PERGAMON

European Journal of Cancer 36 (2000) 1488–1497

European Journal of Cancer

www.ejconline.com

Randomised trial of paclitaxel versus doxorubicin as first-line chemotherapy for advanced breast cancer: quality of life evaluation using the EORTC QLQ-C30 and the Rotterdam Symptom Checklist

J.A. Kramer [a], D. Curran [a,*], M. Piccart [b], J.C.J.M. de Haes [c], P.F. Bruning [d], J.G.M. Klijn [e], M. Bontenbal [e], C. van Pottelsberghe [a], M. Groenvold [f], R. Paridaens [g]

[a] *European Organization for Research and Treatment of Cancer (EORTC) Data Center, Avenue Mounier 83, Bte 11, B-1200, Brussels, Belgium*
[b] *Institut Jules Bordet, Brussels, Belgium*
[c] *Academic Medical Center, Amsterdam, The Netherlands*
[d] *The Netherlands Cancer Institute, Amsterdam, The Netherlands*
[e] *Daniel den Hoed Cancer Center, Rotterdam, The Netherlands*
[f] *Department of Palliative Medicine, Bispebjerg Hospital, Copenhagen, Denmark*
[g] *Department of Oncology, University Hospital Gasthuisberg, Leuven, Belgium*

Received 13 December 1999; received in revised form 28 February 2000; accepted 4 May 2000

Abstract

The aim of the study was to compare the quality of life (QL) of patients treated with single-agent paclitaxel versus doxorubicin as first-line chemotherapy for advanced breast cancer. 331 patients with advanced breast cancer were randomised, with 294 eligible for analysis. Patients completed both the EORTC QLQ-C30 questionnaire and the Rotterdam Symptom Checklist (RSCL) with six additional items, at baseline and after the third, fifth and seventh cycles of chemotherapy. A significant difference in progression-free survival in favour of doxorubicin caused a bias in the data with differences in expected completion rates of questionnaires beyond cycle three. Therefore, statistical comparisons were performed only for the first three cycles. Baseline compliance was 64% and 61% for the QLQ-C30 and RSCL questionnaires, respectively. Doxorubicin was associated with significantly more nausea/vomiting ($P=0.001$), loss of appetite ($P=0.010$) and a greater burden of disease and treatment ($P=0.044$), but with less bone pain ($P=0.042$) and rash ($P=0.045$) than paclitaxel. Both treatments were associated with improved emotional function and reduction in psychological distress at cycle 3. Longitudinal data suggested that doxorubicin was associated with less pain, specifically bone pain. Doxorubicin was more active but may have had more side-effects during the first three cycles. Long-term QL outcomes could not be assessed.

Keywords: Breast cancer; Quality of life; EORTC QLQ-C30; Rotterdam Symptom Checklist

1. Introduction

Despite clinical responses to standard endocrine and cytotoxic regimens, the median duration of survival from identification of metastases in women with breast cancer is 2–3 years [1]. Chemotherapy is widely used in the management of such patients, with the aims of prolonging progression-free survival, relieving symptoms and delaying or preventing anticipated problems, and thereby optimising quality of life with the disease. Quality of life (QL) has been included as an outcome measure in a few randomised trials comparing different chemotherapy regimens in advanced breast cancer. These studies have yielded data about patient-reported effects of treatment which complement reports of response rates and toxicity profiles. Coates and colleagues [2] used visual analogue scales to assess QL and observed that over approximately 9 weeks between baseline and the end of the third cycle of chemotherapy (doxorubicin + cyclophosphamide versus cyclophosphamide, methotrexate, 5-fluorouracil and prednisone), physical well-being, pain, mood and appetite all improved, as did the global QL measure. Only nausea/vomiting worsened. Richards and associates [3] used the Rotterdam Symptom Checklist (RSCL) to assess QL during treatment with two different schedules of doxorubicin,

* Corresponding author. Tel.: +353-1-2161-284 fax: +353-1-2161-200.
E-mail address: currand@iconirl.com (D. Curran).

0959-8049/00/$ - see front matter
PII: S0959-8049(00)00134-9

and observed that over 12 weeks after randomisation psychological scores improved in those receiving 3-weekly treatment but not in those having weekly treatment. Pain and breathlessness improved with both regimens, whereas vomiting and diarrhoea worsened during treatment but returned to pretreatment levels after completion of chemotherapy. Ramirez and associates [4] studied a consecutive series of women with advanced breast cancer who received any of a variety of chemotherapy regimens as first-line treatment for advanced disease. The main outcome measure was overall well being assessed at a post-treatment interview (4–6 months after starting treatment), while functional status and symptoms were measured using the RSCL before, during and after treatment. Approximately equal proportions felt better, worse and the same after chemotherapy. Those who felt better experienced a reduction in psychological distress and pain, less fatigue and a trend towards improvement in functional status. The authors of each of these studies concluded that chemotherapy seems to be beneficial in advanced breast cancer even if it does not improve overall survival, because it gains some responses and tends not to impair QL. However, none of these papers adequately considered the limitations of the datasets examined, such as short follow-up, problems with attrition and compliance and the resultant bias.

The aim of the present study was to compare single-agent paclitaxel versus doxorubicin as first-line chemotherapy for advanced breast cancer. In the event that no substantial differences were observed in response rates or progression-free survival between the treatments, QL was considered an important outcome measure of the study. The intention was to measure QL longitudinally to determine treatment-related changes over time both on first-line treatment and after crossover to the alternative regimen as second-line therapy following progression. When the study was designed, the RSCL [5] was already well established as a measurement tool for QL, whereas the EORTC QLQ-C30 had only recently been validated [6]. Therefore, both instruments were used to assess QL to try to ensure that any true differences between the treatments were detected. In this paper, we demonstrate differences in QL during the first months of treatment, but also the difficulties we encountered in collecting, analysing and interpreting longitudinal QL data in a randomised comparison of two chemotherapy agents in advanced breast cancer.

2. Patents and methods

2.1. Clinical data

EORTC trial 10923 was a prospective, randomised phase II/III crossover study in advanced breast cancer, in which the efficacy of paclitaxel 200 mg/m^2 as a 3-h infusion every 3 weeks until progression followed by doxorubicin 75 mg/m^2 (maximum seven cycles) as an intravenous (i.v.) bolus every 3 weeks was compared with the reverse regimen of doxorubicin followed by paclitaxel. The trial was approved by the EORTC Protocol Review Committee and by the ethics committee of each participating centre, and was conducted in compliance with the Helsinki declaration. All patients gave informed consent.

Clinical details of this trial have been reported elsewhere [7]. Briefly, 331 patients from 20 institutions were entered into the trial, 166 randomised to paclitaxel and 165 to doxorubicin. Response was documented clinically before each new cycle of chemotherapy (day 1), whilst mandatory clinical and radiological evaluation of response was made at the end of the third, fifth and seventh cycles.

2.2. QL evaluation

The EORTC QLQ-C30 v.1 consists of five function scales: physical (PF), role (RF), emotional (EF), cognitive (CF) and social (SF); three symptom scales: fatigue (FA), nausea/vomiting (NV) and pain (PA); six single-item scales (dyspnoea (DY), sleep disturbance (SL), appetite loss (AP), constipation (CO), diarrhoea (DI) and financial impact of the disease/treatment (FI), and a global health status/quality of life (QL) scale [6]. Items were scored and scales constructed using recommended procedures [8]. Raw scores were transformed to a linear scale ranging from 0 to 100, with higher scores representing a higher level of functioning or a higher level of symptoms. In the case of missing items within a scale, the scale score was calculated using only those for which values were available, provided at least half of the items in the scale were completed. The financial impact variable was not included in this analysis, because it is difficult to interpret in a multinational clinical trial.

The RSCL consists of four scales: physical symptom distress, psychological distress, activity level and overall valuation of life (a global QL item) [9]. As for the QLQ-C30, most responses are given on a 4-point Likert-type scale and are transformed to a linear scale ranging from 0 to 100. Higher scores correspond to better function and less distress. For this study, six items were appended to the RSCL: bone pain, hot flushes, rash, disruption in relationships because of the illness and treatment, the overall burden of disease and treatment, and a visual analogue scale (VAS) for global QL.

The protocol specified that QL assessment should be undertaken at baseline (not earlier than 2 weeks before the onset of therapy), and not earlier than 2 days before starting the next cycle at the end of cycles 3, 5 and 7 of chemotherapy. Administration of questionnaires during treatment was not allowed. Since patients were asked to

complete both the EORTC QLQ-C30 and the RSCL questionnaires at each assessment point, the order in which these were administered was randomised to exclude an effect of order of presentation on QL scores and to minimise differences in completion rates between questionnaires.

2.3. Compliance

Compliance was calculated as the number of forms received out of the number expected at each assessment point. To define the number of forms expected we calculated the duration of first-line treatment from the date of randomisation as follows: to the date of completion of seven cycles of therapy if a response or stable disease were observed, and to the date of progression or date of death if this occurred while the patient was still on first-line treatment.

It is important to allow for delays that occurred in the administration of chemotherapy cycles and consequently of QL questionnaires to ensure that patients with delayed treatment were not artificially removed from the QL analysis. Therefore, time windows were chosen a priori for accepting QL forms based on the following principles: forms were excluded if they were completed within 1 week after starting a cycle and included if they were completed within a time window that allowed for a 1-week delay in each cycle due to toxicity. Overlap of time windows was avoided. The time windows were defined from randomisation as follows: baseline (2 weeks before to 2 weeks after), cycle 3 (8– < 13 weeks after), cycle 5 (14– < 20 weeks after) and cycle 7 (20– < 29 weeks after). If two forms were completed by a patient within a time window, those completed at unacceptable times (such as within 1 week after starting treatment) were excluded. If two acceptable forms were available within a single time window, the earlier one was used in the analysis because it was likely to be closer to the expected date for which a form was required.

QL data were examined in several ways. QL measured at baseline using the QLQ-C30 was compared with baseline reference data for patients with advanced breast cancer enrolled in clinical trials [10]. Data from both the QLQ-C30 and the RSCL with additional items were used in a cross-sectional comparison of the two treatments at the end of cycle 3, and also to compare changes in QL scores between baseline and cycle 3. A descriptive profile of the longitudinal QL data was obtained for each treatment arm by plotting mean QL scores for each of the subscales by dropout time.

2.4. Statistical considerations

Randomisation was performed using the minimisation technique [11]. The primary endpoint of the trial was progression-free survival. QL, response to treatment and overall survival were secondary endpoints. Survival curves and probabilities for the duration of first-line treatment were estimated using the Kaplan–Meier technique [12].

Treatment comparisons at cycle 3, and change scores between baseline and cycle 3 were performed using the Wilcoxon rank-sum test [13]. Data analysis was performed using Statistical Analysis Software (SAS) [14]. As this was an exploratory analysis a 5% level of significance was used and all significance tests were two-sided.

3. Results

3.1. Clinical data

The clinical results for all randomised patients ($n = 331$) have been reported elsewhere [7]. Briefly, for first-line therapy there was a significant difference in progression-free survival between the two treatments (median 7.5 months for doxorubicin versus 3.9 months for paclitaxel, $P < 0.001$). For those who progressed and subsequently received second-line therapy, the median number of first-line cycles received was 5 and 3 for doxorubicin and paclitaxel, respectively. Table 1 summarises the reasons for discontinuing first-line treatment.

3.2. QL evaluation

Of the 331 patients randomised from 20 institutions between September 1993 and April 1996, 15 were ineligible on clinical grounds (8 with no measurable lesions, 4 with inadequate organ function, 2 with symptomatic cerebral metastases, and 1 with prior chemotherapy for metastatic disease). A further 22 patients were not eligible for the QL analysis (3 did not complete QL questionnaires because translations were not available in their language and 19 used non-validated translations), leaving 294 eligible patients for QL analysis.

Table 1
Reasons for discontinuing first-line treatment (all randomised patients)

Reasons for discontinuing treatment	Number (%) patients	
	Paclitaxel	Doxorubicin
Completion of $\geqslant 7$ cycles with CR, PR, SD	57 (34)	75 (45)
Progression with cross-over	65 (39)	24 (15)
Progression without cross-over	21 (13)	8 (5)
Excessive toxicity	9 (5)	26 (16)
Maximum dose administered	–	6 (4)
Patient's request	3 (2)	8 (5)
Death (malignant and other)	8 (5)	10 (6)
Other	3 (2)	8 (5)
Total	166 (100)	165 (100)

CR, complete response; PR, partial response; SD, stable disease.

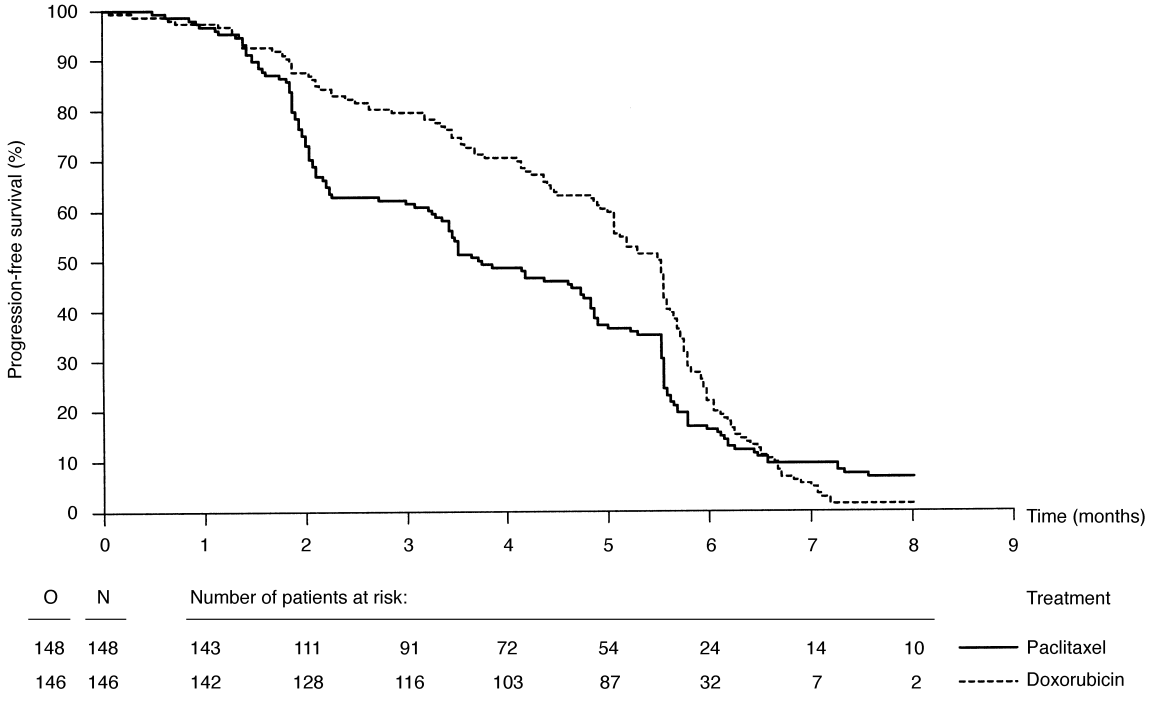
3.3. Compliance

Fig. 1 shows the duration of first-line treatment for the 294 eligible patients. There was greater attrition in the paclitaxel arm at the end of cycle 3 (week 9) when the first mandatory evaluation of response took place according to the protocol. At the end of cycle 7, 91 (62%) patients randomised to doxorubicin remained on first-line treatment, compared with 61 (41%) receiving paclitaxel.

Of all eligible patients, 107 did not have baseline QLQ-C30 questionnaires: 94 for unknown reasons (including all 30 patients recruited by two institutions), 12 whose forms were unacceptable because they were completed after starting the first course of chemotherapy and 1 whose form was completed more than 14 days before randomisation. Of 294 eligible patients, 187 had acceptable baseline QLQ-C30 forms with a compliance rate of 64%. For the RSCL, 180 of 294 eligible patients had acceptable baseline forms, giving a compliance rate of 61%.

Using the time windows selected a priori to take account of treatment delays, 80 and 84 patients had no valid QLQ-C30 and RSCL questionnaires, respectively. Thirty QLQ-C30 and 28 RSCL forms were excluded because of completion within 1 week after starting a cycle of chemotherapy (13 QLQ-C30 forms (11 RSCL) were excluded from the paclitaxel arm and 17 from the doxorubicin arm). Of those excluded, 13 were completed 1 day after starting treatment, and 12/30 (40%) occurred at the baseline assessment. There were no significant differences in the number and pattern of excluded forms between treatment arms.

Delays in administration of each cycle of first-line chemotherapy were investigated by treatment arm (Table 2). The delays did not differ between treatment arms up to and including cycle 3, but by cycle 5 those in the doxorubicin arm had a median delay of 1 week in their treatment cycle, which increased to almost 2 weeks by cycle 7. Those receiving paclitaxel had a median delay of less than 1 week in the administration of cycles 4–7.



Compliance with completion of QL questionnaires during first-line treatment

	Paclitaxel				Doxorubicin			
Cycle schedule	0	3	5	7	0	3	5	7
Expected (*n* on first-line treatment)	148	102	81	61	146	126	110	91
QLQ-C30 received (*n*)	93	61	45	37	94	73	49	35
(% compliance)	(63)	(60)	(56)	(61)	(64)	(58)	(45)	(38)
RSCL received (*n*)	91	61	45	35	89	69	50	35
(% compliance)	(61)	(60)	(56)	(57)	(61)	(55)	(45)	(38)

Fig. 1. Duration of first-line treatment. The accompanying table shows compliance with completion of questionnaires for each treatment arm.

Table 2
Delays in treatment by treatment arm

Cycle of treatment	Expected day of treatment	Median (IQR) day of treatment	
		Paclitaxel	Doxorubicin
2	21	23 (22–26)	23 (22–27)
3	42	45 (43–48)	45 (43–49)
4	63	66 (64–70)	68 (64–74)
5	84	88 (85–91)	91 (86–100)
6	105	109.5 (106–113)	114 (108–126)
7	126	131 (127–135)	139.5 (131–154)

IQR, interquartile range.

Fig. 1 shows compliance rates for completion of QL forms on first-line chemotherapy. Baseline and cycle 3 compliance rates were similar for both treatments. By cycle 7 the number of forms received in both arms combined represented only 24% (72 for QLQ-C30, 70 for RSCL) of those patients initially available at baseline and a minority (47%) of those remaining on first-line treatment. Compliance with completion of QL questionnaires after crossover to second-line treatment was negligible ($<30\%$).

The pattern of missing data and drop-out for completion of QLQ-C30 forms is shown in Table 3. For the RSCL and additional items the patterns were similar. Only 38 patients had complete longitudinal QLQ-C30 data covering the 4 assessment points from baseline to cycle 7 (39 for the RSCL).

3.4. Bias

By the first mandatory evaluation of response at the end of cycle 3 there was greater attrition from first-line treatment in the paclitaxel arm (Fig. 1). Beyond this point, there was a difference in expected rates of completion of QL questionnaires (Fig. 1) between the two arms. By cycle 7 there were 21% fewer patients remaining on first-line treatment in the paclitaxel arm, but their compliance with completion of questionnaires was 23% QLQ-C30 and 19% RSCL higher than in the doxorubicin arm. Because of the low numbers of questionnaires and the risk of bias in the data beyond cycle 3, we did not carry out any statistical comparisons beyond this point.

Table 3
Available QLQ-C30 forms and patterns of missing data

Baseline	Cycle 3	Cycle 5	Cycle 7	Frequency	Per cent
*	*	*	*	38	17.8
*	*	*		30	14.0
*	*			34	15.9
*				61	28.5
*	*		*	11	5.1
*		*	*	8	3.7
*		*		3	1.4
*			*	2	0.9
	*	*	*	6	2.8
	*	*		6	2.8
	*			8	3.7
	*		*	1	0.5
		*	*	3	1.4
			*	3	1.4
187	134	94	72	214	99.9

3.5. Baseline scores for the QLQ-C30

Table 4 shows median and mean baseline scores for the items included in the analysis, together with reference values for advanced breast cancer [10]. For patients in this study, the baseline scores for the functional scales are very similar to the reference values whereas all symptom scores are slightly lower, suggesting that the patients studied were slightly less symptomatic. They also had a higher mean global QL score than the reference sample.

3.6. Comparison of the two treatments at the end of cycle 3

134 patients had QLQ-C30 data available at the end of cycle 3 (61 receiving paclitaxel and 73 receiving doxorubicin). Table 5 shows the proportions of patients in each group of QLQ-C30 scores for each treatment arm at the end of cycle 3. For the single-item symptom scales, columns correspond to the patients' responses 'not at all', 'a little', 'quite a bit' and 'very much'. Significant differences between the treatment arms at the end of cycle 3 were observed for nausea/vomiting ($P=0.001$) and loss of appetite ($P=0.010$) in favour of paclitaxel. At the end of the cycle 3, 49 (82%) of patients receiving paclitaxel had no nausea/vomiting

Table 4
Baseline QLQ-C30 scores ($n=181$–187), with baseline reference values for comparison

	Median	Mean	Reference mean[a]
Physical function	60.0	65.5	64.1
Role function	50.0	63.6	63.4
Emotional function	66.7	63.2	64.8
Cognitive function	75.0	82.4	81.1
Social function	83.3	76.0	70.1
Global QL	66.7	60.1	54.1
Fatigue	33.3	37.2	40.1
Nausea/vomiting	0.0	8.6	12.4
Pain	16.7	33.7	39.0
Dyspnoea	0.0	22.4	28.9
Sleeping difficulty	33.3	32.2	34.6
Appetite loss	0.0	27.3	28.6
Constipation	0.0	15.8	16.9
Diarrhoea	0.0	4.9	8.1

QL, quality of life.
[a] Baseline pretreatment values for advanced breast cancer [10].

compared with 30 (41%) in the doxorubicin arm. No statistically significant differences between the treatments in functional dimensions or in global QL were observed.

Table 5
Cross-sectional comparison of paclitaxel and doxorubicin at cycle 3 using QLQ-C30 scores

	Number (%) patients				*P* value
QLQ-C30 Score[a]	0–25	26–50	51–75	76–100	
Physical function					0.531
Paclitaxel	6 (10)	14 (23)	14 (23)	26 (43)	
Doxorubicin	10 (14)	12 (16)	25 (34)	26 (36)	
Role function					0.053
Paclitaxel	3 (5)	28 (47)	–	29 (48)	
Doxorubicin	13 (18)	32 (44)	–	28 (38)	
Emotional function					0.850
Paclitaxel	5 (8)	7 (12)	25 (42)	23 (38)	
Doxorubicin	6 (8)	12 (17)	22 (31)	32 (44)	
Cognitive function					0.158
Paclitaxel	0 (0)	7 (12)	10 (16)	44 (72)	
Doxorubicin	2 (3)	13 (18)	9 (13)	48 (67)	
Social function					0.534
Paclitaxel	5 (8)	10 (16)	6 (10)	40 (66)	
Doxorubicin	8 (11)	13 (19)	12 (17)	37 (53)	
Global QL					0.188
Paclitaxel	2 (3)	20 (33)	26 (43)	13 (21)	
Doxorubicin	7 (10)	26 (37)	26 (37)	12 (17)	
Fatigue					0.103
Paclitaxel	21 (35)	21 (35)	9 (15)	9 (15)	
Doxorubicin	17 (23)	24 (33)	16 (22)	16 (22)	
Nausea/vomiting					0.001
Paclitaxel	54 (90)	5 (8)	1 (2)	0	
Doxorubicin	51 (70)	15 (21)	5 (7)	2 (3)	
Pain					0.349
Paclitaxel	35 (57)	14 (23)	7 (12)	5 (8)	
Doxorubicin	45 (62)	18 (25)	7 (10)	3 (4)	
Dyspnoea[b]					0.466
Paclitaxel	27 (46)	16 (27)	13 (22)	3 (5)	
Doxorubicin	32 (44)	30 (41)	8 (11)	3 (4)	
Difficulty sleeping[b]					0.088
Paclitaxel	25 (42)	17 (28)	11 (18)	7 (12)	
Doxorubicin	38 (52)	22 (30)	9 (12)	4 (5)	
Lack of appetite[b]					0.010
Paclitaxel	43 (72)	9 (15)	4 (7)	4 (7)	
Doxorubicin	32 (44)	21 (29)	14 (19)	6 (8)	
Constipation[b]					0.164
Paclitaxel	37 (62)	14 (23)	7 (12)	2 (3)	
Doxorubicin	40 (55)	14 (19)	13 (18)	6 (8)	
Diarrhoea[b]					0.378
Paclitaxel	53 (87)	6 (10)	2 (3)	0	
Doxorubicin	57 (80)	11 (15)	3 (4)	0	

[a] As data are missing for some scores paclitaxel arm $n = 59–61$; doxorubicin arm $n = 70–73$.

[b] Single items for which columns correspond to responses 'not at all', 'a little', 'quite a bit'' and 'very much'.

Fewer completed forms were available for the RSCL and appended items. Additionally, French-speaking patients received only the first page of the questionnaire because of an administrative error. Therefore, 25% of the data were missing for certain variables (activity level, overall valuation of life, burden of disease and treatment and disruption of relationships). Significant differences between the treatment arms at the end of cycle 3 were observed for bone pain (worse in the paclitaxel arm, $P = 0.042$), rash (worse with paclitaxel, $P = 0.045$) and burden of disease and treatment (greater in the doxorubicin arm, $P = 0.044$). Bone pain was present in 35/59 (58%) of patients receiving paclitaxel and in 27/66 (41%) on doxorubicin, and rash in 16/60 (27%) (paclitaxel) and 9/68 (13%) (doxorubicin). 'Quite a bit' or 'very much' burden of disease and treatment was reported by 20/46 (43%) (paclitaxel) and 32/51 (63%) receiving doxorubicin.

3.7. Change scores from baseline to cycle 3

Change in QL dimensions was investigated for the period from randomisation to the end of cycle 3. 13 patients had QLQ-C30 questionnaires for baseline and cycle 3 (Table 3). For the RSCL and additional items, 108 patients had questionnaires available for analysis of change scores. For both the QLQ-C30 and RSCL positive change scores indicate improvement and negative scores indicate deterioration from baseline.

Figs 2 and 3 show the mean change scores between baseline and the end of the cycle 3 measured with the QLQ-C30, RSCL and the six additional items. For both treatments, the QLQ-C30 recorded improvements in emotional function and pain but worsening of most of the other variables. There were no statistically significant differences between treatment arms in any variable, but a borderline difference in nausea/vomiting (worse with doxorubicin, $P = 0.055$). The RSCL demonstrated similar improvements in psychological distress (cf. with emotional function on the QLQ-C30). The question specifically about bone pain showed a borderline significant difference between treatment arms, with improvement in bone pain in those receiving doxorubicin and deterioration in those receiving paclitaxel. This finding contrasts with the general questions on pain on the QLQ-C30, which demonstrated mean improvement in both treatment arms.

Fig. 4 shows the mean scores for selected QL dimensions by drop-out time for each treatment arm. There are clear differences in the plots for nausea/vomiting, appetite loss, pain and bone pain between treatment arms. Worse physical function and increasing fatigue are generally apparent in those who dropout early from both treatments, whilst improvement in emotional function and lessening in psychological distress are general trends that occur during the first three cycles of treatment, especially in patients who complete at least five cycles.

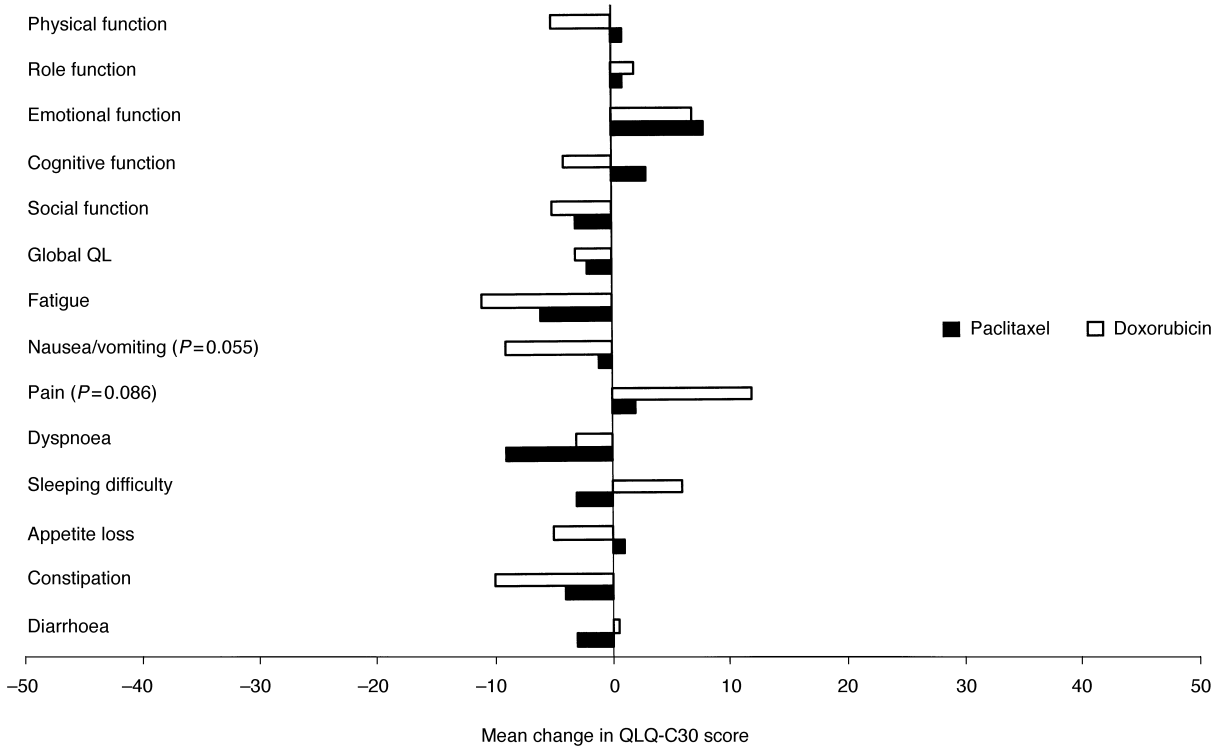


Fig. 2. Mean change in QLQ-C30 scores between baseline and the end of cycle 3.

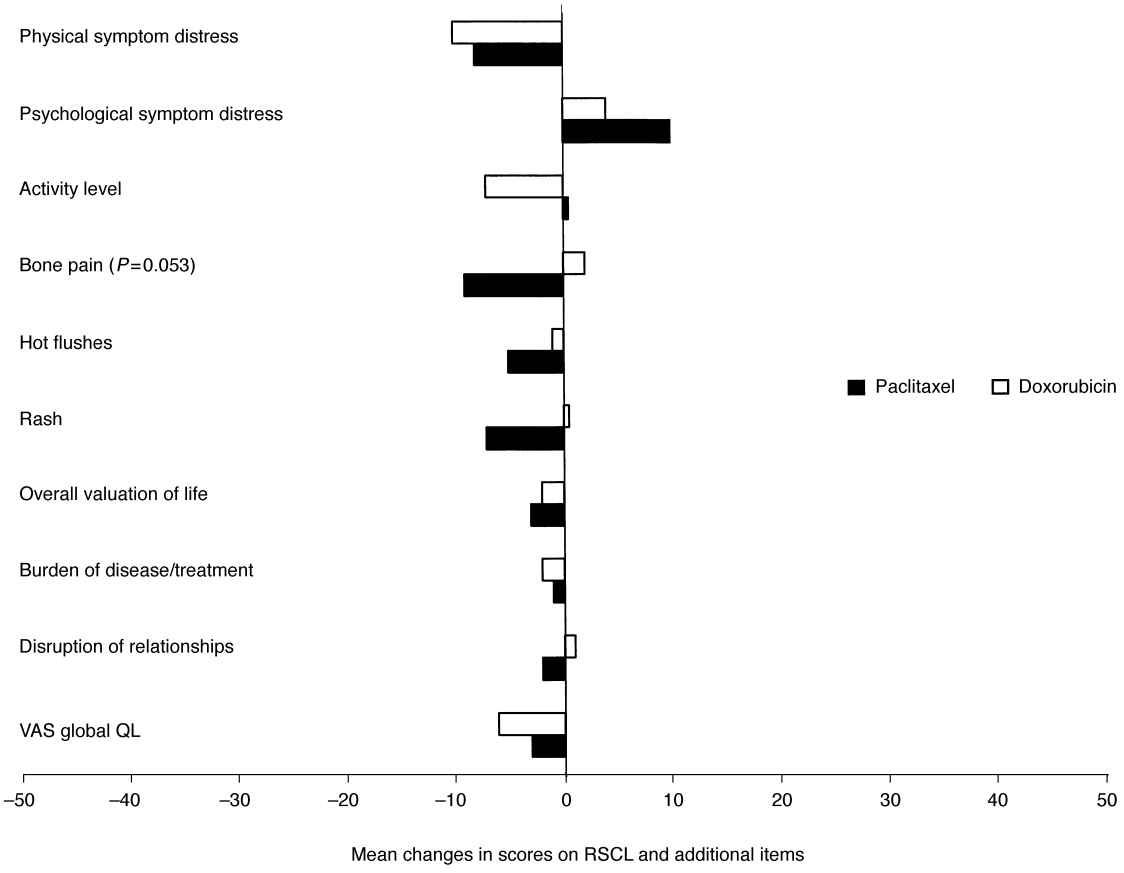


Fig. 3. Mean change in scores on the Rotterdam Symptom Checklist (RSCL) and additional items between baseline and the end of cycle 3.

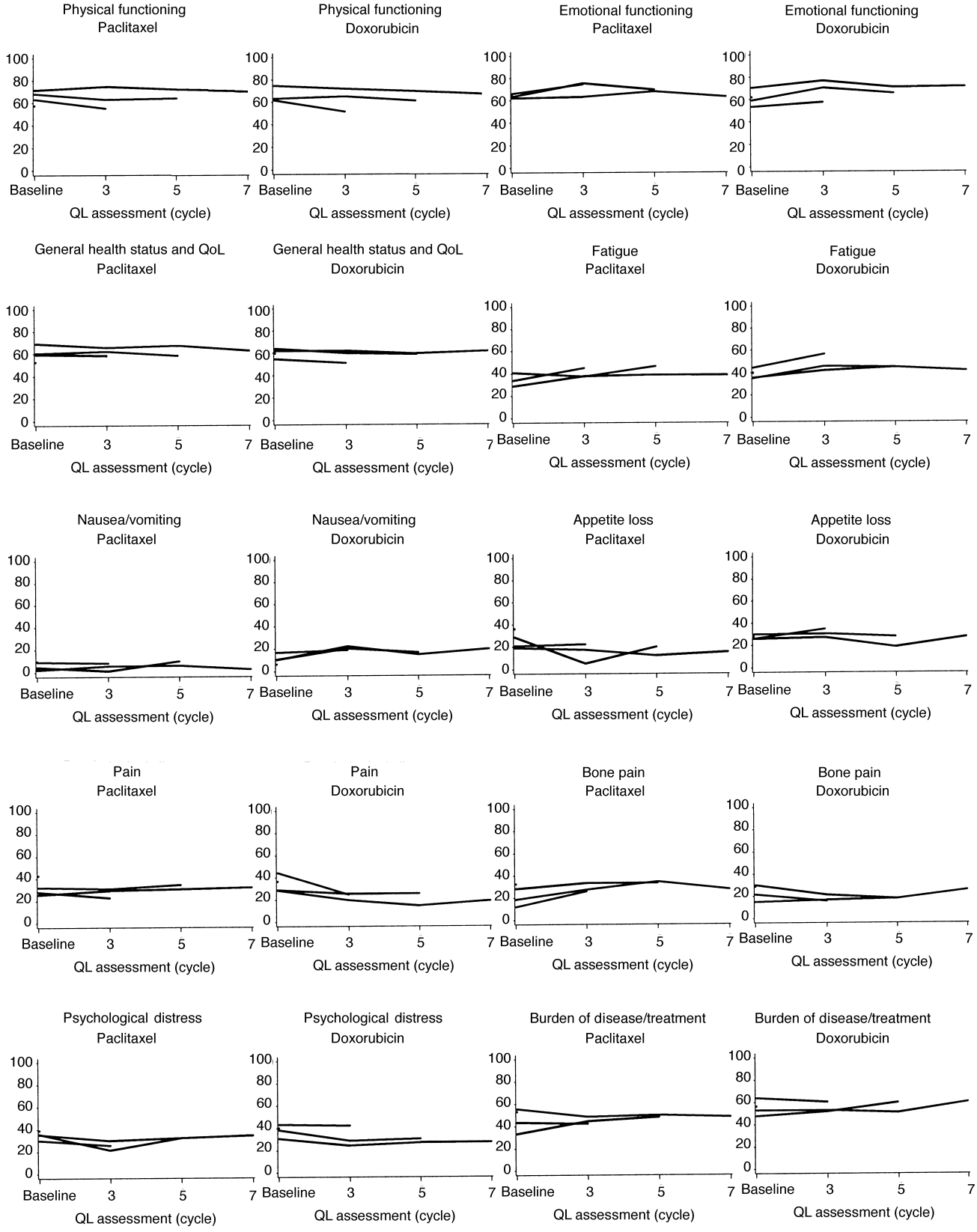


Fig. 4. Mean scores for selected quality of life (QL) dimensions by dropout time.

4. Discussion

The baseline QL scores for patients in this study were similar to those from reference patients with advanced breast cancer from clinical trials, although global QL was higher in the study group and symptom scores were marginally lower. This reflects the selection of patients with relatively good performance status, which is not uncommon in clinical trials. The clinical data unexpectedly showed that patients receiving doxorubicin had a median progression-free survival almost twice as long as that of patients receiving paclitaxel [7] suggesting a difference in antitumour efficacy between the two drugs. The resulting difference in the duration of first-line treatment introduced a potential bias into the QL data. Fewer patients remained in the paclitaxel arm, and these were likely to have been in better health due to selection bias (survival of the fittest) than those remaining in the doxorubicin arm. To minimise bias we allowed for delays in treatment, but even so by cycle 7 there were 21% fewer patients remaining on paclitaxel and their compliance with completion of questionnaires was 23% QLQ-C30 and 19% RSCL higher than for those receiving doxorubicin. It was not possible to adjust for the bias because of relatively poor compliance, particularly after cycle 3.

Many authors have drawn attention to the problems of compliance in the completion of questionnaires in clinical trials, and shown that institutional and administrative factors tend to be more influential than patient factors, at least until performance status deteriorates [15–17]. The present study illustrates that missing data are problematic since evaluation of data quality showed that analyses beyond 9 weeks after randomisation were not justified because of the risk of bias.

Cross-sectional analysis of QLQ-C30 data at the end of cycle 3 showed that patients receiving doxorubicin had significantly more nausea/vomiting and loss of appetite than those in the paclitaxel arm. This toxic effect should be able to be mitigated by greater attention to antiemetic premedication allowing those receiving doxorubicin to feel better with their gains in progression-free survival and less burdened by the disease and treatment. There were more rashes in the paclitaxel arm at cycle 3. Failure to study other treatment-specific toxicities means that we are not able to say which treatment was more toxic from the patients' perspective. The only evidence available is the borderline significantly greater burden of disease and treatment in the doxorubicin arm at the end of cycle 3 and the greater proportion of patients on doxorubicin who discontinued treatment because of excessive toxicity after the third cycle. Therefore, our overall impression is that doxorubicin may have been more toxic but also more active. It is not known whether patients are prepared to accept potentially greater toxicity for a gain of 3 months in progression-free survival with doxorubicin.

We showed with the QLQ-C30 that those in the doxorubicin arm had a trend towards less pain between baseline and cycle 3, and this appeared to be true specifically for bone pain. Fig. 4 indicates a trend towards decreasing mean pain score (i.e. less pain) in the doxorubicin arm and increasing mean pain score in (especially bone pain) the paclitaxel arm in those receiving more than three cycles of treatment, suggesting that the more active treatment has a direct effect on disease-specific pain. In this respect doxorubicin appeared to meet one of the goals of chemotherapy in advanced cancer by relieving this symptom and potentially delaying or preventing problems associated with bone disease. However, the study did not specifically assess paclitaxel's side-effects of peripheral neuropathy and myalgia, which might have contributed to the increase in pain in the paclitaxel arm.

Longitudinal changes between baseline and cycle 3 suggested that only a minority of patients recorded improved scores on QLQ-C30 dimensions. The exceptions were emotional function, which improved in over half of the patients and pain which improved in 40–50%. It could be argued that the benefits in pain control might have been achieved just as well with palliative care, and perhaps with less toxicity. Improvements in emotional function might merely reflect the fact that something was being done irrespective of what it was, and could be an indication of hope in a life-threatening situation [4]. The descriptive presentation of longitudinal changes in QL by dropout time shows the changes already described for emotional function/psychological distress, nausea/vomiting, appetite loss and pain/bone pain, and suggests a trend towards greater burden of disease and treatment and lower general health status and QL for doxorubicin. Our analysis is likely to have overestimated the benefits of both treatments, since patients who deteriorated and did not complete QL evaluations were not included.

Like us, neither Coates and colleagues [2] nor Richards and associates [3] succeeded in evaluating QL beyond 9–12 weeks in the clinical trial setting, whilst Ramirez and associates [4] were able to evaluate non-trial patients after 4–6 months. Because we were only able to analyse QL reliably over a period of 9 weeks (out of a median survival of approximately 1.5 years after entry into the trial) we do not know whether extending the time to progression in the doxorubicin arm relative to paclitaxel maintains QL. To assess the benefits of prolonging time to progression it would be necessary to obtain serial QL evaluations after the disease has progressed.

This study has shown that future research should give detailed attention to data quality in studies with QL endpoints so that a reliable analysis of QL can be performed over the whole period of advanced disease. Validated disease- and treatment-specific questions

should be used, and QL observations need to be extended until death. It is important to identify those patients least likely to respond so as to spare them the burden of treatment, and to ensure that all patients receive optimal treatment for toxicity, pain control and palliative care.

Acknowledgements

This study was sponsored by Bristol-Myers Squibb International Corporation, Pharmaceutical Research Institute. J.A. Kramer was supported at the EORTC by the Lady Grierson Research Fellowship.

References

1. Harris JR, Morrow M, Norton L. Malignant tumors of the breast. Chapter 36.2. In DeVita VT, Hellman S, Rosenberg SA, eds. *Cancer: Principles & Practice of Oncology*, 5th edn. Lippincott-Raven, Philadelphia, 1997, 1602.
2. Coates A, Gebski V, Bishop JF, *et al.* for the Australian–New Zealand Breast Cancer Trials Group, Clinical Oncological Society of Australia. Improving the quality of life during chemotherapy for advanced breast cancer. A comparison of intermittent and continuous treatment strategies. *N Engl J Med* 1987, **317**, 1490–1495
3. Richards MA, Hopwood P, Ramirez AJ, *et al.* Doxorubicin in advanced breast cancer: influence of schedule on response, survival and quality of life. *Eur J Cancer* 1992, **28A**, 1023–1028.
4. Ramirez AJ, Towlson KE, Leaning MS, Richards MA, Rubens RD. Do patients with advanced breast cancer benefit from chemotherapy? *Br J Cancer* 1998, **78**, 1488–1494.
5. de Haes JCJM, van Knippenberg FCE, Neijt JP. Measuring psychological and physical distress in cancer patients: structure and application of the Rotterdam Symptom Checklist. *Br J Cancer* 1990, **62**, 1034–1038.
6. 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.
7. Paridaens R, Biganzoli L, Bruning P, *et al.* Taxol® versus doxorubicin as first-line single agent chemotherapy for metastatic breast cancer — an EORTC randomized study with crossover. *J Clin Oncol* 2000, **18**, 724–733.
8. Fayers P, Aaronson N, Bjordal K, Curran D, Groenvold M. *EORTC QLQ-C30 Scoring Manual.* 2nd edn. Brussels, EORTC Quality of Life Study Group, 1999.
9. de Haes JCM, Olschewski M, Fayers P, *et al. Measuring the quality of life of cancer patients with The Rotterdam Symptom Checklist (RSCL). A Manual.* University of Groningen, The Netherlands, Northern Centre for Healthcare Research, 1996.
10. Fayers P, Weeden S, Curran D, on behalf of the EORTC. *Quality of Life Study Group. EORTC QLQ-C30 Reference Values.* Brussels, EORTC Quality of Life Study Group, 1998.
11. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* 1975, **31**, 103–115.
12. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958, **53**, 457–481.
13. Hollander M, Wolfe DA. *Nonparametrics Statistical Methods.* New York, John Wiley & Sons, 1973.
14. SAS/STAT Software. *Changes and Enhancements 6.11.* Cary NC, USA, SAS Institute Inc., 1996.
15. Hurny C, Bernhard J, Joss R, *et al.* for the Swiss Group for Clinical Cancer Research SAKK. Feasibility of quality of life assessment in a randomized phase III trial of small cell lung cancer — a lesson from the real world. *Ann Oncol* 1992, **3**, 825–831.
16. Hopwood P, Stephens RJ, Machin D, for the MRC. Lung Cancer Working Party. Approaches to the analysis of quality of life data: experiences gained from a Medical Research Council Lung Cancer Working Party palliative chemotherapy trial. *Qual Life Res* 1994, **3**, 339–352.
17. Hopwood P, Harvey A, Davies J, *et al.*, on behalf of the Medical Research Council Lung Cancer Working Party and the CHART Steering Committee. Survey of the administration of quality of life (QL) questionnaires in three multicentre randomised trials in cancer. *Eur J Cancer* 1998, **34**, 49–57.