Pergamon

European Journal of Cancer, Vol. 35, No. 6, pp. 913–920, 1999

Printed in Great Britain
0959-8049/99/$ - see front matter

PII: S0959-8049(99)00028-3

Original Paper

Quality of Life in Postmenopausal Patients with Breast Cancer after Failure of Tamoxifen: Formestane Versus Megestrol Acetate as Second-line Hormonal Treatment

J. Bernhard,[1] M. Castiglione-Gertsch,[1] S.-F. Hsu Schmitz,[1] B. Thürlimann,[2] F. Cavalli,[3] R. Morant,[2] M.F. Fey,[4] H. Bonnefoi,[5] A. Goldhirsch[6] and C. Hürny[7] for the Swiss Group for Clinical Cancer Research (SAKK)

[1]SIAK Coordinating Centre, CH-3008 Bern; [2]Department of Medicine C, Kantonsspital, St Gallen; [3]Cantonal Institute of Oncology, Ospedale San Giovanni, Bellinzona; [4]Institute of Medical Oncology, Bern; [5]Division of Gynaecology, Centre Médical Universitaire, Genève; [6]Cantonal Institute of Oncology, Ospedale Civico, Lugano; and [7]Bürgerspital, St Gallen, Switzerland

The Swiss Group for Clinical Cancer Research (SAKK) compared efficacy and toxicity of formestane (250 mg intramuscularly (i.m.) every 2 weeks) versus megestrol acetate (MGA; 160 mg orally daily) as second-line treatment in postmenopausal patients with advanced breast cancer and disease progression while on tamoxifen treatment in a randomised trial (Thürlimann B, Castiglione M, Hsu Schmitz SF, *et al. Eur J Cancer* 1997, 33, 1017–1024). Quality of life (QL) was evaluated as a secondary endpoint ($n = 177$). Overall, 83% (669/805) of expected QL forms were received, 88% (155/177) at baseline, 88% (402/457) on study treatment, and 65% (112/171) at treatment failure. Patients with no impairment in performance status reported better physical well-being ($P = 0.0001$), mood ($P = 0.0007$) and coping ($P = 0.03$), and less tiredness ($P = 0.0001$) and appetite/sense of taste disturbance ($P = 0.0001$) at baseline. After adjustment for baseline, there was no statistically significant difference in QL by treatment. Baseline QL was strongly predictive for QL under treatment but not for time to treatment failure. In conclusion, the question of whether oestrogen deprivation (e.g. formestane) or addition of progesterone (MGA) has a more beneficial impact on QL needs further investigation. The subjective experience of second-line endocrine treatment varies considerably as a consequence of the large variation in the individual course of the disease and has to be judged on an individual basis.

Key words: advanced breast cancer, formestane, megestrol acetate, aromatase inhibitor, quality of life, prognostic factors
Eur J Cancer, Vol. 35, No. 6, pp. 913–920, 1999

INTRODUCTION

IN METASTATIC breast cancer, currently available treatments are mainly palliative and quality of life (QL) is a primary endpoint. In postmenopausal patients, tamoxifen is the most frequently used endocrine agent in this situation. However, the increasing use of this drug in the adjuvant setting has led to the development of second-line endocrine therapies for patients with advanced disease.

After tamoxifen failure, aromatase inhibitors and progestins are widely used as second-line treatment, mainly in patients who responded to first-line endocrine therapy. Compared with chemotherapy, relatively good palliation has been reported in these patients [1]. Aminogluthetamide, the first non-steroidal aromatase inhibitor, has frequently been used, despite its unfavourable toxicity profile. Recently, more

Correspondence to J. Bernhard, e-mail: jbernhard@sakk.ch
Received 17 Jul. 1998; revised 5 Jan. 1999; accepted 26 Jan. 1999.

selective aromatase inhibitors (steroidal and non-steroidal) have been developed to increase efficacy and reduce toxicity of endocrine treatment in advanced breast cancer.

4-hydroxyandrostendione (formestane) was the first steroidal compound and the first aromatase inhibitor of these more selective second generation agents. Intramuscular (i.m.) formestane has been shown to be effective and well tolerated as second-line treatment in phase I/II studies [2]. As a first-line treatment in postmenopausal patients with advanced breast cancer, equivalent treatment efficacy and side-effects in comparison with tamoxifen have been reported [3].

The Swiss Group for Clinical Cancer Research (SAKK) compared efficacy and toxicity of formestane (250 mg i.m. every 2 weeks) versus megestrol acetate (MGA; 160 mg orally daily) as second-line treatment in postmenopausal patients with advanced breast cancer and disease progression while on tamoxifen treatment in a multicentre randomised trial (SAKK 20/90) [4]. There was no significant difference between the treatments with regard to response rate and time to treatment failure. Overall, toxicity was similar in both arms, but the two drugs had different toxicity profiles [4].

We evaluated patient-rated QL by treatment as the secondary endpoint of this trial. In addition, the predictive value of patient characteristics for QL at baseline and under treatment were investigated. With respect to potential supportive interventions we also investigated whether a patient's QL at baseline would predict QL under second-line treatment.

Coates and colleagues have reported significant independent prognostic value of QL scores for survival in pre- and postmenopausal patients with advanced breast cancer [5]. We examined these associations likewise in this sample of postmenopausal patients with advanced, less endocrine sensitive disease by testing the prognostic value of QL scores at baseline for time to failure.

PATIENTS AND METHODS

The trial

The trial was planned to test differences in time to treatment failure and toxicity between formestane (250 mg i.m. every 2 weeks) and MGA (160 mg orally daily). Details are described elsewhere [4]. In brief, postmenopausal patients with objective evidence of histologically and/or cytologically proven breast cancer and measurable or evaluable advanced disease were eligible. Patients must have had failed prior adjuvant and/or palliative treatment with tamoxifen, irrespective of the best response. Prior adjuvant chemotherapy was allowed, but had to be completed more than 12 months before enrolment in the trial. Further inclusion criteria were SAKK/ECOG performance status 0–2 and appropriateness of endocrine therapy as decided by the treating physician. Informed consent was obtained from all patients. Patients were randomised after being stratified for institution, type of treatment and best response to prior tamoxifen (adjuvant tamoxifen versus palliative tamoxifen with partial response (PR) or complete response (CR) versus palliative tamoxifen with no change (NC) or progressive disease (PD) versus palliative tamoxifen with unknown or not assessable response). In November 1992, the inclusion criteria were modified to allow pretreatment with one chemotherapy regimen for advanced disease. Treatment failure was defined by documented disease progression, patient refusal or unacceptable toxicity. Patients were followed up until treatment failure, thus information on overall survival was not available.

QL assessment

Linear analogue self-assessment (LASA) indicators were used to measure selected components of QL: physical well-being, mood [6, 7] and coping/perceived adjustment (PACIS) [8] as global indicators; tiredness, appetite/sense of taste disturbance [7], hot flushes [9] and dizziness as indicators of symptoms and side-effects. The timeframe was related to the past 2 weeks. Conceptual and methodological issues of this approach are summarised elsewhere [9].

In this advanced disease multicentre trial, a detailed QL assessment was not feasible. We therefore followed a different strategy: given that the primary purpose was to compare the two treatments with regards to patients' QL overall—not to determine specific reactions associated with the two drugs—we defined the limited set of indicators described above to evaluate outcome. Responses on the global LASA indicators are expected to reflect the summation of the individual meaning and importance of various factors for each patient [9]. Although less precise for specific treatment effects, these measures are expected to be sensitive to the wide spectrum of reactions seen in patients on endocrine treatments and should detect these changes on single dimensions, allowing for comparison across treatments. In patients with advanced breast cancer, some of these indicators have shown significant independent prognostic value for survival [5]. They discriminated between effects of endocrine and cytotoxic treatment and between responders and non-responders [10] and between different chemotherapy regimens [11]. The QL form was used in three languages (German, French, Italian).

The QL assessment schedule was defined with regards to the known large variation in time to treatment failure (i.e. due to disease progression) in these patients: QL indicators were assessed at randomisation, 1 month, 3, 5, 7, 9 and 11 months and at treatment failure; this strategy is suitable for summary measures over time. The evaluation at month 1 was designed to assess toxicity in patients with early failure. Patients were instructed by nurses or physicians to fill in the QL forms at clinical visits in the hospital, using a standardised and pretested written example. Reasons for missing QL data were recorded. Sociodemographic data were assessed by nurses or physicians at study entry.

All QL indicators were scored by measuring in millimetres from 0 to 100, with higher numbers reflecting better QL (e.g. less symptoms).

Statistical analyses

Submission rates of QL forms were defined as the ratio of the number of received completed forms to the number of expected forms at baseline, on study treatment and at treatment failure. All available QL data were used for the following analyses.

Covariates at baseline. The impact of biomedical, sociodemographic and language factors on baseline QL scores was explored by analysis of covariance (ANCOVA). Biomedical factors included age (as a continuum), previous tamoxifen treatment (adjuvant versus palliative) and performance status (0 versus 1 or 2) at randomisation. Sociodemographic factors included marital status, living situation, employment status and educational level (see Table 1). We expected younger patients, those with poorer performance status and those living alone to report worse QL. The language of the QL form was included in this analysis because language/culture had a substantial impact on these indicators in the adjuvant setting [12, 13].

Covariates under treatment. The impact of biomedical factors and the time interval (days, in a natural logarithm scale) between baseline and the last available QL assessment on study treatment before treatment failure (LAST assessment) [14] on this LAST assessment was explored by ANCOVA. Biomedical factors included age, previous tamoxifen treatment, performance status at randomisation and treatment arm. The corresponding QL score at baseline was used as an additional covariate in ANCOVA to assess the change between baseline and LAST assessment. Similarly, the impact of these biomedical factors and the time interval between the LAST assessment and failure on the QL at failure was explored; this analysis was adjusted for the LAST score to assess the change between LAST assessment and failure.

QL between treatments. First, all available QL scores were evaluated separately for each time point within each treatment arm. Due to the high attrition rate (i.e. rapidly diminishing patient numbers), QL was evaluated only up to the seventh month. Second, the effect of treatment on QL at each time point on study treatment, including the LAST assessment, and between the LAST assessment and treatment failure was investigated by ANCOVA, with the baseline and LAST scores, respectively, as additional covariates to assess changes.

In addition, the prevalence of weight gain and physician-rated toxicity was investigated by Fisher's exact test. The effects of these toxicities on QL were investigated for the LAST assessment by ANOVA. Given that toxicity may differ between treatment arms, interaction between treatment and toxicity was also included.

QL as a prognostic factor for time to treatment failure. The QL scores at baseline were included in Cox regression analyses of time to treatment failure controlling for age, previous tamoxifen treatment, performance status and liver metastasis at randomisation and treatment arm. We hypothesised that scores of physical well-being, tiredness and coping predict time to treatment failure. Univariate Cox regressions on baseline QL alone were also performed.

Due to the exploratory nature of the analyses, no adjustment was made for multiple testing. To keep the results easily comparable, no transformation was applied to QL scores in some ANCOVA with non-normal distributions or unequal variance. In ANCOVA, the regression coefficient (slope) for a continuous covariate is denoted by β, while the difference in the intercept between groups of a categorical covariate is denoted by $\Delta\beta$. All tests were two-sided.

No special approach was applied to deal with missing values. The statistical inferences might be inappropriate if data were not missing completely at random [15]. However, as described in the Results, local administrative problems were the main reason for missing QL forms and probably not related to outcome. Hence, missing completely at random might be justified. Numbers of patients in analyses varied across time points and QL measures due to attrition or missing data. The LAST assessment was considered as a summary measure over time to account for attrition.

RESULTS

Description of the sample

The biomedical characteristics of the total sample ($n = 177$ eligible randomised patients) are described elsewhere [4]. In summary, approximately 11% of all patients had negative oestrogen, 22% had negative progesterone receptor status, 51% had received tamoxifen as adjuvant and 49% as palliative treatment. The two arms were well balanced at study entry with the exception of weight and tumour localisation: patients in the formestane arm had a significantly higher mean weight and significantly less frequent liver metastases. The median age was 65 years (range 43–87 years) [4].

170 (96%) of the 177 patients completed at least one QL form and were included in this analysis. Overall, 83% (669/805) of expected QL forms were received, 88% (155/177) at baseline, 88% (402/457) sometime during study treatment (varying from 1 to 11 months) and 65% (112/171) at treatment failure. The distribution of reasons for missing forms differed with respect to status of disease: local administrative problems were the main reason for missing data, with 48% before and 39% at treatment failure; refusal by patients was the second most important reason, with 24% before and 34% at failure. There was no statistically significant difference in patient characteristics between those with baseline QL data and those without. Completeness of received QL forms was high, with 99% of all scale scores evaluable for analysis.

Clinical and sociodemographic characteristics of the sample with QL data are summarised in Table 1. More than half (54%) of the patients had no impairment in performance status at randomisation. Half the patients were previously treated with adjuvant tamoxifen. The majority was married (59%) and/or lived together with a partner (62%). Most patients had no professional degree (43%) or received an apprenticeship (44%). Most patients were housewives without an external job (48%) and approximately a third (34%) of all patients was on sick leave or retired.

Table 1. *Patient characteristics at study entry* (n = 170)

	n (%)
Performance status	
0	91 (54)
1 or 2	79 (46)
Previous treatment with tamoxifen	
Adjuvant	85 (50)
Palliative	85 (50)
Treatment arm	
Formestane	89 (52)
Megestrol acetate	81* (48)
Marital status	
Married	101 (59)
Separated/divorced or widowed	46 (27)
Single	23 (14)
Living situation	
With partner	106 (62)
With next of kin or other	21 (12)
Alone	43 (25)
Employment status	
Full or part-time job	31 (18)
On sick leave or retired	58 (34)
Housewife without external job	81 (48)
Training/education†	
No professional degree	72 (43)
Apprenticeship	75 (44)
High school, business school, technical college or academic	22 (13)

*1 patient refused treatment. †No information in 1 case.

Covariates at baseline

The impact of biomedical, sociodemographic and language factors on baseline QL scores was explored. Overall, performance status had the strongest impact: patients with no impairment in performance status reported better physical well-being ($\Delta\beta = 26.4$, $P = 0.0001$), mood ($\Delta\beta = 16.0$, $P = 0.0007$) and coping ($\Delta\beta = 10.3$, $P = 0.03$) and less tiredness ($\Delta\beta = 19.1$, $P = 0.0001$) and appetite/sense of taste disturbance ($\Delta\beta = 17.4$, $P = 0.0001$); $\Delta\beta$ represents the difference in the corresponding baseline QL score between performance status 0 versus 1 or 2. For example, patients with no impairment reported an increase of 16 points in mood on average, compared with those with performance status 1 or 2. Figure 1 presents the distributions of physical well-being, mood and tiredness at baseline grouped by performance status.

The type of previous tamoxifen treatment and age had a less pronounced effect. Patients with adjuvant tamoxifen reported more hot flushes ($\Delta\beta = -8.7$, $P = 0.05$) and worse coping ($\Delta\beta = -11.0$, $P = 0.02$) than those with palliative tamoxifen. Older patients indicated less hot flushes ($\beta = 0.8$, $P = 0.003$) and better coping ($\beta = 0.9$, $P = 0.0005$) than younger patients; concerning the age effect, β (slope) represents the change in hot flushes or coping for each unit (i.e. year) change in age.

Marital status had a consistent impact on physical well-being ($P = 0.03$), hot flushes ($P = 0.04$) and dizziness ($P = 0.01$): compared with married women, divorced/separated or widowed women indicated worse scores ($\Delta\beta = -12.7, -15.3, -10.7$) and single women indicated better scores ($\Delta\beta = 8.5, 1.9, 3.3$) in these primarily physical aspects; there was no effect on mood or coping. Living situation, employment status, educational level and language had no significant impact on baseline QL (data not shown).

Overall, patient characteristics explained a moderate part of variance in baseline QL scores, with R^2 ranging from 11% in dizziness to 25% in physical well-being.

Covariates under treatment

The impact of biomedical factors and time interval (i.e. baseline to LAST assessment) on the LAST assessment ($n = 143$) was explored with adjustment for the corresponding baseline scores. The biomedical factors had no impact (data not shown). In addition, there was no significant association between initial performance status (0 versus 1 or 2) and response (CR/PR) under treatment (data not shown). Examining the impact of biomedical factors for each time point separately revealed some influences. Initial performance status had the most dominant impact during the first 3 months for most of the QL measures (data not shown).

The median time interval between baseline and the LAST assessment across both arms was 99 days, with a range from 14 days to 11 months. Taking into account all randomised patients, irrespective of time to failure, changes between baseline and LAST assessment indicated a marginal improvement in QL (Figure 2); it has to be noted that the LAST assessment is a rather conservative estimate. The length of this interval had an effect on physical well-being ($\beta = 5.5$, $P = 0.01$), appetite/sense of taste disturbance ($\beta = 3.5$, $P = 0.06$) and coping ($\beta = 3.6$, $P = 0.06$); β represents the change in scores for each unit change of interval (days) in a natural logarithm scale. In other words, patients with a longer time interval, reflecting a longer time to treatment failure, reported better scores under treatment. Finally, baseline scores were strongly predictive for the LAST assessment in all QL measures ($0.2 \leq \beta \leq 0.5$, $0.0001 \leq P \leq 0.03$).

Although the initial biomedical factors had no impact on the LAST assessment, they appeared to have an influence on

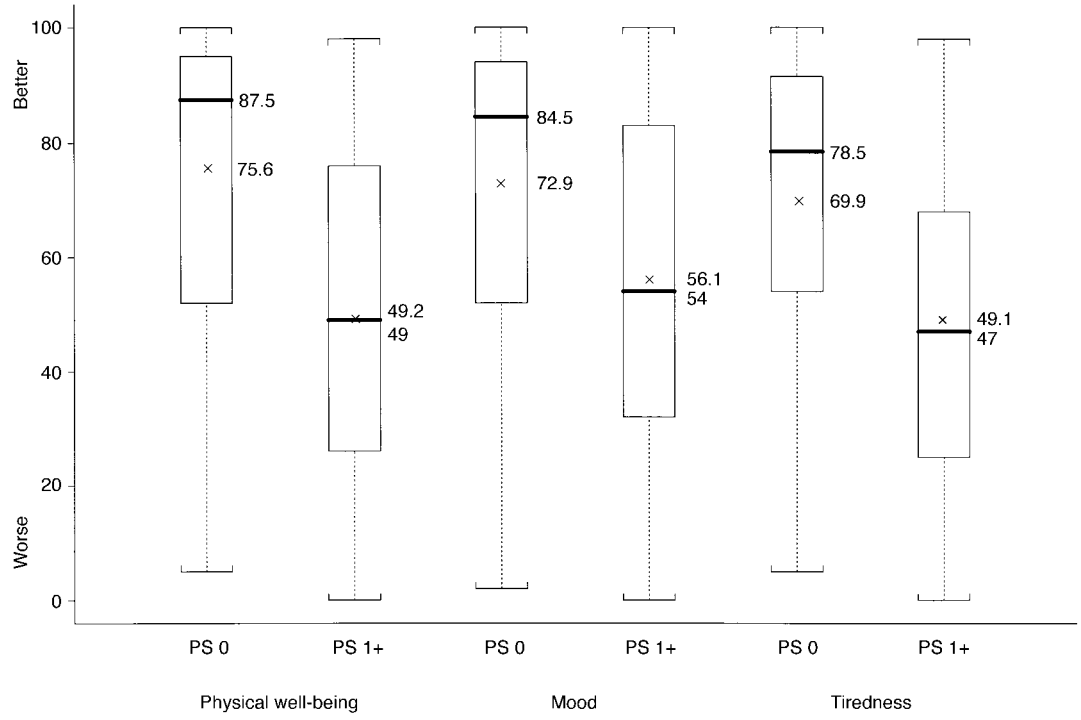

Figure 1. Impact of initial performance status on baseline quality of life (QL) indicators. The bold line indicates the median, 'x' the mean. All QL indicators ranged from 0 to 100, with higher scores reflecting better QL (e.g. less symptoms). The height of the boxes represents the distance between the 25th and 75th percentiles.

QL at treatment failure (after adjustment for LAST assessment scores). Patients with impaired performance status at randomisation reported worse physical well-being ($\Delta\beta = -10.9$, $P = 0.02$) and more tiredness $\Delta\beta = -13.8$, $P = 0.004$) and hot flushes ($\Delta\beta = -8.8$, $P = 0.03$) at failure. Those with previous adjuvant tamoxifen treatment indicated more tiredness ($\Delta\beta = -9.1$, $P = 0.05$). Patients receiving formestane rated more hot flushes than those receiving MGA ($\Delta\beta = -9.1$, $P = 0.03$).

The median time interval between the LAST assessment and treatment failure across both arms was 56 days, with a range from 13 to 676 days. Overall, the QL scores at failure were comparable to those at baseline or slightly worse (data not shown). With increasing interval, patients reported more deterioration in mood ($\beta = -6.7$, $P = 0.02$) but physical measures were not affected. The LAST assessment scores were also strongly predictive for QL at failure in all measures ($0.4 \leq \beta \leq 0.8$, $0.0001 \leq P \leq 0.0006$). It has to be noted that the submission rate of QL forms was considerably lower and the rate of patient refusals was higher at failure than before. The available data may represent a selection bias of patients possibly in a better (emotional) state. QL at failure was most probably considerably worse than can be estimated by the available data; therefore, more deterioration at failure is expected compared with the improvement under treatment. Comparisons of baseline QL between patients with QL data at failure and those without showed no difference; an exception was hot flushes.

In summary, patients' QL under treatment was mainly predicted by baseline scores and the time interval under study treatment.

QL between treatments

There was no statistically significant difference in baseline QL measures between treatments. Figure 3 shows the median coping scores by treatment during the first 7 months. The numbers of available and expected scores are displayed. There was a steady decline in the number of patients on study treatment. The sample decreased by half between months 3 and 5. This is in accordance with the median time to treatment failure (mainly due to disease progression) of 120 days in the formestane and 111 days in the MGA arm in the total sample [4]. The better health status of the remaining patients was reflected in the continuous improvement in coping scores over time (Figure 3), in contrast to the difference between baseline and the LAST assessment in all randomised patients (Figure 2). Considering those patients whose LAST assessment was at month 7 or later, there was also a constant improvement in coping over time, and those receiving MGA had consistently better scores than those receiving formestane, although the difference was not statistically significant.

There were some differences in absolute QL scores between treatments at various time points (higher scores indicating better QL): patients receiving MGA reported less appetite/sense of taste disturbance at month 1 (formestane/

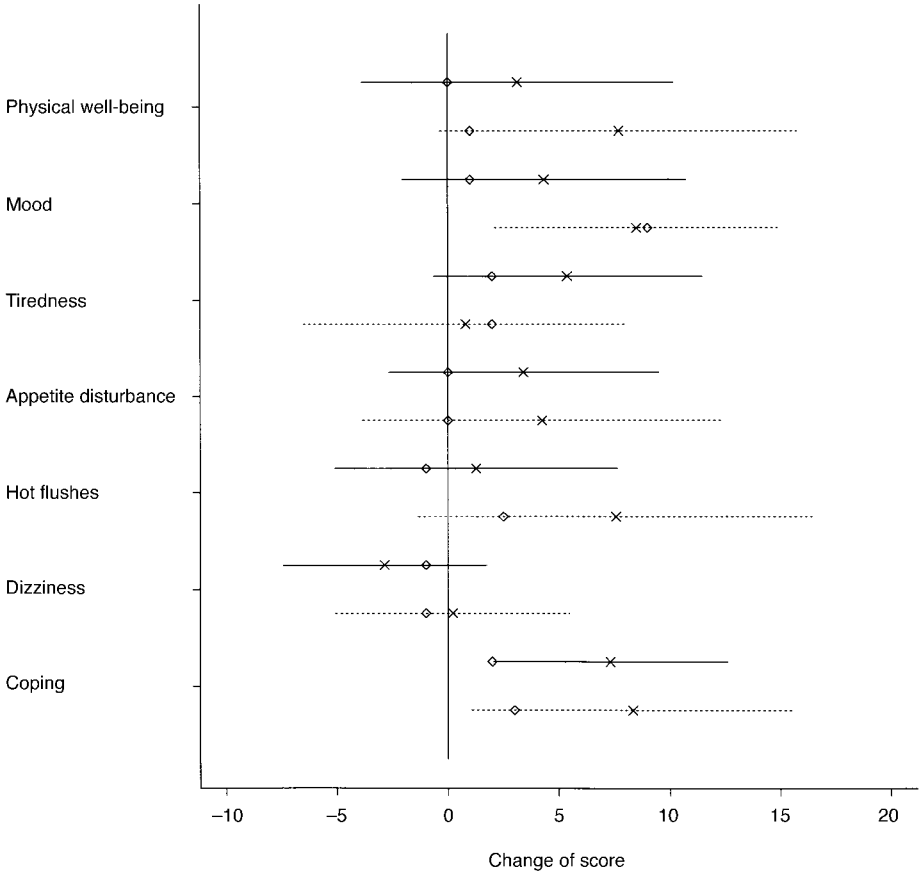

Figure 2. Mean (x) and median (◇) changes with 95% confidence intervals (CI) in quality of life (QL) indicators between baseline and LAST assessment by treatment. Positive differences indicate an improvement. Solid lines, formestane; broken lines; MGA.

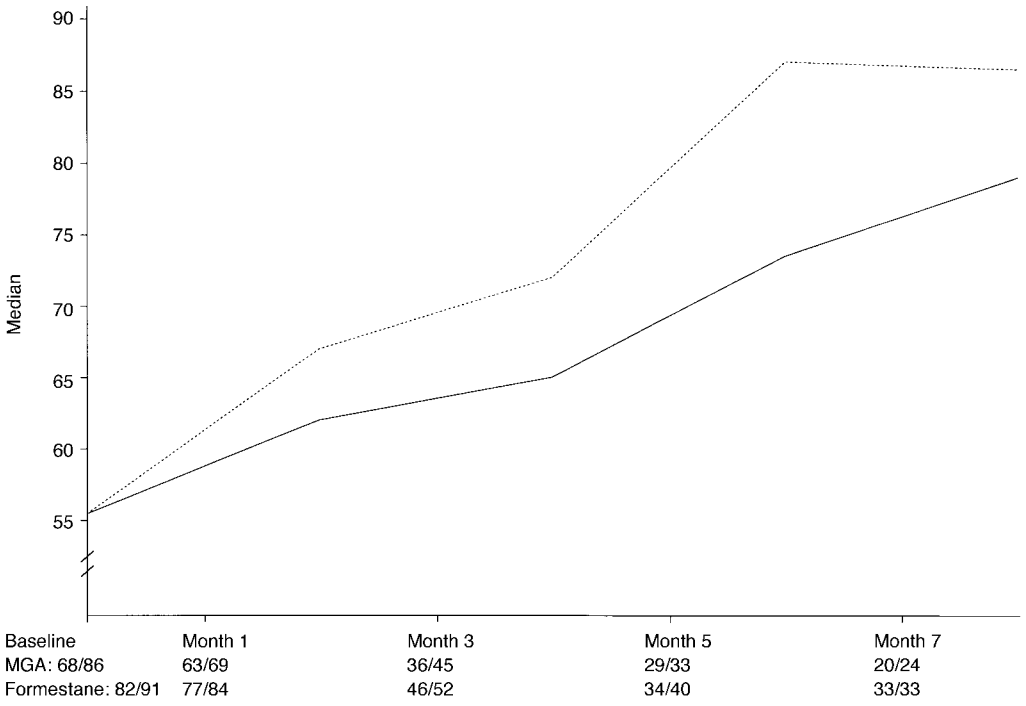

Figure 3. Median coping (PACIS) scores by treatment during the first seven months. The coping indicator ranged from 0 to 100, with higher scores reflecting better coping. Solid line, formestane; broken line, MGA.

MGA: $n = 78/65$, medians = 83.5/90), less tiredness at month 3 ($n = 47/36$, medians = 66/82), better coping at month 5 ($n = 34/29$, medians = 73.5/87) and less hot flushes at month 7 ($n = 33/19$, medians = 86/94). At treatment failure, patients receiving MGA reported less hot flushes ($n = 57/51$, median = 80/88). Although there were consistent trends over time between treatment arms, the results of ANCOVA with adjustment for baseline did not demonstrate a statistically significant treatment effect on QL measures at any time point. The results stayed the same if liver metastasis was also included in the model to account for the unbalanced tumour localisation at baseline.

At the LAST assessment, a weight gain of 3 kg or more was observed in 19 of 61 patients treated with MGA and in 8 of 77 treated with formestane. The effect of weight gain on QL indicators was not significant. Similarly, there was no statistically significant association between hypertension and QL. Severe sleep problems, thromboembolism and tremor/neurological disturbance occurred in ≤ 2 cases in each arm, thus no further analysis was performed.

QL as a prognostic factor for time to treatment failure

The results of the multivariate analyses of time to treatment failure on biomedical factors and QL scores are summarised in Table 2. Performance status at randomisation was the strongest predictor ($P = 0.005$): patients with no impairment in performance status had a longer time to failure. Baseline QL scores did not predict time to failure in univariate (data not shown) or in multivariate analyses.

DISCUSSION

At present, the selection of endocrine agents for patients with advanced breast cancer is mainly based on their toxicity profiles. We evaluated the impact of formestane versus MGA as second-line endocrine treatment on patient-rated QL.

Although there were some differences in QL between treatments consistently in favour of MGA at various time points, after adjustment for baseline QL the differences were no longer statistically significant in this relatively small sample. This finding is partly in agreement with those of the clinical report [4]. Overall, there was no difference in the frequency of physician-rated side-effects by treatment. However, MGA was associated with significantly more WHO toxicity grade ≥ 2 (i.e. moderate, severe or life threatening). This difference was not reflected by the patient-rated measures. It has to be noted that the response format of the LASA scales was not graded as in the WHO ratings, but patients were asked to give an overall estimation. The side-effects of endocrine treatments in advanced breast cancer are probably underestimated by physicians [16], especially in the case of low intensity. Patients may have a different perception

Table 2. Results of multivariate Cox regression of the time to treatment failure on patient characteristics, baseline quality of life and treatment

	Hazard ratio	*P*
Age	0.99	0.19
Performance status		
1 or 2 (compared with 0)	1.75	0.005
Previous tamoxifen treatment		
Palliative (compared with adjuvant)	1.10	0.62
Presence of liver metastasis (compared with none)	1.52	0.09
Treatment		
Megestrol acetate (compared with formestane)	1.29	0.15
Physical well-being	1.00	0.35
Tiredness	1.00	0.38
Coping	1.00	0.68

and give relatively more weight to those toxicities which have a high prevalence and are judged as mild by the treating physician.

A further aspect needs to be considered: the two agents had different modes of application. Taking a tablet daily or having an injection every other week can have different impacts on patient's perceptions of the treatment and disease. Hence, the question of whether oestrogen deprivation (formestane) or addition of progesterone (MGA) has a more beneficial impact on QL cannot be answered by our study.

Recently, two doses of anastrozole (1 and 10 mg orally daily), a highly selective non-steroidal third generation aromatase inhibitor, were compared with MGA (40 mg orally four times daily) in postmenopausal women with advanced breast cancer who progressed after prior treatment with tamoxifen [17]. Patients receiving MGA reported less psychological distress at 12 weeks, but this difference was not apparent at 24 weeks and its clinical relevance remained unclear. As in the current trial, the two drugs had different side-effect profiles which were obviously not reflected in patient-rated measures. The effect of MGA on QL in patients with stage IV breast cancer was also evaluated in a dose–response trial [18]. A comparison of 160, 800 or 1600 mg/day showed fewest side-effects and better QL in the lowest dose (i.e. the same as in the current trial), with an improvement within the first 3 months.

Do postmenopausal patients benefit from second-line hormonal treatment in terms of QL? Taking into account all patients with QL data at baseline and at least one assessment under treatment, irrespective of time to failure (i.e. LAST assessment), a marginal improvement in QL despite side-effects was indicated. Overall, patients with a longer time interval reported better scores than those with early failure. In view of the large variation in time to failure, the net QL benefit can vary considerably with the course of the disease, especially with respect to response of the tumour, and has to be judged on an individual basis. The profile of symptoms and side-effects may suggest a change of treatment at any point in time. The question of whether similar survival time implies similar QL in patients responding to endocrine therapy, as is the case in patients with stable disease for at least 6 months, PR or CR [19], is addressed elsewhere [20].

Factors influencing QL can be recognised at baseline. Besides clinical factors, initial performance status and patients' private situation also had an impact. Divorced/separated or widowed women indicated worse scores in physical aspects, but living alone or with others had no impact. Physical symptoms are known to be reported more by patients under higher psychological distress [21, 22]. The loss of the partner earlier may have contributed to acute distress in these patients.

Importantly and probably characteristic for a palliative situation, baseline scores were strongly predictive for QL under treatment. This is another indication that access to supportive and psychosocial interventions should be offered to patients with advanced disease as early as possible and not only at manifestation of high psychological distress (e.g. at failure of second-line treatment). Interestingly, although the initial biomedical factors had no impact on QL under treatment overall, they appeared to have an influence at failure. This is probably related to the different types of metastasis at study entry [4] which were controlled for under treatment and reflected in QL measures again with further disease progression.

Coates and colleagues reported significant independent prognostic value of baseline QL scores for survival in patients with advanced breast cancer, using some of the same LASA indicators as in the present trial [5]. In our patient population, which was less endocrine sensitive, performance status at randomisation was the strongest predictor for time to treatment failure. In contrast, although baseline QL measures did reflect performance status, they did not carry prognostic information for time to failure. The reason for this finding is not clear.

Submission rates of QL forms were good and comparable with those in similar advanced disease studies [23]. For example, in the anastrozole trial cited above [17], the submission rates were > 90% at baseline but reduced to > 75% at month 3 and > 70% at month 6. However, the QL evaluation, especially the treatment comparison, was clearly limited due to a high rate of treatment failures across the whole observation period, as discussed elsewhere [24]. In addition to evaluations at each time point adjusted for baseline, we also used the last available assessment before treatment failure [14]. This summary measure does not reflect the course over time but is still appropriate in consideration of the large variation in time to failure. The findings of the different approaches were mostly consistent, supporting their validity despite the missing data. Future second-line trials should be designed for larger samples [20], and shorter intervals for QL assessment may increase sensitivity of summary measures over time.

In conclusion, the question of whether oestrogen deprivation (e.g. formestane) or the addition of progesterone (MGA) has a more beneficial impact on QL needs further investigation. The subjective experience of second-line endocrine treatment varies considerably as a consequence of the large variation in the individual course of the disease and has to be judged on an individual basis. Baseline QL was strongly predictive for QL under treatment but did not carry prognostic information for time to failure.

1. Hagen MF, Thürlimann B. Hormontherapie des metastasierenden Mammakarzinoms. *Therapeutische Umschau* 1996, **53**, 820–828.
2. Toyce C. Formestane. *Drugs of the Future* 1993, **18**, 599–600.
3. Pérez Carrion R, Alberola Candel V, Calabresi F, *et al*. Comparison of the selective aromatase inhibitor formestane with tamoxifen as first-line hormonal therapy in postmenopausal women with advanced breast cancer. *Ann Oncol* 1994, **5**, 19–24.
4. Thürlimann B, Castiglione M, Hsu Schmitz SF, *et al*. for the Swiss Group of Clinical Cancer Research (SAKK). Formestane versus megestrol acetate in postmenopausal breast cancer patients after failure of tamoxifen: a phase III prospective randomised cross over trial of second-line hormonal treatment (SAKK 20/90). *Eur J Cancer* 1997, **33**, 1017–1024.
5. Coates A, Gebski V, Signorini D, *et al*. for the Australian New Zealand Breast Cancer Trials Group. Prognostic value of quality-of-life scores during chemotherapy for advanced breast cancer. *J Clin Oncol* 1992, **10**, 1833–1838.
6. Coates A, Fisher Dillenbeck CF, MCNeil DR, *et al*. On the receiving end II. Linear analogue self-assessment (LASA) in evaluation of aspects of the quality of life of cancer patients receiving chemotherapy. *Eur J Cancer Clin Oncol* 1983, **19**, 1633–1637.
7. Coates A, Glasziou P, McNeil D. On the receiving end III. Measurement of quality of life during cancer chemotherapy. *Ann Oncol* 1990, **1**, 213–217.
8. Hürny C, Bernhard J, Bacchi M, *et al*. The Perceived Adjustment to Chronic Illness Scale (PACIS): a global indicator of coping for operable breast cancer patients in clinical trials. Swiss

Group for Clinical Cancer Research (SAKK) and the International Breast Cancer Study Group (IBCSG). *Support Care Cancer* 1993, **1**, 200–208.
9. Bernhard J, Hürny C, Coates AS, *et al*. Quality of life assessment in patients receiving adjuvant therapy for breast cancer: the IBCSG approach. The International Breast Cancer Study Group. *Ann Oncol* 1997, **8**, 825–835.
10. Baum M, Priestman T, West RR, Jones EM. A comparison of subjective responses in a trial comparing endocrine with cytotoxic treatment in advanced carcinoma of the breast. *Eur J Cancer* 1980, **16A**(Suppl.), 223–226.
11. Coates A, Gebski V, Bishop JF, *et al*. for the Australian-New Zealand Breast Cancer Trials Group. 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.
12. Hürny C, Bernhard J, Gelber RD, *et al*. Quality of life measures for patients receiving adjuvant therapy for breast cancer: an international trial. The International Breast Cancer Study Group. *Eur J Cancer* 1992, **28A**, 118–124.
13. Bernhard J, Hürny C, Coates A, *et al*. for the International Breast Cancer Study Group. Factors affecting baseline quality of life in two international adjuvant breast cancer trials. *Br J Cancer* 1998, **78**, 686–693.
14. Tandon PK. Application of global statistics in analyzing quality of life data. *Stat Med* 1990, **9**, 819–827.
15. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. New York, Wiley, 1987.
16. Leonard RCF, Lee L, Harrison ME. Impact of side effects associated with endocrine treatments for advanced breast cancer: clinicians' and patients' perceptions. *The Breast* 1996, **5**, 259–264.
17. Jonat W, Howell A, Blomqvist C, *et al*. A randomised trial comparing two doses of the new selective aromatase inhibitor anastrozole (arimidex) with megestrol acetate in postmenopausal patients with advanced breast cancer. *Eur J Cancer* 1996, **32A**, 404–412.
18. Kornblith AB, Hollis DR, Zuckerman E, *et al*. Effect of megestrol acetate on quality of life in a dose-response trial in women with advanced breast cancer. The Cancer and Leukemia Group B. *J Clin Oncol* 1993, **11**, 2081–2089.
19. Robertson JFR, Willsher PC, Cheung KL, Blamey RW. The clinical relevance of static disease (no change) category for 6 months on endocrine therapy in patients with breast cancer. *Eur J Cancer* 1997, **33**, 1774–1779.
20. Bernhard J, Thürlimann B, Hsu Schmitz S-F, *et al*. Defining clinical benefit in postmenopausal patients with breast cancer under second-line hormonal treatment: does quality of life matter? *J Clin Oncol* (in press).
21. Watson D, Pennebaker JW. Health complaints, stress, and distress: exploring the central role of negative affectivity. *Psychol Rev* 1989, **96**, 234–254.
22. Manne SL, Sabbioni M, Bovbjerg DH, Jacobsen PB, Taylor KL, Redd WH. Coping with chemotherapy for breast cancer. *J Behav Med* 1994, **17**, 41–55.
23. Bernhard J, Cella DF, Coates AS, *et al*. Missing quality of life data in cancer clinical trials: serious problems and challenges. *Stat Med* 1998, **17**, 517–532.
24. Curran D, Bacchi M, Hsu Schmitz SF, *et al*. Identifying the types of missingness in quality of life data from clinical trials. *Stat Med* 1998, **17**, 739–756.