

Health-related quality of life in recurrent platinum-sensitive ovarian cancer—results from the CALYPSO trial

M. Brundage^{1,2*}, M. Gropp³, F. Mefti⁴, K. Mann⁵, B. Lund⁶, V. GebSKI⁷, G. Wolfram⁸, N. Reed⁹, S. Pignata¹⁰, A. Ferrero¹¹, C. Brown¹², E. Eisenhauer^{1,2} & E. Pujade-Lauraine¹³

¹NCIC Clinical Trials Group; ²Department of Medical Oncology, Queen's University, Kingston, Canada; ³Department of Medical Oncology, AGO, Düsseldorf, Germany; ⁴Department of Medical Oncology, GINECO, Saint-Cloud, France; ⁵Department of Medical Oncology, National Health and Medical Research Council Clinical Trials Centre, Sydney, Australia; ⁶Department of Oncology, Aalborg Hospital, Aalborg, Denmark; ⁷National Health and Medical Research Council Clinical Trials Centre, University of Sydney, Sydney, Australia; ⁸Department of Gynecology, Arbeitsgemeinschaft Gynaekologische Onkologie-Austria, Ried, Austria; ⁹Beatson Oncology Centre, Gartnavel General Hospital, Glasgow, UK; ¹⁰Multicenter Italian Trials in Ovarian Cancer (MITO), Napoli; ¹¹Academic Division of Gynecologic Oncology, Mauriziano Hospital, Torino, Italy; ¹²NHMRC Clinical Trials Centre, Sydney, Australia; ¹³Department of Oncology, Hôpital Hôtel-Dieu, Paris, France

Received 20 July 2011; revised 14 November 2011; accepted 15 November 2011

Background: In the CALYPSO trial, carboplatin–pegylated liposomal doxorubicin (CD) demonstrated superior therapeutic index versus carboplatin–paclitaxel (CP) in patients with recurrent ovarian cancer. This paper reports the health-related quality of life (HRQoL) findings.

Materials and methods: HRQoL was measured with the EORTC QoL-QC30 questionnaire and OV28 ovarian cancer module. Mean change scores from baseline in HRQoL subscales (five functional scales and global health status) in each arm and the proportion of patients improved or worsened were calculated every 3 months until 12 months.

Results: Compliance was 90% at baseline and 76%, 64%, 57% at 3, 6, and 9 months, respectively. Baseline HRQoL showed already impaired global scores (mean 62/100) and considerable symptom burden (90% of patients reporting nonzero scores). Global QoL and abdominal symptom scores improved over time in both arms; at 6 months, 36% of patients met criteria for improved symptoms. Treatment with CD resulted in less peripheral neuropathy (9.8 versus 24.2), fewer other chemotherapy side-effects (9.5 versus 16.2), and less impact on body image (3.8 versus 10.4) versus CP (all $P < 0.02$) at 6 months.

Conclusions: These patient-reported outcomes confirm the overall lower toxicity of CD versus CP. The improved disease-related outcomes achieved with CD were not at the expense of QoL.

Key words: carboplatin, paclitaxel, pegylated liposomal doxorubicin, platinum sensitive, recurrent ovarian cancer, quality of life

introduction

The management of recurrent ovarian cancer (ROC) is an ongoing topic of clinical research designed to guide practice. For patients with ovarian cancer who initially present with advanced disease, standard therapy includes a combination of maximal cytoreductive surgery with subsequent chemotherapy [1]. However, despite high response rates to this initial management, many patients experience relapse. Such patients are not curable, hence, the goals of therapy for disease

recurrence focus on improving both length of life and quality of life (QoL) [2].

Patients with ROC are conventionally classified as having disease i.e. either 'platinum sensitive', when relapse occurs >6 months after completion of initial platinum-based therapy, or 'platinum resistant/refractory', when time from treatment completion to relapse is <6 months [2]. In the setting of platinum-sensitive ROC, combination chemotherapy with carboplatin and paclitaxel (CP) was considered the standard of care when the present clinical trial was developed (2005) [2]. This collaborative, multinational randomized phase III noninferiority trial, CALYPSO, tested the efficacy and safety of pegylated liposomal doxorubicin and carboplatin (CD) compared with standard CP in 976 patients with platinum-sensitive relapsed or recurrent ovarian cancer, concluding that

*Correspondence to: Dr M. Brundage, Division of Cancer Care and Epidemiology, Queen's Cancer Research Institute, Level 2, 10 Stuart Street, Kingston, ON K7L 3N6, Canada. Tel: +1-613-533-6895; Fax: +1-613-533-6794; E-mail: michael.brundage@krcc.on.ca

treatment with CD was associated with longer progression-free survival (PFS) and less overall toxicity [3].

This report summarizes the QoL findings of CALYPSO and places these findings in the context of other primary and secondary trial outcomes. With a median follow-up of 22 months, PFS was prolonged with CD [hazard ratio (HR) 0.821, 95% confidence interval (CI) 0.72–0.94, $P = 0.005$] versus CP; median PFS was 11.3 months with CD versus 9.4 months with CP. Treatment with CD was also shown to have less grade >2 non-hematologic toxic effects leading to early treatment discontinuation. Specifically, there was less hypersensitivity to carboplatin (15% versus 6%, $P < 0.01$), less frequent alopecia (84% versus 7%), and less grade ≥ 2 sensory neuropathy (27% versus 5%) compared to CP. In contrast, CP was associated with lower rates of other toxic effects, including hand-foot syndrome (HFS; 12% versus 2%), nausea (35% versus 24%), and mucositis (14% versus 7%) [3].

A systematic review of previously reported clinical trials was used to develop evidence-based practice guidelines addressing the optimal choice of chemotherapy for ROC [2]. This review, which was published in 2007, included five trials [4–8] that exclusively accrued patients with platinum-sensitive ROC, two of which demonstrated the superiority of combination CP over single-agent carboplatin [5, 6]. One trial showed the combination of carboplatin and gemcitabine yielded significantly higher response rates and longer PFS compared with carboplatin alone [4]. None of these trials showed differences in overall QoL outcomes between randomization groups [2]. Given the advantage in PFS demonstrated for CD in the CALYPSO study [3], the QoL outcomes of the trial are highly relevant to inform treatment decision making in the management of patients with platinum-sensitive ROC.

methods

The details of this GCIG phase III trial and summary of the overall scientific basis of the study have been reported previously [3]. Briefly, the trial accrued 976 patients, of whom 467 were randomized to CD and 509 to CP. A centrally randomized, open-label noninferiority design was used to test whether the combination of pegylated liposomal doxorubicin (30 mg/m² i.v. on day 1) and carboplatin (AUC 5 based on glomerular filtration rate calculated from serum creatinine values) administered i.v. at 4-week intervals was no >23% inferior in PFS compared with paclitaxel (175 mg/m² i.v. on day 1) and carboplatin (AUC 5 i.v. on day 1) at 3-week intervals. Random assignment was carried out in permuted blocks of six, and patients were stratified based on therapy-free interval from last chemotherapy (6–12 versus > 12 months), measurable disease (yes versus no), and center. In the absence of unacceptable toxicity or disease progression, patients were treated for a total of six courses of therapy or until progression.

patients

Eligible patients were over 18 years old with a histologically confirmed diagnosis of ovarian, fallopian tube, or extra-ovarian papillary serous cancer, and with disease progression occurring >6 months after treatment with a first- or second-line platinum-based chemotherapy regimen. Other key eligibility criteria included an Eastern Cooperative Oncology Group (ECOG) performance status of two or less; life expectancy of at least 12 weeks; and adequate bone marrow, renal, and hepatic function. Patients with preexisting neuropathy grade >1 were excluded. The majority of

patients completed six cycles of chemotherapy (396 and 391, CD and CP, respectively), the two most frequent causes of treatment discontinuation being toxicity or progressive disease. The baseline characteristics of randomized patients have been reported previously [3]; briefly, the majority of patients had ECOG performance status of zero or one (95%), stage III/IV disease (85%), with one previous course of chemotherapy (85%).

toxicity and health-related QoL outcome assessment

Toxicity was assessed using the National Cancer Institute–Common Toxicity Criteria for adverse events (NCI–CTCAE version 3.0). All patients were expected to complete the study health-related QoL (HRQoL) assessments. HRQoL was measured by the EORTC QLQ-C30 core instrument [9], supplemented by the OV28 ovarian cancer-specific module [10]. The core instrument has been validated in the general cancer population and specifically in ovarian cancer patients [11] and consists of a global QoL scale, five functional scales, and nine symptom scales. The ovarian cancer-specific module consists of seven symptom scales; supplemental Table S1 (available at *Annals of Oncology* online) illustrates the subject content of questionnaire items relevant to each of the seven scales. As is the case for the core questionnaire, the ovarian module has been shown to be valid, reliable, and sensitive to change in patients with ovarian cancer receiving chemotherapy [10]. The rationale for choosing these instruments was based on their validated status, their content scope that addresses global, functional, and symptom domains of high relevance to patients with ROC, and their previous use in this cancer setting [10]. The study protocol required assessment of HRQoL at randomization (baseline) and at the 3-, 6-, 9-, and 12-month assessments. HRQoL assessment was not carried out after documented disease progression; therefore, no additional cancer therapies were given in those patients who provided HRQoL data at follow-up times.

statistical considerations

Analyses of the HRQoL data were undertaken in keeping with the general analysis approach of the NCIC Clinical Trials Group (NCIC CTG) [12]. In these analyses, baseline completion rates are calculated for each treatment arm, and compliance rates at subsequent time points are defined as the proportion of patients completing HRQoL assessment conditional on baseline HRQoL submission and no disease progression [12]. Between treatment arms, Student's *t*-tests were used to test for statistically significant differences in mean baseline HRQoL scores.

Mean change scores at each time point were calculated for all patients with baseline HRQoL data. Statistical differences between arms were tested with *t*-tests at each time point. The multidimensional aspect of the HRQoL instruments allowed two main HRQoL hypotheses to be evaluated, namely, changes in HRQoL scores with treatment would reflect either improvement in ovarian cancer-related symptoms due to disease response or adverse effects of treatment-related toxicity. HRQoL data were typically not collected in patients once their disease had progressed.

HRQoL 'response' analysis represents the proportion of patients who experience a clinically significant change in HRQoL score on a particular functional or symptom scale compared with their baseline score. This approach has been promoted as a preferred way of illustrating the impact of treatment on HRQoL, in that the proportion of patients improving in HRQoL is seen as analogous to the proportion of patients with a tumor complete or partial response [13]. For the purposes of this analysis, a previously validated minimal important difference of 10 points (on a 100-point scale) was used to define the proportion of patients with a clinically significant change [12, 14]. This magnitude of difference is clinically conservative, with some estimates of minimal important change being as low as 5–7 points, depending on the underlying variance of the data [14]. Patients with an HRQoL change score of ≥ 10 in a direction reflecting

deteriorating functioning or increased symptoms at a given time point (compared with baseline) were classified as 'worsened', whereas those with an change for the better of ≥ 10 points were classified as 'improved'. Those with HRQoL scores at a given time point within 10 points of baseline score were classified as 'stable'. Chi-square test of proportions was used to determine statistical significance of the proportion of patients in each category based on these assignments. The focus of this report is based on 3-month response scores because less missing data were evident at this time point, and patients on treatment in both arms had received three or more cycles of chemotherapy by this time.

All analyses of HRQoL end points were restricted to those patients who completed baseline HRQoL assessments and had at least one HRQoL form completed in follow-up. Given the extent of missing assessments in follow-up, between-arm comparisons of average difference in HRQoL scores over time, adjusted for baseline levels, were also estimated using generalized estimating equations. Given the drop of 12-month compliance rates to fewer than half of expected rates (see below), the report was limited to data collected up to the 9-month follow-up period.

Missing value patterns (missingness) were investigated by examining associations between baseline variables significantly associated with progression [3] (these being therapy-free interval, pretreatment CA125, treatment, measurable disease status, and involved disease sites) and missing HRQoL outcomes, using multivariate logistic regression.

results

compliance rates

Overall, 458 (90.0%) patients of those randomized to CP and 421 (90.1%) patients of those randomized to CD completed baseline HRQoL assessments. Compliance (CP versus CD) at 3 months was 73.5% versus 79.3%, at 6 months was 60.3% versus 68.3%, at 9 months was 56.1% versus 57.3%, and at 12 months was 49.7% versus 50.6%, among the patients that had completed baseline HRQoL.

baseline HRQoL and changes in mean HRQoL scores over time

Table 1 lists the baseline HRQoL scores for the functional and symptom scales of the core QLQ-C30 instrument as well as the seven symptom scales of the OV28 module. As expected by the randomization process, baseline scores were well balanced between treatment arms with no significant between-arm differences detected.

Table 1 also illustrates the change in HRQoL scores on each domain and symptom scale at 3 and at 6 months, respectively. Statistically significant between-arm differences at 3 months were seen for the global and physical function scales; however, the magnitude of the differences between arms was clinically modest (4.8 points for global HRQoL and 3.7 points for physical function). These changes are graphically illustrated in Figures 1 and 2, which show the change in mean scores over time for the core instrument function scales and for selected symptom scales, respectively.

patient response analysis

Figure 3 illustrates the proportion of patients in each response category (improved, stable, or worsened) for each HRQoL functional domain or symptom scale at 3 months after randomization. Significant differences between treatment arms

in these response proportions were seen; a higher proportion of patients treated with CD were seen to have clinically significantly improved global QoL scores, and a smaller proportion of patients treated with CD were seen to have worsened pain, dyspnea, insomnia, neuropathy, body image, sexual functioning, and other chemotherapy side-effect symptom scores compared with those treated with CP. Smaller proportions of patients treated with CP were seen to have worsened sexual functioning and nausea/vomiting scores compared with those treated with CD.

analysis of HRQoL with a focus on abdominal and gastrointestinal symptoms

The randomization of patients on the CALYPSO trial took place either on the basis of tumor marker progression or measurable disease ($n = 533$). However, given the nature of the study population, it was not unexpected that the majority of patients (89.7%) reported abdominal or gastrointestinal (GI) symptoms on the OV28 symptom scale at baseline. An analysis of the patients who were symptomatic at baseline (i.e. before the development of side-effects from randomization-allocated chemotherapy) was undertaken. Overall, 409 (48%) of patients had symptom scores of up to and including 30/100, and 356 (42%) had symptom scores $>30/100$ at baseline. The detailed nature of the HRQoL data reported on the OV28 symptom module allowed analysis of HRQoL issues that reflect the impact of symptoms that are not effectively captured by conventional toxicity outcomes, which tend to focus on adverse effects of therapy, rather than symptoms related to disease burden.

Table 2 illustrates the relationship between the severity of patients' baseline symptoms, and whether or not each patient had measurable disease at study entry (baseline). It is noteworthy that a large number of patients classified as not having measurable disease were nonetheless symptomatic. Examining HRQoL responses in asymptomatic patients (Figure 4) reveals that many symptomatic patients reported improved symptoms after 3 months: over one-third of patients with nonzero baseline scores improved in both CP and CD arms, and over 60% of patients with baseline symptom scores >30 responded to treatment (no significant between-arm differences were present).

impact of missing data

Seventy-nine percent (361/458) of patients in the CP arm and 84% (352/421) of patients in the CD arm had QoL data at baseline and at another time point in the study. These proportions were not statistically significantly different, and a multivariate logistic regression examining the association between missingness and therapy-free interval, CA125, treatment, measurable disease, and involved disease sites was not significant. This suggests that the pattern of missing QoL data does not substantially alter the conclusions of the analyses.

discussion

The CALYPSO study was designed as a noninferiority trial that demonstrated not only noninferiority but also superiority of

Table 1. Baseline HRQoL mean scores and change in mean scores

Item/domain	Baseline scores				3-month change ^a				P value	6-month change ^a				P value
	CP		CD		CP		CD			CP		CD		
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)		N	Mean (SD)	N	Mean (SD)	
Functional scales scores														
Physical functioning	452	79.5 (20.7)	414	79.8 (20.1)	313	-7.4 (18.1)	309	-3.7 (18.8)	0.01	228	-1.1 (20.6)	242	-2.8 (19.1)	0.36
Role functioning	447	72.6 (30.4)	413	72.3 (31.9)	310	-9.2 (31.2)	305	-4.5 (33.8)	0.07	225	2.4 (34.7)	241	-1.7 (31.5)	0.18
Emotional functioning	447	63.4 (25.6)	410	64.4 (25.2)	308	6.5 (21.8)	303	8.3 (22.6)	0.31	225	6.8 (25.3)	239	5.1 (23.7)	0.46
Cognitive functioning	448	83.9 (20.2)	412	83.8 (20.2)	309	-6.4 (20.3)	305	-3.6 (19.1)	0.08	226	-3.1 (21.3)	242	-5.2 (22.6)	0.31
Social functioning	445	74.2 (28.2)	411	78.4 (27.2)	303	-5.2 (28.3)	304	-5.1 (28.5)	0.95	224	0.1 (31.2)	242	-4.3 (23.8)	0.09
Global health														
Global health status score	447	62.2 (23.0)	408	61.4 (24.2)	307	-2.2 (22.7)	301	2.6 (26.0)	0.01	227	4.8 (24.4)	238	2.4 (26.4)	0.31
Symptoms scales scores														
Fatigue	450	34.7 (25.7)	413	34.4 (27.5)	310	-9.4 (25.3)	306	-6.3 (27.0)	0.13	225	1.2 (25.9)	243	-4.6 (27.6)	0.02
Nausea and vomiting	449	8.0 (17.3)	413	10.9 (20.7)	309	-3.5 (22.9)	308	-8.4 (26.1)	0.01	225	3.2 (21.9)	244	-0.3 (24.9)	0.11
Pain	450	27.1 (28.4)	414	25.9 (28.2)	310	1.3 (31.2)	306	6.2 (28.7)	0.04	227	6.4 (31.3)	243	6.0 (29.3)	0.86
Dyspnea	444	17.9 (25.3)	409	19.4 (27.8)	307	-11.8 (29.6)	305	-3.2 (30.0)	<0.001	224	-4.5 (30.1)	244	-5.7 (29.4)	0.64
Insomnia	447	36.8 (32.9)	413	36.6 (32.8)	306	2.2 (31.7)	304	5.7 (31.4)	0.16	223	3.6 (30.2)	242	4.6 (32.6)	0.73
Appetite loss	445	18.7 (29.2)	413	21.5 (30.8)	305	-2.3 (31.5)	306	0.6 (30.1)	0.24	225	7.3 (27.9)	242	5.3 (31.3)	0.47
Constipation	449	22.6 (31.1)	409	23.6 (31.7)	309	-4.5 (33.2)	301	-5.5 (34.9)	0.74	227	3.6 (34.2)	239	-2.7 (34.3)	0.05
Diarrhea	447	10.3 (21.1)	409	13.5 (24.8)	307	0.5 (24.0)	300	3.4 (25.2)	0.15	227	-1.0 (24.5)	238	3.8 (22.5)	0.03
Financial difficulties	441	14.7 (27.3)	405	12.4 (24.8)	303	-4.6 (26.8)	297	-2.4 (22.6)	0.28	221	-2.6 (30.5)	239	-2.0 (22.9)	0.81
QLQ-OV28														
Abdominal/gastrointestinal symptoms	442	29.1 (22.6)	411	30.3 (24.0)	306	5.0 (22.0)	305	4.7 (21.7)	0.83	223	6.8 (23.1)	238	5.1 (23.9)	0.43
Peripheral neuropathy														
Peripheral neuropathy	434	17.7 (22.0)	402	15.3 (20.6)	300	-27.4 (26.8)	297	-6.1 (18.9)	<0.001	218	-24.2 (30.5)	233	-9.8 (20.1)	<0.001
Other chemotherapy side-effect														
Other chemotherapy side-effect	435	15.0 (14.9)	405	14.2 (15.1)	301	-24.7 (18.4)	301	-7.6 (16.8)	<0.001	219	-16.2 (19.9)	236	-9.5 (15.9)	<0.001
Hormonal/menopausal symptoms														
Hormonal/menopausal symptoms	435	26.4 (28.0)	405	24.2 (28.6)	300	-1.6 (24.2)	301	-0.6 (28.3)	0.62	219	-2.4 (28.4)	235	-2.9 (28.4)	0.84
Body image														
Body image	431	23.9 (27.6)	401	24.3 (28.0)	297	-12.2 (29.3)	292	-1.2 (28.2)	<0.001	212	-10.4 (31.4)	234	-3.8 (27.0)	0.02
Attitude to disease and treatment														
Attitude to disease and treatment	432	57.2 (28.2)	397	56.3 (28.5)	295	-0.4 (25.4)	290	1.7 (25.8)	0.32	216	0.0 (28.0)	228	1.8 (24.3)	0.48
Sexual functioning														
Sexual functioning	385	20.4 (23.5)	358	16.3 (21.7)	241	4.5 (17.8)	232	0.4 (18.5)	0.02	173	0.8 (19.7)	187	0.2 (18.5)	0.78

P values shown are from a *t*-test.

^aPositive values indicate an increase in improvement, while negative values indicate a worsening.

CD, carboplatin–pegylated liposomal doxorubicin; CP, carboplatin–paclitaxel; HRQoL, health-related quality of life; SD, standard deviation.

CD. Patients randomized to CD experienced longer PFS and less severe overall toxicity than patients randomized to CP. The HRQoL data collected support the hypothesis that the PFS benefits of CD did not come at the expense of a negative QoL impact relative to CP. In fact, the proportion of patients reporting improved global HRQoL at 3 months was significantly higher in patients randomized to CD.

The largest magnitude changes in HRQoL from baseline were seen in patient-reported symptom scores; these outcomes may reflect both disease response to treatment (treatment benefit) and/or adverse sequelae of treatment (toxicity). Baseline symptoms reported on scales assessing abdominal pain and GI symptoms, fatigue, and attitudes/concerns were presumed to be

predominantly disease related, although these symptoms potentially also reflected the impact of previous treatment. Other baseline symptoms (e.g. neuropathy or body image) were presumed to reflect persistent treatment-related toxicity from prior therapies. Both treatment arms in this study were effective in improving abdominal symptom scores (as assessed by both mean change score analysis and proportion of patients improving), whereas, neither treatment had a meaningful effect in improving mean fatigue or attitude/concern scores over time.

Treatment-related toxic effects in this study were reflected in those symptom scores that deteriorated significantly after initiation of treatment. In keeping with the primary CALYPSO

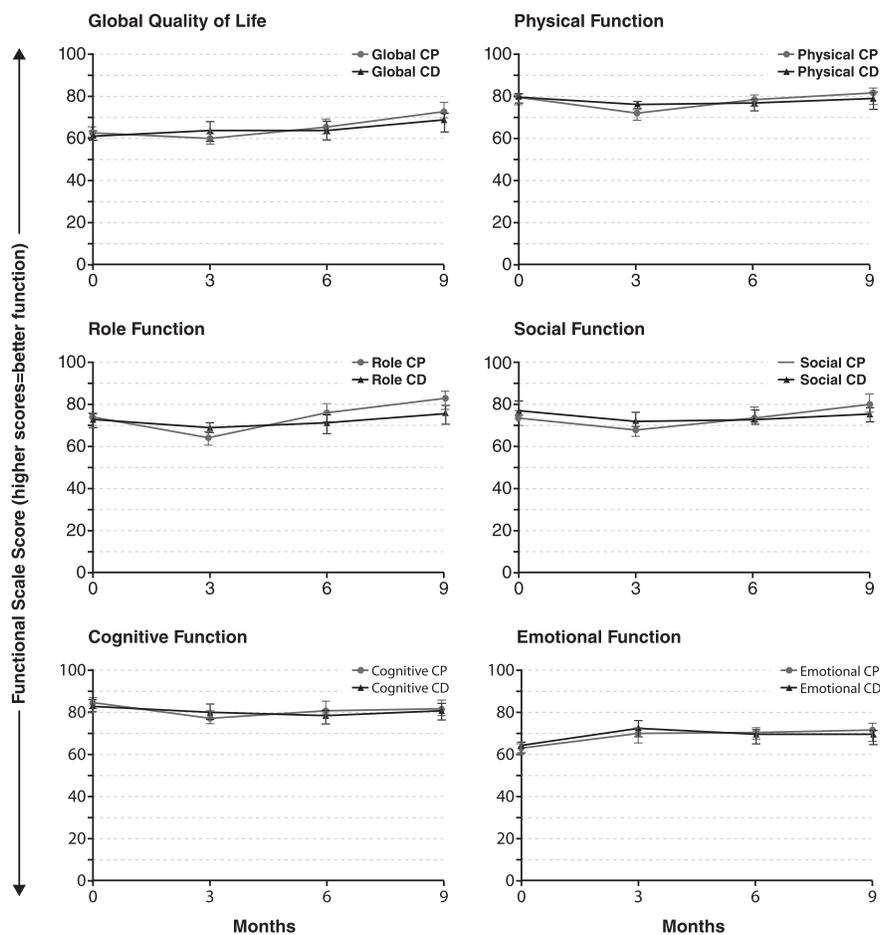


Figure 1. Baseline and changes in mean scores for health-related quality of life function scales. Error bars represent 95% confidence interval for the estimated mean values. CD, carboplatin-pegylated liposomal doxorubicin; CP, carboplatin-paclitaxel.

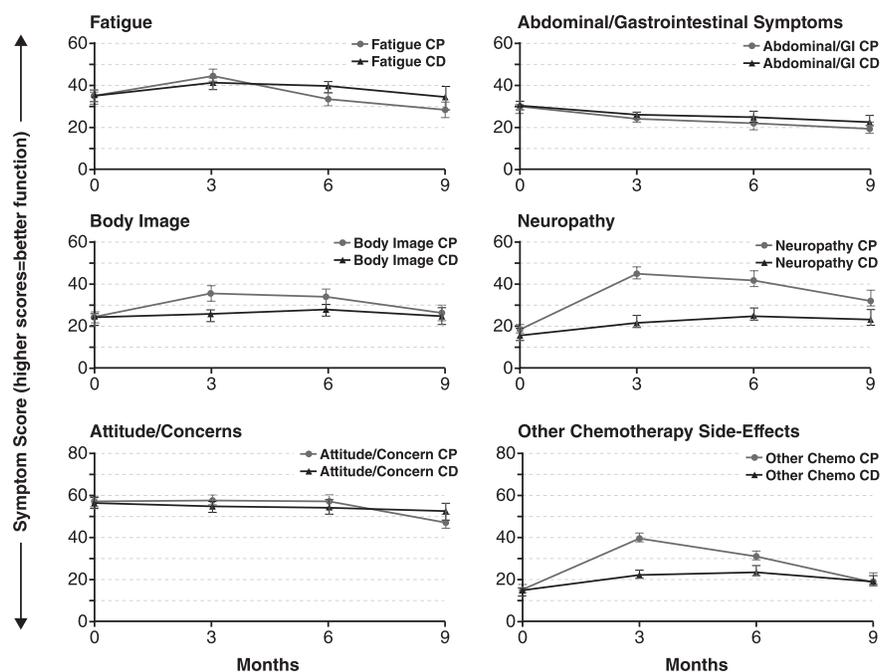


Figure 2. Baseline and changes in mean scores for selected health-related quality of life symptom scales. Error bars represent 95% confidence interval for the estimated mean values. CD, carboplatin-pegylated liposomal doxorubicin; CP, carboplatin-paclitaxel.

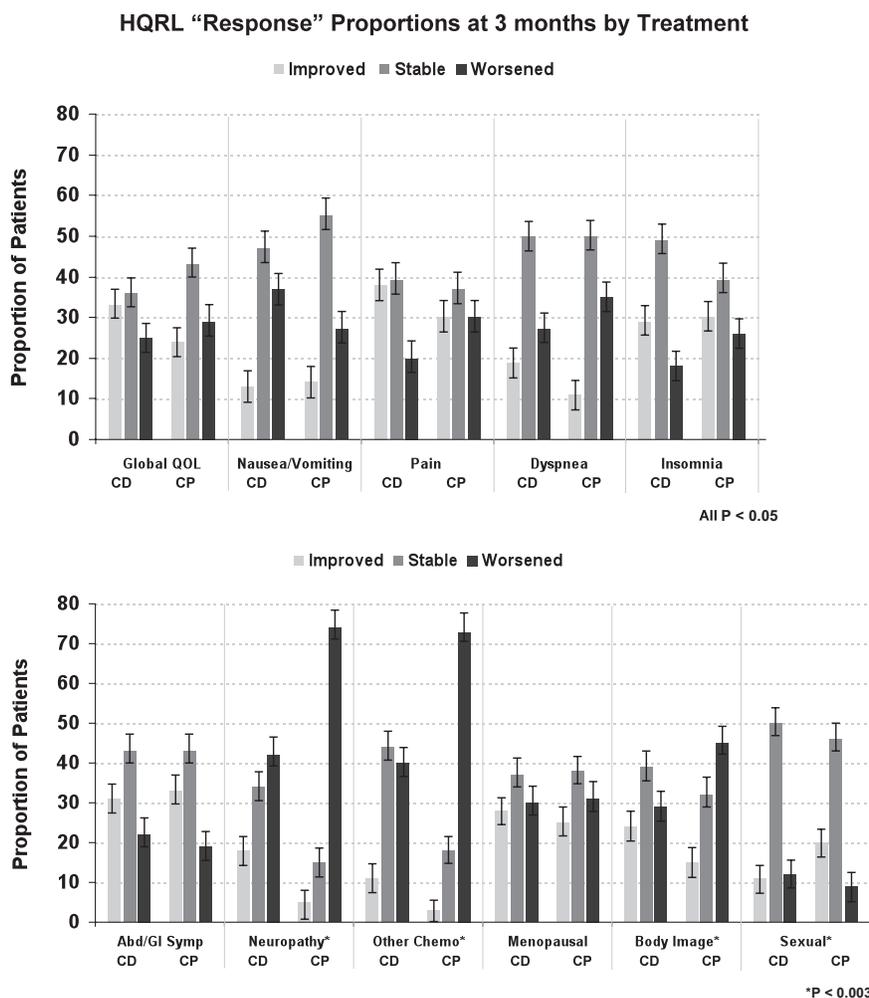


Figure 3. Proportions of patients responding to treatment by patient-reported outcome. Error bars represent 95% confidence interval for the estimated proportions. CD, carboplatin–pegylated liposomal doxorubicin; CP, carboplatin–paclitaxel; HRQoL, health-related quality of life.

Table 2. Association between baseline symptom scores and measurable disease

	Measurable disease at baseline		Total
	Present	Absent	
Abdominal pain and GI symptom score			
No symptoms	58 (11%)	30 (9%)	88 (10%)
Symptom score 1–30	274 (51%)	135 (42%)	409 (48%)
Symptom score >30	201 (38%)	155 (48%)	356 (42%)
Total	533	320	853

GI, gastrointestinal.

toxicity data, patients randomized to CP reported lower rates of nausea and vomiting worsening but also reported higher rates of neuropathy, compromised body image (reflecting alopecia and other aspects of body image), and other chemotherapy side-effects.

Despite the similar between-treatment patterns in toxicity rates and HRQoL findings, there were four important ways in which measuring QoL in the CALYPSO trial added value to the

study and its interpretation. First, the HRQoL data provide evidence that many patients derive symptom control benefit associated with treatment, in addition to the delay in disease progression afforded by chemotherapy for recurrent disease. The patient-reported outcomes are the only study end points that directly evaluate these palliative benefits.

Second, the HRQoL data provide different estimates of the magnitude of adverse aspects of treatment compared with conventional toxicity data. For some symptoms, reported toxicity rates were comparable to the proportion of patients reporting significantly worsened scores. For example, grade 2 nausea rates of 24% (CP) and 35% (CD) were comparable to the proportion of patients reporting worsened nausea after 3 months (27% and 37%, respectively). For other symptoms, HRQoL rates were somewhat higher than toxicity rates: fatigue, e.g., was reported as grade 2 in up to 40% of patients but was self-reported as worsened in 49% of patients. For neuropathy, grade 2 events were reported in 27% of CP patients, but 75% of patients in that arm reported significantly worse scores.

Third, the scope of issues identified by the HRQoL scores may help clinicians understand the nature of the issues experienced by patients in this setting, which in turn may

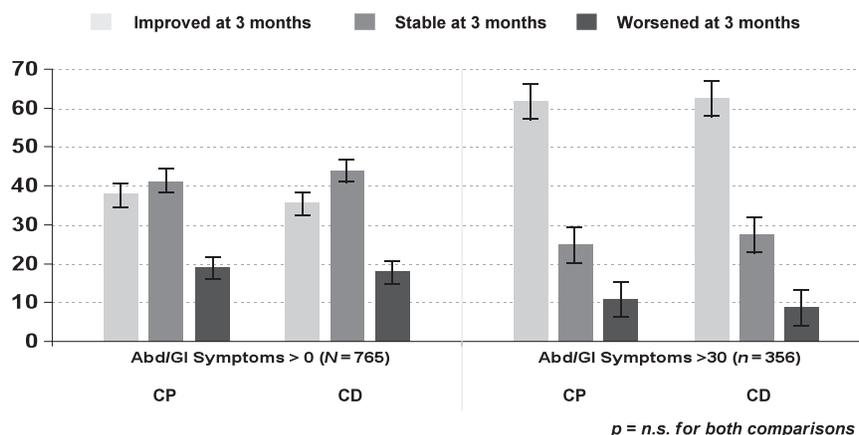


Figure 4. Abdominal and GI symptom responses by severity of symptoms as baseline. Error bars represent 95% confidence interval for the estimated proportions. CD, carboplatin–pegylated liposomal doxorubicin; CP, carboplatin–paclitaxel; GI, gastrointestinal.

provide for increased supportive care resources or better supportive care strategies. The high prevalence of fatigue, patient concerns, sexual functional impairment, and other issues relating to HRQoL may be useful in informing the development of enhanced services for ovarian cancer patients and may also indicate the potential value of routinely measuring or screening for these symptoms in clinical practice outside of a clinical trial setting [15].

Finally, scores on the core HRQoL function scales illustrate that the treatment with the less severe toxicity impact was also associated with a small but detectable improvement in global QoL scores at 3 months after randomization. While this impact is of borderline clinical significance, there is no evidence suggesting that the improvement in PFS observed with CD was at the expense of patients' overall QoL or their functioning on the main HRQoL domains.

The interpretation of the HRQoL findings in this study is not without some limitations. First, the extent of missing data likely biases estimates of mean improvement in HRQoL with time (because those patients with deteriorating HRQoL scores are more likely to be missing). This bias results in some spurious improvement in mean scores with time in both treatment arms. It is for this reason that we focused on the 3-month end point (for which data were reasonably complete) to estimate the HRQoL benefits of treatment. In addition, we point out that the proportion of patients reporting improved HRQoL status compared with baseline is not biased by the absence of scores from those patients who have deteriorated. Hence, the estimated proportion of patients improved in each domain is less subject to bias in clinical interpretation than is the average improvement across patients. Second, although the study evaluated improvement in baseline symptoms, some baseline symptoms may have been due to previous treatment or comorbid conditions that were not disease related. Our evaluation of patients with baseline scores in excess of 30/100 was undertaken to address the subset of patients who were more likely to have disease-related symptoms. Third, some HRQoL descriptive scales contain several subitems (e.g. chemotherapy side-effects, illustrated in Table 1); hence, changes in the composite scores may be due to changes in any

combination of the contributing items. Nonetheless, differences in these composite scale scores between treatments are highly clinically relevant.

conclusions

Treatment of patients with platinum-sensitive ROC with either CP or CD improves patient-reported symptoms of disease in many dimensions. Patient-reported outcomes reflecting toxicity were often higher than corresponding toxicity rates, and more treatment-related symptoms were observed with CP. Treatment with CD was associated with statistically significantly superior global and physical function scores at 3 months; however, the between-arm differences in these scores were of modest clinical significance. These patient-reported outcomes confirm the overall lower toxicity of CD compared with CP and underscore that the improved disease-related outcomes (PFS) yielded by treatment with CD were not achieved at the expense of impaired QoL.

acknowledgements

We would like to acknowledge the women who participated in this study as well as the investigators and their staff, Phillips Gilmore for technical assistance in preparation of the manuscript, and the helpful contributions of Dr Lori Brotto. Dr Brundage is a senior researcher supported by Cancer Care Ontario and the Ministry of Health and Long Term Care, ON, Canada. Study previously presented at the 2010 American Society of Clinical Oncology Annual Meeting: Brundage M, Gropp M, Mefti F et al. Health-related quality of life (HRQL) and progression-free survival (PFS) in patients with recurrent ovarian cancer: results from the CALYPSO trial. *J Clin Oncol* 2010; 28 (Suppl): 15s (Abstr 5044).

funding

This research was supported by a research grant from Merck, formerly Schering Plough Corp. (GCIG006-1a).

disclosure

All authors of this paper declare that there are no conflicts of interest involved in this paper except for SP who received honoraria from Schering Plough.

references

1. Covens A, Carey M, Bryson P et al. Systematic review of first-line chemotherapy for newly diagnosed postoperative patients with stage II, III, or IV epithelial ovarian cancer. *Gynecol Oncol* 2002; 85: 71–80.
2. Fung-Kee-Fung M, Oliver T, Elit L et al. Optimal chemotherapy treatment for women with recurrent ovarian cancer. *Curr Oncol* 2007; 14: 195–208.
3. Pujade-Lauraine E, Wagner U, Avall-Lundqvist E et al. Pegylated liposomal doxorubicin and carboplatin compared with paclitaxel and carboplatin for patients with platinum-sensitive ovarian cancer in late relapse. *J Clin Oncol* 2010; 28: 3323–3329.
4. Pfisterer J, Plante M, Vergote I et al. Gemcitabine plus carboplatin compared with carboplatin in patients with platinum-sensitive recurrent ovarian cancer: an intergroup trial of the AGO-OVAR, the NCIC CTG, and the EORTC GCG. *J Clin Oncol* 2006; 24: 4699–4707.
5. Parmar MK, Ledermann JA, Colombo N 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: 2099–2106.
6. Gonzalez-Martin AJ, Calvo E, Bover I 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: 749–755.
7. Cantu MG, Buda A, Parma G et al. Randomized controlled trial of single-agent paclitaxel versus cyclophosphamide, doxorubicin, and cisplatin in patients with recurrent ovarian cancer who responded to first-line platinum-based regimens. *J Clin Oncol* 2002; 20: 1232–1237.
8. Bolis G, Scarfone G, Giardina G et al. Carboplatin alone vs carboplatin plus epidoxorubicin as second-line therapy for cisplatin- or carboplatin-sensitive ovarian cancer. *Gynecol Oncol* 2001; 81: 3–9.
9. Aaronson NK, Ahmedzai S, Bergman B et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993; 85: 365–376.
10. Cull A, Howat S, Greimel E et al. EORTC Quality of Life Group [European Organization for Research and Treatment of Cancer]; Scottish Gynaecological Cancer Trials Group. Development of a European organization for research and treatment of cancer questionnaire module to assess the quality of life of ovarian cancer patients in clinical trials: a progress report. *Eur J Cancer* 2001; 37: 47–53.
11. Osoba D, Zee B, Pater J et al. Psychometric properties and responsiveness of the EORTC quality of life questionnaire (QLQ-C30) in patients with breast, ovarian and lung cancer. *Qual Life Res* 1994; 3: 353–364.
12. Osoba D, Bezjak A, Brundage M et al. Quality of Life Committee of the NCIC CTG analysis and interpretation of health-related quality-of-life data from clinical trials: basic approach of the National Cancer Institute of Canada Clinical Trials Group. *Eur J Cancer* 2005; 41: 280–287.
13. Guyatt G, Schunemann H. How can quality of life researchers make their work more useful to health workers and their patients? *Qual Life Res* 2007; 16: 1097–1105.
14. Sloan JA, Frost MH, Berzon R et al. Clinical Significance Consensus Meeting Group. The clinical significance of quality of life assessments in oncology: a summary for clinicians. *Support Care Cancer* 2006; 14: 988–998.
15. Greenhalgh J. The applications of PROs in clinical practice: what are they, do they work, and why? *Qual Life Res* 2009; 18: 115–123.