautery (e.g. Argon Beam Coagulator) is acceptable.
5.0 Surgical evaluation of the pelvic and paraortic node bearing areas requires resection if not performed on initial staging/debulking procedure. If incomplete nodal resection was previously documented, unresected areas should be excised.
6.0 Solid organ metastases (spleen and liver) should be considered for resection. Treatment by Radio Frequency Ablation (RFA) is acceptable.

Goal: Surgical goal of cytoreduction is to reduce volume of residual disease to smallest quantity possible (no visible residual).

Reporting: The size (two dimensions) and location of residual disease will be recorded.