Early Changes in CA125 After Treatment with Pegylated Liposomal Doxorubicin or Topotecan Do Not Always Reflect Best Response in Recurrent Ovarian Cancer Patients

ROBERT L. COLEMAN, a ALAN GORDON, b JAMES BARTER, c STEVEN SUN, d WAYNE RACKOFF, d THOMAS J. HERZOGe

aThe University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA; bArizona Gynecologic Oncology, Phoenix, Arizona, USA; cWomen’s Health Specialists, Bethesda, Maryland, USA; dJohnson & Johnson Pharmaceutical Research & Development, L.L.C., Raritan, New Jersey, USA; eColumbia University, New York, New York, USA

Key Words. CA125 • Biomarkers • Ovarian cancer • Response to therapy • Pegylated liposomal doxorubicin Topotecan

ABSTRACT
Purpose. To examine early changes in CA125 relative to objective response in patients with recurrent ovarian cancer treated with pegylated liposomal doxorubicin (PLD) or topotecan and to compare the CA125 trends between the two chemotherapeutics.

Patients and Methods. Patients with recurrent ovarian cancer, all of whom had measurable or evaluable disease, were randomized to receive 50 mg/m² PLD every 28 days (n = 239) or 1.5 mg/m² topotecan for 5 days every 21 days (n = 235) as part of a previously reported multicenter study. CA125 measurements were obtained prior to therapy and with each cycle of administration. Assessable patients underwent radiographic evaluation for response after two cycles of therapy. Objective responses were compared to trends in CA125 values at the end of cycles 1 and 2. CA125 changes were categorized as baseline (±10%), ±10%–25% variance, and >25% variance.

Results. Among patients treated with PLD, 50% of complete responders (CR) and 41% of partial responders (PR) had increases in CA125 from baseline to cycle 1. Increases in CA125 were also seen in topotecan-treated patients; however, fewer patients had increases (20% and 8%, respectively). Overall, 15% of responding patients (CR + PR) receiving PLD and 6% receiving topotecan had elevated CA125 after two cycles of therapy. For those patients achieving a partial response, 19% of PLD-treated patients and 8% of topotecan-treated patients had CA125 levels above baseline at cycle 2.

Conclusions. Considerable intrapatient variation in CA125 values is present among responding patients. Early increases in CA125 may not predict ultimate outcome, especially in PLD-treated patients. The Oncologist 2007;12:72–78

INTRODUCTION
Despite the well-characterized limitations in the interpretation of a solitary CA125 value, this biomarker is widely used to prospectively evaluate therapeutic efficacy and monitor disease status among ovarian cancer patients [1, 2]. More than 80% of patients with advanced ovarian cancer...
have elevated CA125 values at the beginning of therapy. Trends in CA125 values during the administration of each cycle of therapy are easy to follow. In general, a negative slope in serial determinations is reassuring and often considered as evidence of therapeutic efficacy. A positive slope is disconcerting and suggests the persistence of disease or emergence of drug resistance. Decisions to stop or continue an individual therapeutic are often made in reference to these patterns. In addition, patients under remission surveillance experiencing a rise from a normal baseline value often undergo confirmatory evaluation to document recurrence. Several authors have provided evidence that positive slopes in serial CA125 values among patients in clinical remission are associated with earlier documentation of disease progression with a median lead time to diagnosis of 2–4 months [3–8]. It has been further described that, although a normal reference range may be used in the interpretation of the test, individual patients typically have a “nadir” or baseline to which subsequent values may be more accurately compared. In this regard, values that are increasing but still within the “normal” reference range may be associated with the onset of recurrence [9, 10]. It is not surprising that clinicians and patients are acutely interested in these values.

The predictive nature of CA125 values has led several authors and research organizations to adopt CA125-based response and progression criteria for patients undergoing surveillance and therapy. Rustin et al. first proposed a biomarker-based algorithm with two definition thresholds for response among patients receiving platinum-based frontline chemotherapy [7, 11]. When applied to two prospective studies, the criteria produced a false-positive CA125 response in just 0.3% of patients. Bridgewater et al. also subsequently tested the Rustin response criteria among patients receiving paclitaxel and cisplatin [6]. False-positive rates were 2.9% for paclitaxel and 2.2% for cisplatin in patients undergoing frontline and recurrence therapy. Gronlund et al. compared the prognostic value of CA125 for survival to the parameters outlined in the Response Evaluation Criteria in Solid Tumors (RECIST). They reported that CA125 criteria were 2.6 times better than RECIST at predicting survival and independently predictive of survival in multivariate analysis [12]. The CA125 bioassay is informative in patients with both measurable and nonmeasurable disease. Therefore, a validated algorithm could increase the number of patients eligible for clinical studies and provide a more precise estimation of the efficacy of cytotoxic therapeutics. Rustin et al. argued that CA125 response criteria are sensitive enough to be used in a standard two-stage phase II clinical study design where minimum probabilities of response are mandated for subsequent cohort accrual [13]. However, few studies have evaluated the biomarker response criteria by individual chemotherapeutics, and none have specifically evaluated the criteria among patients treated with pegylated liposomal doxorubicin (PLD), a commonly used agent in recurrent ovarian cancer. Discordance between response assessment by radiographic imaging and CA125 values has been reported with high-dose 5-fluorouracil [14]. The misclassification of response in one-third of treated cases prompted the authors to recommend caution when assessing efficacy on the basis of CA125 alone.

We have observed that serial CA125 measures in patients receiving PLD may not be as reflective of response as seen with the topoisomerase I inhibitor topotecan [15]. In this review of recurrent ovarian cancer patients, 33 of 40 topotecan-treated responders had declining CA125 values from the first cycle of therapy until the observation of response; among the 17 PLD-treated responders, only 7 had a similar trending pattern, with some demonstrating significant increases from baseline through the second cycle of therapy. We sought to confirm this observation among a large group of patients undergoing first-relapse treatment with either PLD or topotecan as part of a randomized controlled study and for whom response was determined by traditional radiographic measures and independent of CA125 values.

**METHODS**

To evaluate the trends in serial CA125 values among patients receiving PLD or topotecan, we retrospectively reviewed biomarker and response data from patients enrolled in a phase III randomized clinical study comparing single-agent PLD with single-agent topotecan. The details of this study have been reported but are briefly outlined for clarity [16, 17].

**Patients**

Eligible patients were required to have histologically confirmed recurrent or progressive epithelial ovarian carcinoma. All patients were required to have received no more than one platinum-based chemotherapy regimen for treatment and have either measurable or measurable and assessable disease. Measurable disease was defined by standard World Health Organization criteria. Assessable disease included at least one of the following: a unidimensionally measurable lesion, mass with margins not clearly defined, lesion with both diameters 0.5 cm or less on radiographic imaging, palpable lesion under 2 cm, or malignant ascites or pleural effusion in conjunction with serum CA125 values of 100 U/ml or more in the absence of cirrhosis.
Treatment Plan
In this randomized phase III study, patients were assigned randomly in a 1:1 ratio to receive either 50 mg/m² PLD intravenously every 28 days or 1.5 mg/m² topotecan intravenously daily for 5 days every 21 days. Patients were allowed to remain on therapy in the absence of progression for up to 1 year. Dose modifications were made for toxicities. Two strata were prospectively considered: platinum sensitivity, based on a 6-month definition, and tumor bulk (above or below 5 cm).

Assessment of Response and Design
Tumor measurements were obtained by radiographic imaging within 30 days of entry and after every two cycles of therapy thereafter. A physical exam, chemistries, and CA125 values were obtained before every cycle. Subsequent radiographic assessments of response were to be conducted by the same methodology as that obtained at baseline. Definition of response was based on objective tumor measurements. A complete response (CR) was defined as the complete disappearance of all measurable and assessable disease, no new lesions, and no disease-related symptoms. A partial response (PR) was defined as a 50% or greater decrease in the sum of the products of bidimensional perpendicular diameters of all measurable lesions. Progression of assessable disease and new lesions were not allowed under the designation of a PR. Any patient who achieved a CR or PR underwent repeat radiographic assessment at least 4 weeks later to confirm the response. Progressive disease was defined as a 50% or greater increase in the sum of products of bidimensionally measured lesions over the smallest sum obtained at best response, the reappearance of any lesion that had disappeared, clear worsening of any assessable disease, death or deteriorating condition, or the appearance of any new lesion or site. Patients were classified as having stable disease (SD) if they did not qualify for CR, PR, or progressive disease.

A summary of patient demographics is presented in Table 1. The primary efficacy endpoint was overall survival; 474 patients were included in the intention-to-treat (ITT) population and analyzed. The results from this study have been reported and are summarized in Table 2.

CA125 and Response Evaluation
For the purposes of the current study, CA125 values obtained at baseline and with each of the first two cycles of both PLD and topotecan were reviewed relative to a patient’s best response on therapy. Serial CA125 values were obtained from assays performed at local or central laboratories according to standard practice. The percent change from baseline by agent (PLD and topotecan) and course (first and second) was calculated for each category of best response (CR, PR, SD, and progressive disease). Tuxen et al. have documented that intrapatient variability in serial CA125 measures is approximately 24% [5]. Therefore, five categories of percent CA125 change from baseline were assigned as follows: baseline ± 10%, an increase or decrease of 11%–25%, and an increase or decrease of more than 25%.

Statistical Evaluation
Descriptive statistics such as percent change of CA125 are presented. Differences in median CA125 values by agent and response category were evaluated using the Mann-Whitney U statistic. Since CA125-value percent changes from baseline between agents is a post hoc evaluation of the primary database, no formal statistical comparisons are reported.

RESULTS
Four hundred nine of 474 patients (86%) had both baseline CA125 values and a best tumor response assessment. All CA125 analyses are based on this population of 409 subjects. CA125 values within the normal range in the presence of evaluable disease were classified as noninformative if no trends were seen during treatment. At baseline, CA125 values were similar between patients receiving PLD and topotecan (Fig. 1). At the completion of cycle 1, 177 of 204 PLD patients (87%) and 178 of 205 (87%) topotecan patients had paired measurements. After cycle 2, 167 of 204 (82%) PLD patients and all 178 topotecan patients had serial measurements for evaluation. The primary criteria for dropout before the first evaluation time point were death, toxicity, and rapid progression.

Figures 2A and 2B present the magnitude of CA125 changes by the categorical criteria outlined above. As is appreciated, CA125 values after the first cycle of therapy were widely distributed within each response category. Of note, a large percentage of patients undergoing treatment with either agent had increased CA125 from baseline after the first cycle. The number of patients with an increased CA125 from pretreatment baseline was less after the second cycle of treatment. However, at this time point, 15% of responding patients treated with PLD and 6% of responding patients treated with topotecan had CA125 values at or above their pretreatment baseline at the first radiological evaluation point (after cycle 2). Of the 40 PLD-treated patients who had a response (CR + PR), 17 (43%) had rising CA125 values following the first cycle of therapy. The magnitude of this increase was 25% or greater in one of five responders. The ratio reduced following cycle 2; however, 1 in 10 PLD responders still recorded a CA125 value that
was greater than 25% from their pretreatment baseline. Of the eight PLD-treated patients who had a CR, four had a rise in CA125 after the first cycle. All patients in this group had decreasing values at the end of cycle 2.

Of the 35 topotecan-treated patients with a response (CR + PR), 4 (11%) had an increase in CA125 after cycle 1. Similar to PLD-treated patients, the number of women with significant (25% or greater) rises in CA125 from baseline reduced after cycle 2. At this time point, two patients (both PRs) had a 25% or greater increase from baseline in their CA125 value at the time of their first radiographic evaluation (after cycle 2). Among the 10 patients with a CR in the topotecan-treated group, 2 had rising CA125 values after the first cycle; all patients in this group had decreasing values following cycle 2. Small numbers of responders in this trial precluded the interpretation of CA125 patterns by randomization strata (platinum sensitivity and tumor bulk).

DISCUSSION
This study was undertaken to evaluate the consistency in CA125 trends relative to objective response in a homogeneously defined population. These data were ideally suited to address the hypothesis, as the clinical study was (a) large—more than 470 assessable patients; (b) randomized in design—balancing known and unknown confounding variables; (c) stratified for prognostic determinants—in particular, platinum sensitivity and disease bulk; and (d) characterized by radiographically defined and verified objective response. Furthermore,

---

| Table 1. Demographics for patients enrolled in the phase III clinical studya [16, 17] |
|---------------------------------|---------------------------------|---------------------------------|
| Characteristic                  | PLD (n = 239)                  | Topotecan (n = 235)             |
| Median age, years (range)       | 60 (27–87)                     | 60 (25–85)                      |
| Baseline Karnofsky PS, n (%)    |                                 |                                 |
| <80                             | 39 (16.3)                      | 37 (15.7)                       |
| ≥80                             | 199 (83.3)                     | 195 (83.0)                      |
| Prior platinum + taxanes, n (%) | 177 (74)                       | 169 (72)                        |
| Platinum-sensitive, n (%)       | 109 (45.6)                     | 110 (46.8)                      |
| Platinum-refractory, n (%)      | 130 (54.4)                     | 125 (53.2)                      |
| Disease bulk (>5 cm), n (%)     | 108 (45.2)                     | 105 (44.7)                      |
| Median DFI, months (range)      | 7.0 (0.9–82.1)                 | 6.7 (0.5–109.6)                 |

aSlight differences in demographic patient numbers are reported between [16] and [17]. Abbreviations: DFI, drug-free interval; PLD, pegylated liposomal doxorubicin; PS, performance status.

| Table 2. Summary of primary and secondary endpoints reported in the phase III clinical study [16, 17] |
|---------------------------------|---------------------------------|-------------------------------|
| Endpoint                        | PLD (n = 239)                  | Topotecan (n = 235)           | p value |
| Response rate, %                | 19.7                           | 17.0                          | .390    |
| Median PFS, weeks               | 16.1                           | 17.0                          | .095    |
| Median OS,a weeks               | 62.7                           | 59.7                          | .05     |

aMedian OS was updated and reported in [16]. Abbreviations: OS, overall survival; PFS, progression-free survival; PLD, pegylated liposomal doxorubicin.
robust data with measurable disease were available to correlate with biomarker values. The results of this analysis demonstrate that PLD- and topotecan-treated patients who achieve an objective response may have increasing CA125 values after the first cycle of therapy and that 15% of PLD-treated patients will have CA125 values increased from baseline after two cycles of therapy. Although most responders were ultimately observed to have decreasing CA125 values, 6% of topotecan-treated patients and 10% of PLD-treated patients were found to have a rise of greater than 25% from baseline following cycle 2. This is an important observation in that in the absence of findings on physical examination or radiographic imaging, a decision to continue a chemotherapeutic is based almost entirely on the trends in biomarker values.

Clinically, the serial decline in sequential CA125 determinations is reassuring to both clinicians and patients and is often reflective of response to therapy. However, trending patterns such as a rise from baseline after cycle 1 or cycle 2 are more difficult to interpret and usually instigate radiographic confirmation. In this study, the proportion of responders with CA125 values above baseline decreased from cycle 1 to cycle 2, suggesting that the slope of the latter two determinations was negative relative to cycle 1. In general, this too is reassuring to clinicians and patients. However, the current study demonstrates that the early patterns of change in the recurrent population may not accurately reflect ultimate best response and may be chemotherapeutic-dependent. This supports the frequent recommendation to administer at least two cycles of any new therapeutic, in the absence of limiting toxicity, to determine its efficacy. Several phase II studies have clearly documented that the median number of courses administered at best response ranges from two to four, particularly with topotecan and PLD [18, 19]. The median number of cycles administered to best response in this study was two (median 57 days) for both topotecan and PLD. We did not continue the analysis of CA125 trends beyond the second infusion because the trends were often reversed by the onset of progressive disease. However, the recommendation to administer two to four cycles of therapy seems appropriate based on our observations at the conclusion of cycle 2. Similar recommendations are outlined by the National Comprehensive Cancer Network [20].

The findings reported herein confirm earlier observations among a less homogeneous population of recurrent ovarian cancer patients. In that study of 120 recurrent ovarian or primary peritoneal cancer patients, Gossner et al. reported that 59% of PLD-treated and 18% of topotecan-treated patients who achieved a clinical response had increasing CA125 values during the first two cycles before meeting criteria for response at a median of three cycles of treatment (both agents) [15]. In that study, a 20% reduction in CA125 by cycle 2 predicted response in 85% of topotecan-treated patients; a 10% or more reduction in CA125 identified 71% of PLD-treated patients who achieved a response. However, the widely variant CA125 trends in cycle 1, particularly among PLD-treated patients, prevented the development of a rule for early identification of response.

The trends in CA125 values observed for PLD-treated patients in this study may be specific to this agent. Rustin et al. proposed that CA125 trending rules could be used as a surrogate of objective response in phase II novel chemotherapeutic clinical studies with early stopping rules [13]. In that study, a previously validated response criterion was applied to 14 drugs studied in 25 treatment cohorts from 19 clinical studies. Chemotherapeutics evaluated in this series were commonly used agents in ovarian cancer, including altretamine, paclitaxel, docetaxel, etoposide, gemcitabine,
cisplatin, oxaliplatin, and topotecan. In 20 of 25 treatment cohorts, response by CA125 criteria was concordant with objective response. Biomarker response was defined by a 75% or greater reduction in CA125 from baseline over three samples or a 50% or greater reduction in CA125 over four serial determinations.

Adjustments for missing values, extra determinations, varying upper limits of normal, and natural variations in CA125, however, complicate the algorithm and require mechanical computation. The Gynecological Cancer Intergroup has recommended a simplified variant of these criteria for patients undergoing treatment of relapsed disease, defining response as a 50% reduction in CA125, confirmed by a third value determined a minimum of 3 weeks later [21]. The baseline value must be twice the upper limit of normal. These criteria are useful and highlight the need for serial (or trending) determinations to make an accurate estimation of clinical activity. However, whether these or other biomarker-defined criteria are appropriately applied to differing cohorts of serially treated patients, such as acquired platinum and/or taxane resistance or in the third or more line of chemotherapy administration, is unknown. In addition, neither criterion is recommended for patients with baseline CA125 values in the normal range.

In this analysis, we chose to categorize the magnitude of CA125 to account for the analytical imprecision of the test previously documented by Tuxen et al. Among 26 patients in clinical remission, this group determined that the analytical imprecision of serial values was approximately 12%, and intrapatient variability averaged 24% [5]. Interpatient variability was approximately 43%. Based on these findings, the investigators were able to mathematically determine the critical thresholds for response/progression in both those with normal (one-sided) or elevated (two-sided) CA125 baseline values. In a follow-up study validating the model, Tuxen et al. determined that the efficiency of serial biomarker values for documenting progression among 255 stage IC-IV ovarian cancer patients undergoing platinum-based primary chemotherapy in the North Thames Ovary Trial was 92% [22]. The mean lead time to objective progression was 35 days. Prospective evaluation of this and other biomarker-defined response criteria are being performed for validation. Although not universally accepted in the cooperative group setting, biomarker assessment of response has been advocated by the European Group on Tumor Markers and is being used in some current phase II clinical studies [1].

This study demonstrates that early changes in CA125 values may not reflect the ultimate clinical response as determined by radiographic measures. The patterns of these changes were different between the agents being tested, and it was not uncommon to see persistently elevated CA125 values after cycle 2 among responders. Caution should be exercised when using CA125 to assess the response to PLD during the first two cycles of treatment.

**DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST**

W.R. and S.S. own stock in and have served as officer or member of the board for Johnson & Johnson. J.B. and A.G. have acted as consultants for Johnson & Johnson. A.G., T.J.H., W.R., and S.S. have received support from Johnson & Johnson. R.L.C. and T.J.H. have acted as consultants for GlaxoSmithKline and Johnson & Johnson. R.L.C., A.G., and T.J.H. have received support from GlaxoSmithKline.

**ACKNOWLEDGMENTS**

This work was presented at the 37th Annual Meeting of the Society of Gynecologic Oncologists, March 22–26, 2006, Palm Springs, CA, USA.

**REFERENCES**


