Eur J Cancer Clin Oncol, Vol. 25, No. 1, pp. 51–56, 1989.
Printed in Great Britain

0277–5379/89$3.00 + 0.00

Factors at Presentation Influencing the Prognosis in Breast Cancer

MAIRA CALEFFI,* IAN S. FENTIMAN* and BRIAN G. BIRKHEAD†

**Clinical Oncology Unit, Guy's Hospital, London, U.K. and †Clinical Operation Research Unit, Department of Statistical Science, University College London, U.K.*

Abstract—*Presentation data on* 607 *breast cancer patients treated by a variety of modalities at Guy's Hospital, London, have been analysed using the Cox proportional hazards model, to identify factors associated with length of overall survival. When deaths attributable to causes other than breast cancer were treated as censoring events, the significant factors were found to be* stage, mode of treatment, menstrual status *and* tumour size. *The analysis was repeated for the subgroup of* 326 *patients treated by modified radical mastectomy. Four variables*: stage, age at menarche, menstrual status *and* age *were found to be significantly associated with both overall survival and length of distant recurrence-free interval. These factors have been combined to create a prognostic index which has been used to define subgroups of patients with different prognosis. The index has been validated on a separate group of* 457 *modified radical mastectomy patients treated at the same hospital.*

INTRODUCTION

THERE IS an extensive literature discussing the factors at presentation which are of prognostic significance for the survival of patients with primary breast cancer. Several authors have attempted to construct staging criteria based on these factors to define groups carrying similar prognosis [1–3]. In the early 1960s the TNM classification system was introduced [4] and subsequent attempts to improve on its discriminatory power have been manifold. Cutler *et al.* [5] were among the first to study the effects on outcome of factors other than stage, including size of the primary tumour, histologic type, nuclear grade and the number of axillary nodes involved at presentation. They reported an inverse correlation between each of these and survival. In another study [6], Williams *et al.* reported a significant relationship between histological grade and lymph node status and the length of local recurrence-free interval.

It has also been suggested [7] that patients with slowly growing tumours have a survival advantage over those with more aggressive disease. Fisher *et al.* [8], however, failed to find any significant effect due to the length of symptoms prior to treatment on either overall survival or recurrence-free survival.

The importance of menstrual status has been the focus of many studies. MacMahon *et al.* [9] showed that, with the exception of patients under 35 years at presentation, premenopausal patients have better survival than postmenopausal patients. Other authors have reported a third group presenting within 5 years of the menopause, who have worst survival of all [10, 11]. The poor prognosis of perimenopausal patients has not been found in other series [12, 13].

With regard to age, Adami *et al.* [14] have recently shown that relative survival declines markedly after the age of 49 years, with women in the 50–59 years age group having the worst prognosis. Other studies [15–17] suggest that patients under 35 years have an even poorer survival.

Several risk factors for breast cancer have been identified in epidemiological studies, and the prognostic importance of some of these in the established disease has also been investigated. Wang *et al.* [18], for example, found that postmenopausal breast cancer patients with three or more children had significantly shorter disease-free interval and overall survival than those with fewer children.

In addition, Juret *et al.* [19] reported poorer survival prognosis for patients who started menstruating early (at less than 11 years of age), but other

Accepted 12 July 1988.
Correspondence and requests for reprints should be addressed to: I.S. Fentiman, Clinical Oncology Unit, Guy's Hospital, London SE1 9RT, U.K.
This work was funded by the Department of Health and Social Security and the Imperial Cancer Research Fund.

studies [18, 20] have failed to confirm this. Other intrinsic tumour factors such as histological grade and hormone receptor status have been found to be associated with time to recurrence and overall survival time [20–23].

While several prognostic factors for breast cancer have been reported, many of those mentioned above have been identified using univariate analysis techniques. These methods test separately each factor for significance, but take no simultaneous account of the prognostic importance of any others. In this paper, a multivariate analysis (based on the Cox proportional hazards model) has been used to examine the independent contributions of several factors in an attempt to explain the variability in the observed survival and recurrence-free intervals of a large series of patients. The factors identified as being significant have been combined into prognostic indexes which have been used to define groups of patients with different prognoses. In order that these scoring systems could have wide practical utility, the list of factors tested has been restricted to those which can be known at the time of initial treatment assignment.

PATIENTS AND METHODS

Patients

The study was based on a sample of 607 breast cancer patients selected from a sequential series of 1542 referred to the Breast Clinic, Guy's Hospital, London, between January 1974 and December 1979. The sample was originally chosen as a database for a study of breast cancer in the elderly, and comprised all patients over the age of 70 years referred during this period (230 patients), and randomly selected samples of the same size from patients in the age groups 30–49 years and 50–69 years. After excluding patients who had received any previous treatment, 607 patients remained.

Treatment

Patients were treated by a variety of currently accepted modalities which can be grouped under four major headings: mastectomy, non-mastectomy, radiotherapy and palliative treatment. A small number of these patients also received melphalan (23 patients), tamoxifen (three patients) or chemotherapy (four patients) as adjuvant treatment.

Variables measured

Time to recurrence [24] and survival were computed from the date of first treatment. Cause of death was extracted from patient records or, if necessary, was obtained from the General Practitioner.

Disease stage was determined using the TNM classification: the size of the primary tumour was taken as the maximum diameter measured in two dimensions; lymph node status was determined clinically, as not all patients had full axillary clearance as part of their treatment.

Postmenopausal patients were defined as those with 'absence of periods for the last 6 months without other known cause'. Patients who had undergone a hysterectomy before the natural menopause were coded as premenopausal if aged <50 years at the time of breast cancer diagnosis (nine patients), and postmenopausal if aged >50 years (14 patients).

Other variables recorded in this study were: length of symptoms, age at menarche, age at menopause, age at presentation, number of babies, age at first pregnancy and family history of breast cancer.

Follow-up

All patients were followed-up at 3 monthly intervals for 2 years after the beginning of treatment, then 6 monthly for a further 3 years, and annually thereafter. All analysis has been performed with respect to a common follow-up date of 31 December 1984.

Statistical analysis

The aim was to identify prognostic factors for survival time and the duration of disease-free intervals before local and distant disease recurrence. Survival time was investigated in the first instance by taking death due to any cause as the end-point, and then separately using only deaths attributable to the recurrence or persistence of cancer (deaths due to other causes then being treated as censoring events). Classification of deaths due to causes other than those directly attributable to breast cancer was determined from case-note information and death certificate information where possible. The number of such deaths so identified (64 patients) in the whole group of 607 patients compares well with the expected number (60.3) in a normal population of the same size with the same age distribution.

While survival time was defined for all patients, time to recurrence, whether local or distant, was determined only for those patients who underwent modified radical mastectomy.

All univariate analyses in this study used the log-rank test as the basis for comparing survival curves. Multivariate analysis for prognostic factors employed the Cox regression model for survival data [25]. In this method each candidate factor was assessed individually for its association with the length of the interval under consideration, and the one most strongly related was entered into the model. The remaining factors were then examined having allowed for the first factor, and the one which improved the prediction most was selected

and added to the model. The process was continued in this way (a 'step-up' approach) until all factors not yet entered into the model were found to make no statistically significant improvement in explaining the data.

Associated with each factor (Z_i) included in the final model is a number, the regression coefficient (B_i). The sign of this coefficient indicates the direction of prognostic effect: a positive coefficient corresponds to worsening prognosis as the value of the factor increases, and vice versa. For a given patient, if the value of each covariate in the final model is multiplied by its corresponding regression coefficient, and these products are added, a score is produced ($S = B_iZ_i + B_nZ_n$). In this study, these prognostic scores have been used to define subgroups with different prognoses. The scoring systems have been tested using another data set.

RESULTS

A univariate analysis of the effect of treatment in the four categories defined earlier revealed a significant trend of decreasing survival in the direction of radical to palliative therapy (Fig. 1). The four treatments were ordered and arbitrarily weighted using the integers 1–4 for subsequent multivariate analysis. Menstrual status was analysed univariately in three categories: postmenopausal, late postmenopausal and perimenopausal; 'late post' and 'peri' being defined as those patients presenting 'more than' or 'within' 5 years of their menopause, respectively. A significant trend (log-rank) in overall survival (Fig. 2) and recurrence-free survival was found. For further analysis menstrual status was coded with weight 1, 2 or 3 in these categories.

All patients—survival

Regression models. When treatment was included in the Cox regression, the factors found to explain best the variation in the observed survival data of the whole group of patients were stage, treatment and age [see Table 1(a)]. In this phase of the analysis deaths due to all causes were included. The significant factors are listed in the table in the order in which they entered the regression model, and the corresponding regression coefficients specifying the final model are given.

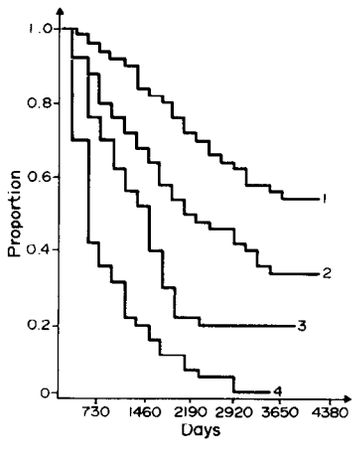

Fig. 1. Overall survival according to treatment modality: 1 = *mastectomy* 1 = *non-mastectomy*, 3 = *radiotherapy*, 4 = *palliative.*

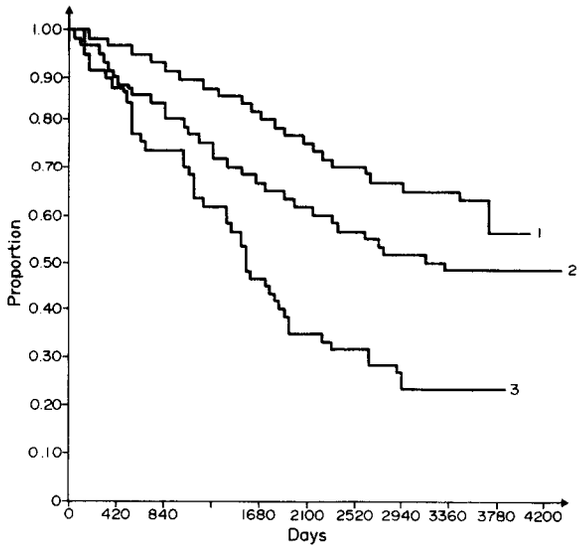

Fig. 2. Overall survival according to menstrual status: 1 = *premenopausal*, 2 = *late postmenopausal*, 3 = *perimenopausal.*

When deaths not attributable directly to breast cancer were treated as censored observations [Table 1(b)], age no longer significantly contributed to prognosis, and was replaced in the model by menstrual status (with perimenopausal patients doing worst) and tumour size. There was also evidence that 'age at menarche' carries prognostic information and this factor would have entered the model (with a negative regression coefficient) had the inclusion criteria been $P < 0.075$ rather than $P < 0.05$.

Table 1. *Survival (all causes)*

i	Factor (Z_i)	Regression coefficient (B_i)	[Standard error]
1	Stage	0.59	[0.09]
2	Treatment	0.34	[0.07]
3	Age	0.02	[0.004]

Table 1(*b*). *Survival (deaths due to breast cancer)*

i	Factor (Z_i)	Regression coefficient (B_i)	[Standard error]
1	Stage	0.46	[0.13]
2	Treatment	0.42	[0.08]
3	Menstrual status	0.41	[0.15]
4	Size	0.26	[0.11]

The inclusion of the treatment variable in these analyses is useful only by way of explaining the observed data of the whole group. This variable, however, reflects not only a 'true' treatment effect (if any) on survival, but also subsumes the effects of patient management policies which determine how a patient with particular characteristics will be treated. It is therefore not possible to disaggregate the prognostic effects of such characteristics from those of the treatment itself. If interest is in eventually using a regression model predictively to help decide treatment assignment, it is necessary to examine prognostic factors within groups of similarly treated patients. In this study, sufficient patients (326) were treated by modified radical mastectomy, and this group was analysed separately.

Patients treated by mastectomy

Regression models. Three hundred and twenty-six patients were treated by modified radical mastectomy, and of these, 100 had died of disease-related causes by the analysis date. From Table 2(a) it can be seen that stage was again significantly associated with survival (treating deaths due to other causes as censored). Age at menarche was the next variable to enter the model, with earlier menarche carrying a worse prognosis. Menstrual status was also found to be significantly important (perimenopausal patients faring worse). Finally, age was entered to improve the model further, with older patients having better prognosis after allowing for the other factors already included.

Time from first treatment to distant disease recurrence was also investigated and the same prognostic factors identified for survival were again found to be important [see Table 2(b)], since distant recurrence is a major determinant of survival. Stage (which itself was highly correlated with size) was the only variable found to be significantly associated with local recurrence-free interval.

Development of prognostic scores

Using the models for survival and distant recurrence-free interval, two prognostic scoring indexes were constructed in the way described earlier, and in each case used to define three subgroups of patients with different prognoses. Figures 3 and 4 show the discrimination provided by these scoring systems. The curves I–III in each figure represent patients with the highest 10%, the middle 80%, and the lowest 10% of scores, respectively. The ranges of scores corresponding to each patient subgroup are shown in Table 3. The *P*-values relate to significance of trend.

Table 2(*a*). *Survival (deaths due to breast cancer)*

i	Factor (Z_i)	Regression coefficient (B_i)	[Standard error]
1	Stage	0.87	[0.17]
2	Age at menarche	−0.13	[0.06]
3	Menstrual status	1.07	[0.32]
4	Age	−0.03	[0.01]

Table 2(*b*). *Distant recurrence-free interval*

i	Factor (Z_i)	Regression coefficient (B_i)	[Standard error]
1	Stage	0.98	[0.16]
2	Age at menarche	−0.13	[0.05]
3	Menstrual status	1.12	[0.30]
4	Age	−0.03	[0.01]

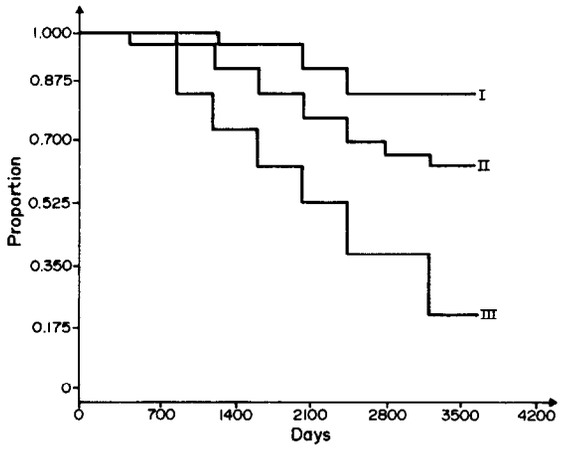

Fig. 3. *Prognostic groups—survival (deaths due to breast cancer): Group I = highest* 10% *of scores, Group II = middle* 80% *of scores, Group III = lowest* 100% *of scores.*

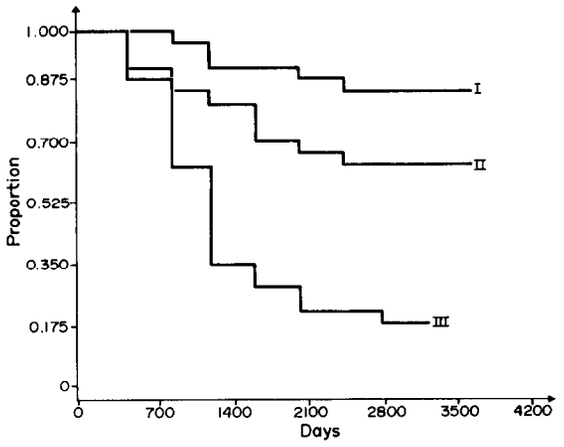

Fig. 4. *Prognostic groups—distant recurrence-free interval.*

Table 3.

Prognostic group	Overall survival	Distant disease recurrence-free interval
I	$S < -1.00$	$S < 0.80$
II	$-1.00 < S < 1.40$	$-0.80 < S < 1.60$
III	$S > 1.40$	$S > 1.60$

Validation of prognostic scoring systems

For scoring systems such as those described above to be used prospectively to identify patients who may fare poorly (or very well), they must first be tested on data sets other than that from which they were derived. Data on a further 457 patients treated by modified radical mastectomy at the same hospital were therefore gathered. The median follow-up (6 years) of this group was similar to that (7 years) of patients in the original sample. For each of these new patients, values of the four prognostic variables were entered into the scoring formula for overall survival. According to her score, each patient was classified into one of three prognostic groups defined by the intervals shown in Table 3. Figure 5 shows the survival of the three groups constructed in this way. Figure 6 shows the results of repeating the procedure using the distant recurrence-free index. In both cases the scoring systems have provided the desired discrimination.

DISCUSSION

The survival experience of a group of 607 breast cancer patients has been described with reference to clinical factors available at presentation, and the treatment they received. The analysis was multivariate, allowing at each stage the independent contribution of each factor in explaining the variability of the data to be assessed, after allowing for any significant factors already identified. Stage and treatment were both found to be strongly associated with survival when all patients were considered. Analysing all causes of death, age was the only other factor which significantly improved the model. The older a patient (within a given disease stage and under a given form of treatment) the worse the prognosis. This would be expected, since older patients have shorter remaining life expectancy and are more prone to concomitant disease. When deaths are restricted to those directly attributable to breast cancer (all other deaths being treated as censored), this age-effect disappears and menstrual status and size of tumour enter the model. In line with the findings of others [10, 11] the perimenopausal patients (within a given stage and treatment group) fare worse than pre- and late postmenopausal patients. Although size is taken into account when assessing stage, it has been found in this study to contribute additional independent prognostic information to that provided by stage alone.

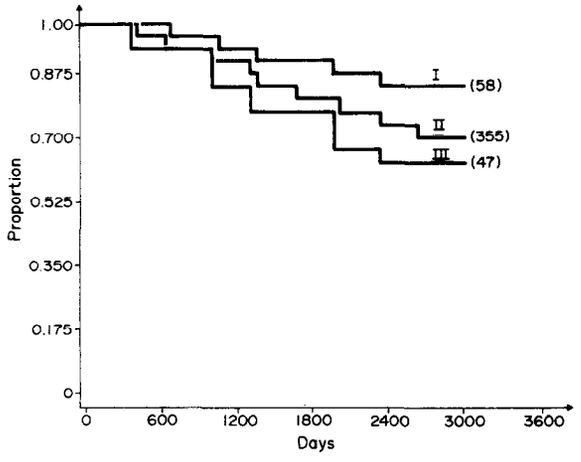

Fig. 5. Validation of prognostic index—survival.

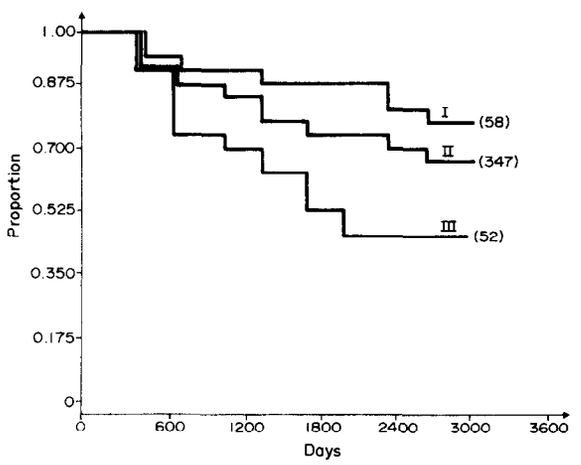

Fig. 6. Validation of prognostic index—distant recurrence-free interval.

While it was felt important to include the treatment variable in a descriptive analysis of survival experience, it is informative to consider the prognostic effect of other variables among similarly treated patients so that those who benefit most or least from a given form of treatment can be identified. The analysis was therefore repeated for the largest group of like-treated patients—the modified radical mastectomy group. As for the whole series of patients, stage and menstrual status were identified as being significant predictors of survival (disease-related deaths only), but in addition the risk factor age at menarche, which was previously borderline, was now found to be highly significant, early menarche being associated with poorer survival. This finding is not without precedent, Juret *et al.* [19] reported a similarly strong association. The hypothesis which he suggested, that this might reflect a correlation with the extent of nodal involvement, cannot however be answered using the data available in this study. An association (cohort effect) between age and age at menarche has also been conjectured [26], but the strength of this association in our data would not account for the level of prognostic significance found here. In fact, age itself also entered the model, with an independent effect, but after stage, age at menarche and menstrual status, with a tendency for younger patients to fare worse. Thus when considering disease-related mortality, these results suggest that older patients have no survival disadvantage simply by virtue of their age.

The same four covariates which were identified for survival were also found to be prognostic of length of distant disease-free interval for the radical mastectomy group.

Prognostic scoring systems based on these factors were constucted and validated (on a different data set) for both survival and distant recurrence-free interval. These indexes can be used prospectively to identify a subset of patients with early disease who fare relatively badly after modified radical mastectomy, and who may benefit most from the combination of local and systemic treatment.

The results of this study include the important finding that two hormonally related variables (*viz* menstrual status and age at menarche) are strongly associated with survival (and recurrence-free interval) of breast cancer patients. It should be noted, however, that because the emphasis here was on constructing a clinically useful prognostic classification, other variables such as tumour grade, nodal involvement and oestrogen receptor status, with which these co-variates may correlate, have not been considered.

Acknowledgements—The authors wish to thank Mrs. Diane Allen of the Imperial Cancer Research Fund, for her help in assembling and checking the data for this study.

REFERENCES

1. Haagensen CD. The treatment and results in cancer of the breast at the Presbyterian Hospital, New York. *Am J Roentgenol* 1949, **62**, 328–334.
2. Gordon-Taylor G. Cancer of the breast. *Br Med J* 1938, **2**, 1071–1073.
3. Zippin C, Kohn HI. An evaluation of the proposed international clinical staging system for cancer of the breast. *J Natl Cancer Inst* 1960, **25**, 13–24.
4. International Union Against Cancer. Committee on Clinical Stage Classification and Applied Statistics. The TNM System. General Rules. 1964.
5. Cutler SJ, Black MM, Mork T, Harvei S, Freeman C. Further observations on prognostic factors in cancer of the female breast. *Cancer* 1969, **24**, 653–667.
6. Williams MR, Hinton CP, Todd JH, Morgan DAL, Elston CW, Blamey RW. The prediction of local or regional recurrence after simple mastectomy for operable breast cancer. *Br J Surg* 1985, **72**, 721–723.
7. Boyd NF, Meakin JW, Hayward JL, Brown TC. Clinical estimation of the growth rate of breast cancer. *Cancer* 1981, **48**, 1037–1042.
8. Fisher ER, Redmond C, Fisher B. A perspective concerning the relation of duration of symptoms to treatment failure in patients with breast cancer. *Cancer* 1977, **40**, 3160–3167.
9. MacMahon B, List ND, Eisenberg H. Relationship of survival of breast cancer patients to parity and menopausal status. In: Forrest AP, Kunkler PB, eds. *Prognostic Factors in Breast Cancer*. Edinburgh, Livingstone, 1968, 56–64.
10. Langlands AO, Kerr GR. Prognosis in breast cancer: the effect of age and menstrual status. *Clin Oncol* 1979, **5**, 123–133.
11. Cutler SJ, Asire AJ, Taylor SG III. An evaluation of ovarian status as a prognostic factor in disseminated cancer of the breast. *Cancer* 1970, **26**, 938–943.
12. Hyman B, Myers MH, Schottenfeld D. The relationship of menstrual status and other risk factors to recurrence of carcinoma of the breast. *Am J Epidemiol* 1972, **96**, 173–182.
13. Peters MV. The role of local excision and radiation in early breast cancer. In: Stoll BA, ed. *Breast Cancer Management, Early and Late*. Chicago, Year Book Medical Publisher, 1977, 171.
14. Adami HO, Malker B, Holmberg L, Persson I, Stone B. The relation between survival and age at diagnosis in breast cancer. *N Engl J Med* 1986, **315**, 559–563.
15. Kleinfeld G, Haagensen CD, Cooley E. Age and menstrual status as prognostic factors in carcinoma of the breast. *Ann Surg* 1963, **157**, 600–605.
16. Falkson G, Gelman RS, Pretorius JF. Age as a prognostic factor in recurrent breast cancer. *J Clin Oncol* 1986, **4**, 663–671.
17. Host H, Lund E. Age as a prognostic factor in breast cancer. *Cancer* 1986, **57**, 2217–2221.
18. Wang DY, Rubens RD, Allen DS *et al.* Influence of reproductive history on age at diagnosis of breast cancer and prognosis. *Int J Cancer* 1985, **36**, 427–432.
19. Juret P, Couette JE, Mandard AM *et al.* Age at menarche as a prognostic factor in human breast cancer. *Eur J Cancer* 1976, **12**, 701–704.
20. Fisher ER. Prognostic and therapeutic significance of pathological features of breast cancer. *NCI Monogr* 1986, **1**, 29–33.
21. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer. A study of 1409 cases which 259 have been followed for 15 years. *Br J Cancer* 1957, **11**, 359–377.
22. Clark GM, McGuire WL, Hubay CA, Pearson OH, Marshall JS. Progesterone receptors as a prognostic factor in Stage II breast cancer. *N Engl J Med* 1983, **309**, 1343–1347.
23. Adami HO, Graffman S, Lindgren A, Sallstrom J. Prognostic implication of oestrogen receptor content in breast cancer. *Breast Cancer Res Treat* 1985, **5**, 293–300.
24. Hayward JL, Meakin JW, Stewart HJ. Assessment of response and recurrence in breast cancer. *Semin Oncol* 1978, **5**, 445–449.
25. Cox DR. Regression models and life tables. *J Stat Soc* 1972, Sec B **34**, 187–220.
26. Tanner JA. Trend towards earlier menarche in London, Oslo, Copenhagen, The Netherlands and Hungary. *Nature* 1973, **243**, 95–97.