Using the Expected Survival to Explain Differences Between the Results of Randomized Trials: A Case in Advanced Ovarian Cancer

By Marc Buyse, Tomasz Burzykowski, Mahesh Parmar, Valter Torri, George Omura, Nicoletta Colombo, Chris Williams, Pierfranco Conte, and Jan Vermorken, for the International Collaborative Ovarian Neoplasm Collaborators and the Ovarian Cancer Meta-Analysis Project

Purpose: A meta-analysis of randomized trials in advanced ovarian cancer showed a longer survival with cyclophosphamide, doxorubicin, and cisplatin (CAP) than with cyclophosphamide and cisplatin (CP; P = .009). In contrast, the results of the large International Collaborative Ovarian Neoplasm Study (ICON2) showed no survival difference between CAP and carboplatin (P = .98). In this article, we show how these discrepant results can be reconciled through the estimation of expected survival curves.

Materials and Methods: A proportional hazards model, fitted to the meta-analysis data, was used to construct the expected survival curve for each treatment arm of the ICON2 trial. Expected survival curves were compared with observed survival curves in the ICON2 trial at all time points using a nonparametric test.

Results: The prognostic model for survival obtained in the meta-analysis included extent of residual disease, age, histologic grade, and International Federation of Gynecology and Obstetrics stage. When this model was applied to the ICON2 data, there was no difference between the expected and observed curves in the CAP arm. In contrast, the observed survival curve for carboplatin was far superior to the expected survival curve for CP (P < .01).

Conclusion: These analyses provide indirect evidence that better results are achieved with carboplatin alone at an optimally tolerated dose, compared with the CP combination at a cisplatin dose of 50 to 60 mg/m^2. The expected survival may provide valuable insight when direct comparisons between randomized groups yield discrepant results across different studies.

J Clin Oncol 21:1682-1687. © 2003 by American Society of Clinical Oncology.

O FTEN, THE results of different studies addressing similar therapeutic questions yield conflicting results, which makes it difficult to draw any definite conclusion about the therapies under investigation. For instance, one experiment may show survival benefit from a certain therapy, whereas another similar experiment may fail to do so. Meta-analysis has been advocated as a way to combine evidence from several experiments addressing the same therapeutic question.[1] Critics of meta-analysis have pointed out that results of trials were sometimes contradicted by those of subsequent large confirmatory trials.[2,3] Although this observation invalidates neither the meta-analysis nor the randomized trial, contradictory results are unsettling and deserve to be investigated further.[4,5] In this article, we consider such a situation in advanced ovarian cancer.

A meta-analysis was undertaken in 1989 to evaluate the role of anthracyclines in the treatment of women with advanced ovarian tumors. Six randomized trials compared a standard regimen consisting of cyclophosphamide and cisplatin (CP) with cyclophosphamide and cisplatin plus the anthracycline doxorubicin (CAP). The meta-analysis of these trials, which is based on individual patient data supplied by the principal investigators, showed that CAP yielded a higher rate of tumor response and a longer survival than CP.[6] These results, which seemed to warrant the use of anthracyclines in the treatment of advanced ovarian cancer, led to a multinational randomized trial that was started in 1991 to compare CAP (the better regimen in the meta-analysis) with carboplatin alone. In this trial, known as the Second International Collaborative Ovarian Neoplasm Study (ICON2), carboplatin was chosen instead of CP because it was believed that an optimally tolerated dose of a single-agent platinum would give results similar to those of platinum-based combinations.[7] In addition, carboplatin (the most widely used platinum salt in the United Kingdom) was likely to have an efficacy similar to that of cisplatin, with far less toxicity.[8]

An update of the meta-analysis, performed after a median follow-up of more than 10 years in the four larger trials, confirmed the survival benefit of CAP over CP (hazard ratio [HR] = 0.84; P = .009).[9] In contrast, the results of the ICON2 trial, after a median follow-up of approximately 2.5 years, showed no survival difference between CAP and carboplatin (HR = 1.0; P = .98).[10] An obvious explanation for these apparently conflicting results is that single-agent carboplatin, at

Table 1. Main Characteristics of the Trials in the Meta-Analysis[9] and of the ICON2 Trial[10]

Characteristic	Meta-Analysis	ICON2
No. of trials	4	1
No. of countries	3	9
No. of patients	1,197	1,526
Accrual period	April 1980-November 1986	January 1991-July 1996
Median follow-up of alive patients, months	126	30
Treatment doses in CAP arm, mg/m^2	Cyclophosphamide, 500-650	Cyclophosphamide, 500
	Doxorubicin, 40-50	Doxorubicin, 50
	Cisplatin, 50-60	Cisplatin, 50
Treatment doses in CP or carboplatin arm	Cyclophosphamide, 500-1,000 mg/m^2	Carboplatin, 5 (GFR + 25) mg*
	Cisplatin, 50-60 mg/m^2	

Abbreviations: CAP, cyclophosphamide, doxorubicin, cisplatin; CP, cyclophosphamide, cisplatin; GFR, glomerular filtration rate; ICON2, Second International Collaborative Ovarian Neoplasm Study.
*Formula from Calvert et al.[11]

an optimally tolerated dose, is better in terms of overall survival than cyclophosphamide combined with cisplatin at the dose of 50 mg/m^2. However, there are alternative possible explanations for the discrepancy, including differences in prognostic mix between the two patient series, imbalances with respect to important prognostic factors in either series, shorter follow-up time in the ICON2 trial, better treatments after disease progression in the more recent ICON2 trial, unreliability of the results of either the meta-analysis or the ICON2 trial, or simply the play of chance. Therefore, without additional analyses, there may be doubts about the proper interpretation of the discrepancy. In this article, we show how the discrepancy can be explored through the estimation of the expected survival that takes into account patient prognostic features in both the meta-analysis and the randomized trial.

MATERIALS AND METHODS

Patient Data

We used individual patient data from the meta-analysis and from the ICON2 trial, both of which are described in detail in previous publications.[6,9,10] All patients were considered in the present article, whether eligible or not and whether properly treated or not. Items requested for every patient included baseline clinical characteristics (patient identification, institution, date of random assignment to treatment, age, performance status [not available in ICON2], extent of residual disease after debulking surgery, histologic cell type, histologic grade (cell differentiation), and International Federation of Gynecology and Obstetrics [FIGO] stage), treatment assigned by randomization, and the outcome of interest (date of death or last visit and survival status). Survival time was considered from the day of random treatment assignment to the day of death regardless of the cause of death. Table 1 compares some characteristics of the meta-analysis[9] with the ICON2 trial.[10] The dose of carboplatin in the ICON2 study was determined using the area under the curve method of Calvert et al[11] to find the optimally tolerated dose for each patient.

Statistical Analysis

All analyses were based on individual patient data using an intent-to-treat approach. Observed survival curves were estimated using the product-limit method and compared using the log-rank test, as in the original publications on the meta-analysis and the ICON2 trial.[6,10] For the meta-analysis, the log-rank test was stratified by trial. A proportional hazards regression model was fitted on the data of the meta-analysis, starting from the full model with all baseline characteristics and treatment, and eliminating covariates from the model using a step-down procedure.[12] Patients with missing values for the

retained covariates were excluded from the estimation of the model parameters. The model was then used to construct the expected survival curve for each treatment arm of the ICON2 trial; that is, the curve one would have observed if the estimated model were correct and the mortality rate of the corresponding treatment arm of the meta-analysis were operating.[13,14] The expected survival curve was calculated by integrating the expected hazard rate over time, with the expected hazard rate at time t defined as the average of the hazards of all individuals who were still under observation at that time.[14,15] The expected survival curve was compared with the observed survival curve at each time point through a nonparametric test.[16] A Kolmogorov-Smirnov type test was used to assess the overall significance of any difference between the observed and expected curve.[17] The level of statistical significance was set at $P = .01$.

RESULTS

Observed Survival Curves

Figure 1 shows the survival curves for CP and CAP in the meta-analysis. The difference between the curves was statistically significant (HR = 0.84; $P = .009$). Figure 2 shows the survival curves for carboplatin and CAP in the ICON2 trial. There was no apparent difference between the curves (HR = 1.00; $P = .98$).

Prognostic Model for Survival

Table 2 shows the distribution of prognostic factors in the meta-analysis and the ICON2 trial. The prognostic mix differed
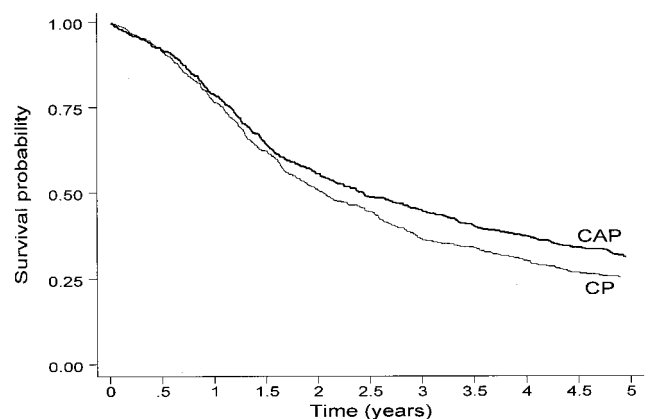
Fig 1. Observed survival in the cyclophosphamide, doxorubicin, and cisplatin (CAP) and the cyclophosphamide and cisplatin (CP) arms of the meta-analysis ($P = .009$; thick line, CAP; thin line, CP).
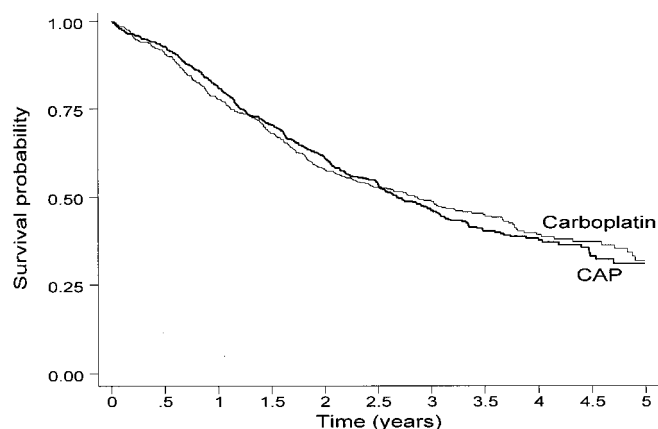
Fig 2. Observed survival in the cyclophosphamide, doxorubicin, and cisplatin (CAP) and carboplatin arms of the Second International Collaborative Ovarian Neoplasm Study (P = .98; thick line, CAP; thin line, carboplatin).
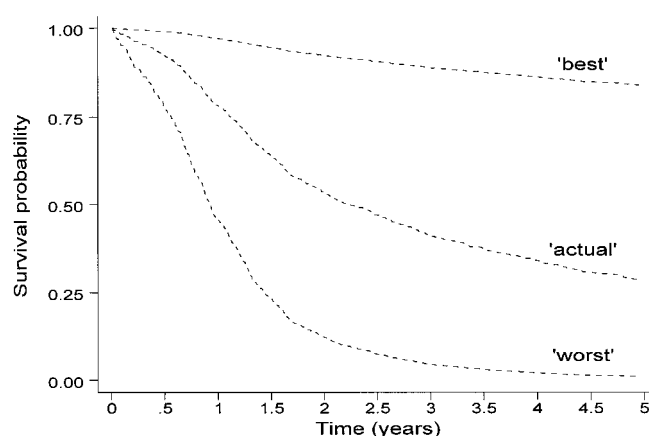

Fig 3. Expected survival for the patients actually entered onto the Second International Collaborative Ovarian Neoplasm Study (actual), and assuming all patients had the most favorable (best) or least favorable (worst) combinations of residual disease, age, International Federation of Gynecology and Obstetrics stage, and histologic grade.

somewhat between the meta-analysis and the ICON2 trial, but treatment arms seemed to be well balanced with respect to all prognostic factors in both patient series. When all patients included in the meta-analysis were used, the prognostic model for survival included the following factors (P < .01): extent of residual disease (HR = 1.54; SE = 0.08), age as a continuous variable (HR = 1.018; SE = 0.003), histologic grade (HR = 1.17; SE = 0.06), FIGO stage (HR = 1.36; SE = 0.12), and treatment (HR = 0.83; SE = 0.06). Performance status was also significant (P < .01) but was not retained in the model because it had not been collected in the ICON2 trial. No interaction term between any of the factors was statistically significant. Of the

Table 2. Main Patient Characteristics in the Meta-Analysis[9] and the ICON2 Trial[10]

	Meta-Analysis (%)		ICON2 (%)	
	CAP (n = 590)	CP (n = 608)	CAP (n = 766)	Carboplatin (n = 760)
Age, years				
< 55	44	45	34	32
55-65	38	39	35	36
> 65	18	16	31	32
Extent of residual disease				
No residual disease	16	15	31	31
≤ 2 cm	39	39	25	24
> 2 cm	43	45	44	45
Missing	2	1	—	—
Histologic grade				
Poor	40	37	46	50
Intermediate	32	35	32	30
Good	18	16	13	11
Missing	10	12	9	9
FIGO stage				
I	—	—	11	12
II	2	3	12	11
III	83	84	63	62
IV	14	12	14	15
Missing	1	1	—	—

Abbreviations: ICON2, Second International Collaborative Ovarian Neoplasm Study; CAP, cyclophosphamide, doxorubicin, cisplatin; CP, cyclophosphamide, cisplatin; FIGO, International Federation of Gynecology and Obstetrics.

1,198 patients included in the meta-analysis, 1,054 had complete data for the factors included in the model.

Expected Survival Curves

With this prognostic model for survival, expected survival curves could be constructed for any combination of the prognostic factors retained as significant in the model. Figure 3 shows the expected survival curves for 1,387 patients entered on the ICON2 trial, under three different assumptions: (1) who had complete data for the factors used in the model, using actually observed values of the prognostic factors (actual); (2) if all the patients had the best possible values for all prognostic factors in the model: no residual disease, age of 18 years, well-differentiated histology, and FIGO stage I (best); and (3) if all the patients had the worst possible values for all prognostic factors in the model: bulky residual disease, age of 85 years, poorly differentiated histology, and FIGO stage IV (worst). In all three cases, the original treatment assignment of the patients was used. Figure 3 illustrates that the expected survival curve depends heavily on the prognostic factors. The actual curve reflects the distribution of the prognostic factors in the ICON2 study presented in Table 2. Both the worst and best curves are markedly different from the actual curve; the 5-year survival probabilities are 30% for the patients actually entered onto the ICON2 trial (Table 2) versus 85% for a patient with the best characteristics and close to 0% for a patient with the worst characteristics.

Comparison of Observed and Expected Survival Curves

Expected survival curves were also constructed separately for each treatment arm of the ICON2 trial for the combination of the prognostic factors actually observed in that arm of the trial. Figure 4 shows the observed and expected survival curve in the CAP arm of the ICON2 trial. There was relatively little difference between the expected and observed curves. Figure 5 shows the observed survival curve in the carboplatin arm of the ICON2 trial and the expected survival curve using the CP arm of the
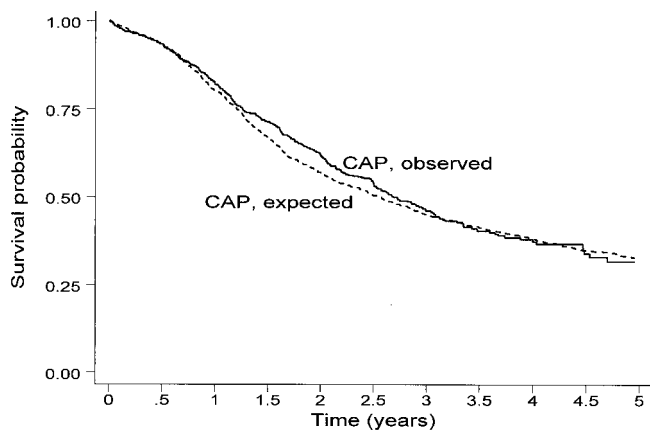
Fig 4. Observed survival in the cyclophosphamide, doxorubicin, and cisplatin (CAP) arm of the Second International Collaborative Ovarian Neoplasm Study, and expected survival on the basis of the data of the CAP arm from the meta-analysis. (——) Observed; (- - - -) expected.

Fig 6. Test statistic for the comparison of the observed and expected survival curves (thick line, observed cyclophosphamide, doxorubicin, and cisplatin [CAP] v expected CAP; thin line, observed carboplatin v expected cyclophosphamide and cisplatin [CP]). The test statistic for observed carboplatin versus expected CP was statistically significant at the 0.01 level after 1.25 years.

meta-analysis. There was a large difference between the expected and the observed curves in favor of the carboplatin arm of the ICON2 trial. Figure 6 shows the test statistic comparing the observed and the expected survival curves over time. There was a tendency for the observed and expected survival curves for CAP to differ between 1 and 3 years, but the separation between them was not statistically significant. In contrast, the observed survival curve for carboplatin and the expected survival curve for CP differed significantly at all times after approximately 1.25 years ($P < .01$).

DISCUSSION

These analyses provide highly suggestive evidence that the difference between the survival comparisons in the meta-analysis and in the ICON2 trial is because of the better results achieved with carboplatin alone at an optimally tolerated dose compared with the CP combination at a cisplatin dose of 50 to 60 mg/m^2. Indeed, after all known covariates were taken into account, the survival curve of the CAP arm in the ICON2 trial
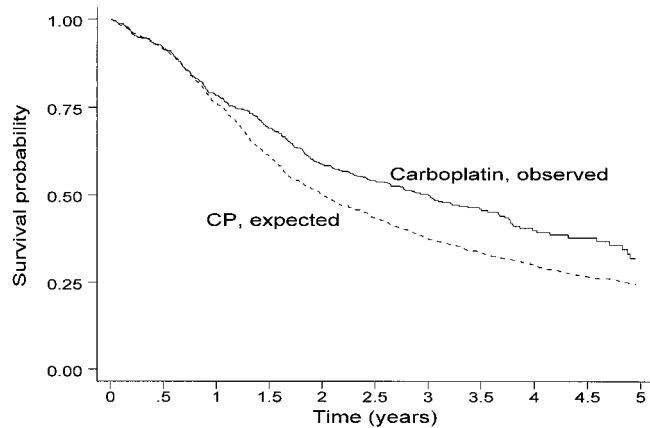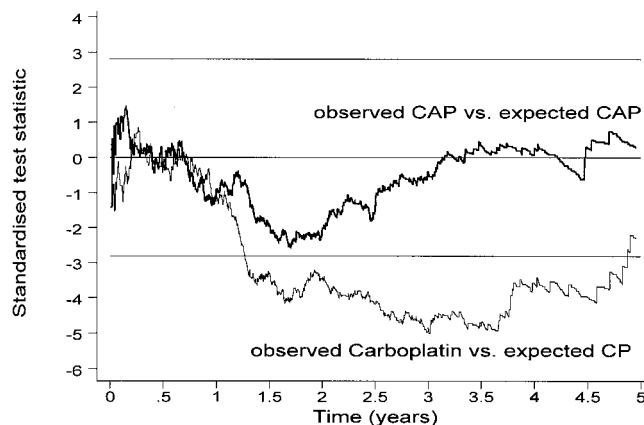

Fig 5. Observed survival in the carboplatin arm of the Second International Collaborative Ovarian Neoplasm Study, and expected survival on the basis of the cyclophosphamide and cisplatin (CP) arm of the meta-analysis. (——) Observed; (- - - -) expected.

was close to the curve that would have been expected had the mortality rates seen in the CAP arm of the meta-analysis been operating (Fig 4). The two curves were only slightly separated by the play of chance between 1 and 3 years. In contrast, the survival curve of the carboplatin arm in the ICON2 trial was much better than the curve that would have been expected had the mortality rates seen in the CP arm of the meta-analysis been operating (Fig 5). The separation between the two curves was highly significant at all times after 1.25 years ($P < .01$; Fig 6). Although the comparison between carboplatin and CP is indirect and, therefore, far less reliable than a direct comparison between randomized arms, it is calibrated by the presence of a common CAP arm that exhibited similar survival in both the meta-analysis and the ICON2 trial.

Taken together, the results of the meta-analysis and of the ICON2 trial indicate that the dose of platinum may be important. Direct evidence from randomized trials comparing doses of platinum, however, is inconclusive because of major differences between the trials, insufficient numbers of observations, and the play of chance.[18] Our results also indicate that the addition of an anthracycline to CP might compensate for the insufficient dose of platinum in the CP arm. Other trials that are ongoing will provide further evidence on any added benefits of anthracyclines.[19]

The comparison of survival curves across different experiments requires knowledge of prognostic factors for individual patients in both experiments.[20] Indeed, as shown in Fig 3, the expected survival curve depends heavily on the prognostic factors found to be significant in the survival model for advanced ovarian cancer (extent of residual disease, age, FIGO stage, and histologic grade). Had performance status been added to the model, the separation of the survival curves presented in Fig 3 would have been even more pronounced. In other diseases for which prognostic factors are less well known or less predictive of the outcome of interest, it may be more difficult to find agreement between independent series of identically treated patients. In any case, our results show remarkable agreement

Table 3. Randomized Evidence on the Treatment of Advanced Ovarian Cancer and Possible Uses of the Expected Survival

Source	Survival Results
Meta-analysis[9]	CAP > CP
ICON2[10]	CAP ≈ Carbo
Expected survival using data from the meta-analysis and ICON2	Carbo > CP
GOG-111[21]	PT > CP
OV10[22]	PT > CP
ICON3[23]	Carbo + paclitaxel ≈ Carbo or CAP
Expected survival using data from GOG-111, OV10, and ICON3?	Carbo or CAP > CP?

Abbreviations: CAP, cyclophosphamide, doxorubicin, cisplatin; CP, cyclophosphamide, cisplatin; Carbo, carboplatin; PT, cisplatin, paclitaxel; >, statistically superior; ≈, not statistically different; ICON, International Collaborative Ovarian Neoplasm; GOG, Gynecologic Oncology Group.

between the survival experience of identically treated patients in a meta-analysis of several small trials and in a large-scale confirmatory trial. The results also demonstrate the importance of performing meta-analysis on the basis of individual patient data, which implies a willingness to share data on the part of the principal investigators of all relevant trials.

The data available in this study included survival as well as important prognostic information in both the meta-analysis and the ICON2 trial. Without such data, there would have been no reliable way of comparing the survival results of the meta-analysis with the results of the ICON2 trial because these series had different distributions of prognostic factors known to have a major impact on survival (Table 2). Even after taking all known

prognostic factors into account, however, we could still have observed a difference between identically treated patients (with CAP) in the meta-analysis and in the ICON2 trial, if only because the ICON2 trial was performed 10 years after the trials included in the meta-analysis. Had that been the case, the reason the other treatment arms differed would have been left unanswered or would be speculative, at best.

Although the results presented here may be particularly clear-cut because the two CAP groups exhibited almost identical results, the expected survival approach may provide a generally useful approach to reconcile the results of independent randomized trials of similar therapies. This approach may provide valuable insight when direct comparisons between randomized groups yield ambiguous or discrepant results. A case in point concerns the role of taxanes in the treatment of advanced ovarian cancer. Two trials, one conducted by the Gynecologic Oncology Group in the United States[21] and the other conducted by a European Canadian Intergroup,[22] showed that the combination of cisplatin and paclitaxel was superior to CP. The recently published ICON3 trial[23] found no benefit of the combination of carboplatin and paclitaxel over either carboplatin alone or CAP (Table 3). The results discussed above indicate that CP may be inferior to both carboplatin alone and CAP. Consequently, the lack of benefit in ICON3 may be because the control group in this trial was superior to the control group used in the two previous trials.[24] This hypothesis could be tested formally using the expected survival approach, as demonstrated in the situation of the meta-analysis and the ICON2 trial analyzed above (Fig 6 and Table 3).

ACKNOWLEDGMENT

REFERENCES

1. Collins R, Gray R, Godwin J, et al: Avoidance of large biases and large random errors in the assessment of moderate treatment effects: The need for systematic overviews. Stat Med 6:245-250, 1987

2. Villar J, Carroli G, Belizan JM: Predictive ability of meta-analyses of randomised controlled trials. Lancet 345:772-776, 1995

3. LeLorier J, Grégoire G, Benhaddad A, et al: Discrepancies between meta-analyses and subsequent large randomized, controlled trials. N Engl J Med 337:536-542, 1997

4. Cappelleri JC, Ioannidis JPA, Schmid CH, et al: Large trials vs meta-analysis of smaller trials: How do their results compare? J Am Med Assoc 276:1332-1338, 1996

5. Ioannidis JPA, Cappelleri JC, Lau J: Meta-analyses and large randomized, controlled trials. N Engl J Med 338:59, 1998 (letter)

6. Ovarian Cancer Meta-Analysis Project: Cyclophosphamide plus cisplatin versus cyclophosphamide, doxorubicin and cisplatin chemotherapy of ovarian carcinoma: A meta-analysis. J Clin Oncol 9:1668-1674, 1991

7. Ghersi D, Parmar MKB, Stewart LA, et al: Early ovarian cancer and the ICON trials. Eur J Cancer 28A:1297, 1992

8. Advanced Ovarian Cancer Trialists Group: Chemotherapy in advanced ovarian cancer: An overview of randomised clinical trials. Br Med J 303:884-893, 1991

9. Ovarian Cancer Meta-Analysis Project: Cyclophosphamide plus cisplatin versus cyclophosphamide, doxorubicin and cisplatin chemotherapy of ovarian carcinoma: A meta-analysis. Classic Papers Current Comments 3:237-243, 1998

10. The ICON Collaborators: ICON-2: A randomised trial of single-agent carboplatin against the 3-drug combination of CAP (cyclophosphamide, doxorubicin and cisplatin) in women with ovarian cancer. Lancet 352:1571-1576, 1998

11. Calvert AH, Newell DR, Gumbrell LA, et al: Carboplatin dosage: Prospective evaluation of simple formula based on renal function. J Clin Oncol 7:1748-1756, 1989

12. Cox DR: Regression models and life tables. J R Stat Soc B 34:187-220, 1972

13. Thomsen BL, Keiding N, Altman DG: A note on the calculation of expected survival. Stat Med 10:733-738, 1991

14. Therneau TM, Grambsch PM: Expected survival, in Modelling Survival Data: Extending the Cox Model. New York, NY, Springer-Verlag, 2000

15. Andersen PK, Vaeth M: Simple parametric and nonparametric models for excess and relative mortality. Biometrics 45:523-535, 1989

16. Nielsen B: Expected survival in the Cox model. Scand J Stat 24:275-287, 1997

17. Schumacher M: Two sample tests of Cramér-von Mises and Kolmogorov-Smirnov type for randomly censored data. Int Stat Rev 52:263-281, 1984

18. Thigpen JT: Dose-intensity in ovarian carcinoma: Hold, enough? J Clin Oncol 15:1291-1293, 1997

19. Vermorken JB, Harper PG, Buyse M: The role of anthracyclines in epithelial ovarian cancer. Ann Oncol 10:S43-S50, 1999 (suppl 1)

20. Simon R: Importance of prognostic factors in cancer clinical trials. Cancer Treat Rep 68:185-192, 1984

21. McGuire WP, Hoskins WJ, Brady MF, et al: Cyclophosphamide and cisplatin compared with paclitaxel and cisplatin in patients with stage 3 and 4 ovarian cancer. N Engl J Med 334:1-6, 1996

22. Piccart MJ, Bertelsen K, James K, et al: Randomized Intergroup trial of cisplatin-paclitaxel versus cisplatin-cyclophosphamide in women with advanced epithelial ovarian cancer. J Natl Cancer Inst 92:699-708, 2000

23. The International Collaborative Ovarian Neoplasm (ICON) Group: Paclitaxel plus carboplatin versus standard chemotherapy with either carboplatin or cyclophosphamide, doxorubicin and cisplatin in women with ovarian cancer: The ICON3 randomized trial. Lancet 360:505-515, 2002

24. Sandercock J, Parmar MKB, Torri V, et al: First-line treatment for advanced ovarian cancer: Paclitaxel, platinum and the evidence. Br J Cancer 87:815-824, 2002