Eur J Cancer Clin Oncol, Vol. 19, No. 12, pp. 1693–1697, 1983.
Printed in Great Britain.

0277-5379/83$3.00+0.00
Pergamon Press Ltd.

Prognostic Factors and Tumour Markers in Early Breast Cancer; a Commentary

R. D. BULBROOK

Department of Clinical Endocrinology, Imperial Cancer Research Fund, Lincoln's Inn Fields, London, U.K.

INTRODUCTION

THIS article is not a review of the literature on prognostic factors and tumour markers in early breast cancer but is a commentary based on some fifteen years' work on the periphery of the field. The advantage of a commentary over the usual classical review of the literature is that it is obvious from the start that the author has a particular axe to grind and is not impartial. Even formal reviews can be slanted in a number of ways, depending on the interests and bias of the reviewer. For example, the basic biological implications of some of the research in this area is not without interest: other workers may be interested in the taxonomy of the subject; the orderly listing of whether particular factors are tumour- or host-specific or whether a marker might be classified as a hormone or not. There is room for all those different viewpoints in cancer research. The main practical aim of work on predictive factors and markers is to see whether they have any immediate use in improving the efficiency of the treatments now available. In realistic terms, it would be desirable to be able to determine one simple figure: the probability that a patient will develop metastatic disease at a given rate (or, less satisfactorily, within a given time, and there is a distinction between these two options). The probability needs to be defined with considerable accuracy, and a finding which scrapes home with the minimal acceptable statistical significance is hardly likely to be useful if the predicted recurrence rate is going to be used for deciding a particular course of treatment. In other words, a probability of 0.05 that, say, serum IgE levels are higher in patients with rapid recurrence than in those with slower rates, based on a series of 20 patients followed for 2 yr, may be acceptable as a basis for a publication but it might be less convincing as a determinant of whether a patient should be given aggressive chemotherapy at the time of mastectomy.

In practical terms, it has to be admitted that work in the last decade has failed to provide realistic prognostic indices. It has not been established, beyond reasonable doubt, that action taken as a result of measurements of either single or multiple markers in a medical centre of excellence have added one day to the disease-free interval after mastectomy, or to survival, compared to that found in the generality of patients treated in peripheral hospitals where such facilities are not available. (Pathological staging and histological grading are excluded: hormone receptor site measurements are reviewed elsewhere in this volume.) It might now be useful to look into some of the reasons for failure.

HISTORICAL DEVELOPMENT

With the exception of staging and grading, most work on prognostic factors is of recent origin. Until quite recently, the primary treatment of breast cancer was so fossilized that there was little or no need for prognostic indices. If a patient fulfilled a few rudimentary criteria of operability, she was treated by the classical radical mastectomy described by Halstead in 1894 [1]. As Meyer points out [2], the title of Halstead's paper included the phrase "operations for the *cure* of cancer of the breast", which meant that the primacy of the radical operation was unchallenged for nearly 50 yr. Pathological staging and, later, histological grading, if measured at all, were carried out as an academic exercise.

The surgical position started to change in 1948 with the publication of McWhirter's study [3] of simple mastectomy combined with intensive irradiation, but there was still no pressing necessity for predicting the probable clinical course of the disease.

The position was different for metastatic disease. Oophorectomy had been introduced by Beatson in 1896 and still remains the major first-

line treatment for pre-menopausal women with metastases. Additive therapy with testosterone was introduced in 1939 and with oestrogens in 1944, with a remission rate from these treatments of about 30%, and with side-effects that were gradually tolerable there was little pressure to search for methods of selecting potentially responsive patients. But when adrenalectomy was used in 1952 and hypophysectomy in 1953 it was widely felt that the morbidity and mortality from these treatments were such that selection for major ablative surgery would be highly desirable [4]. The limited research resources that were available were concentrated on this problem and little work was done on the primary operable disease. This work culminated in the development of the oestrogen receptor assay, which provided a reasonably accurate method for predicting response. However, ablative surgery has been almost entirely superseded by combination chemotherapy, and the demise of major surgical treatments was inevitably accompanied by the virtual cessation of research on prediction (although the side-effects of chemotherapy are hardly negligible). This is a timely warning that the discovery of one good new treatment may lead to a similar disappearance of the present thriving prognostication industry.

While work on the advanced disease was at its height, research on prognosis in the early disease was desultory. In our own case, involvement in two of the earliest trials of adjuvant endocrine therapy [5], using testosterone in one of them and radiation menopause with or without prednisolone in the other, led us inevitably to an investigation of possible prognostic indicators. But the major stimulus to the field came with the introduction of adjuvant chemotherapy. Prolongation of the disease-free interval was reported by Fisher *et al.* [6], using L-Pam, and shortly afterwards Bonadonna *et al.* [7] produced similar results using CMF. Nodal involvement was already being used to select patients for adjuvant therapy, but the morbidity associated with these treatments and the possibility of unknown long-term side-effects made it imperative to look for other prognostic indicators.

Note that only 7 yr have elapsed since the adjuvant results were published. This is a relatively short time in the context of work on prediction, which is obviously long-term. One of the reasons, therefore, for the failure so far to develop useful marker assays lies in the historical development of the field. As a concrete example, given young patients with stage 1 disease, some 30–40% will recur within 10 yr and it would be advantageous to identify these high-risk women since their disease will lead to a substantial life-deficit [8]. This cannot be done within the usual span of a 3-yr research grant.

STAGE AND GRADE

Reference has already been made to the use of pathological staging for the selection of patients for adjuvant chemotherapy. Given a complete axillary clearance and a diligent search for nodes, the actual number of involved nodes provides a remarkable index of recurrence rates. In our own series from Guy's Hospital, London, patients with no involved nodes recur at approximately 3%/yr. The rate for patients with one positive node is 7%/yr and for each additional node there is a 2–3% increment. There appears to be no artificial cut-off point (e.g. <4 nodes vs $\geqslant 4$ nodes). Histological grading, in the hands of a single pathologist, is also a highly effective predictor. Little work has been done on combinations of stage and grade, but these two factors alone identify women with very little risk of recurrence and women with such a high-risk that the term 'early breast cancer' is inappropriate. The problem is that the bulk of the patients fall between these extremes so that another prognostic method is required. However, there is little point in ignoring excellent pathological indices of risk of recurrence when searching for new markers. This immediately brings us to the next problem—that large numbers of patients are required.

NUMBERS

In addition to stage and grade, there is the strong possibility that the usefulness of a prognostic indicator under investigation will vary with other factors such as age or menopausal status. This means that large numbers of patients should be investigated, but it is not uncommon to find descriptions in the literature of series of under 10 cases. There may be a magic marker waiting to be discovered of such power that a small series would be satisfactory, but with the current markers it is probable that 500 or more patients may be required if it is seriously intended to predict the clinical course of the disease. Numbers of this order cause two problems. The first is that few surgeons see enough patients, which means that multi-centre trials are necessary, and these often end up as unsatisfactory compromises. The second problem is that it takes time to accumulate a big series of patients and conditions do not remain static. Our own experience has been that during the course of a collaborative experiment involving 4 promising prognostic factors, based on 700 patients, a trial of adjuvant L-Pam was introduced, followed by a

trial of CMF and finally a trial of breast conservation.

TIME

If patients are standardized as far as possible into homogenous groups, then recurrence rates are approximately log-linear for the first 10 yr after mastectomy. This means that the probability of recurrence does not alter with time. If 10% of a particular amount of patients recur within 1 yr of mastectomy, then 10% of the disease-free survivors will recur at, say, 10 yr after mastectomy. It is a widely held belief that if a patient survives without recurrence for 5 yr, there is a good chance that she has been cured. This belief is based on the fact that the clinician sees few new recurrences in 5- or 10-yr survivors, for the obvious reason that the population has already been decimated.

Many of the published series have very short follow-up periods: a mean of 12 or 18 months is not uncommon. It is hard to see how much reliance can be placed on such data.

BIAS

The most usual and serious cause of bias is retrospective selection of a disproportionate number of recurrent cases for investigation. If serial patients are taken, about half will be in pathological stage 1, with a recurrence rate of 3%/yr. Even in stage 2 patients a rate of 20%/yr is not reached unless there is massive nodal involvement (>10 nodes positive). If the analytical technique for measuring the predictor is complicated, it is tempting to measure the prognostic factor only in recurred patients and a small sample of non-recurrences. Under these circumstances the apparent recurrence rate for patients with an unfavourable status may be 30% or more per yr. If the predictive factor is then measured in a prospective series where all patients are necessarily included, then the real rate of recurrence associated with a particular marker status may be only 5 or 10%. While the retrospective rate might persuade a clinician to opt for aggressive therapy, the prospective rate may not do so.

LACK OF TRANSFERABILITY OF METHODS: NON-REPRODUCIBILITY

As a rule of thumb, if an analytical method cannot be set up during a period of 3 months of intensive work, then the chances are that the fault lies with the originator of the method. One unnoticed and uncontrolled variable may be enough to prevent the establishment of a new method. Not only may analytical methods be difficult to transfer from laboratory to laboratory, but there is sometimes considerable 'within laboratory' variation. Our own experience has been that if blind samples are sent to laboratories specialising in measuring particular markers, then they are often unable to repeat their original findings. The errors may be large enough to be classified as comedy rather than science. For example, the results for a duplicate prolactin assay (the duplicate being coded and separated in time of assay) were reported as 0.0 and 76.0 ng/ml. In the first instance the interpretation would be a successful hypophysectomy; in the second, either a pituitary tumour or pregnancy. Not unnaturally, there is a marked reluctance to publish such findings. A lack of international standards adds to the confusion.

CHOICE OF MARKER

Our own experience comprises work on the indices listed in Table 1. These papers exemplify all the failings listed above, but we would claim that an experiment now in progress, and described in the following section, indicates that we may have learned from our mistakes. In Table 2 a series of possible markers, taken at random from the literature, is listed. All that needs to be said here is that the use of any of these predictors is not mandatory. There is still no convincing evidence, based on a substantial body of data from different laboratories, that any of the factors listed consistently show a marked relationship with recurrence rates.

The situation was summed up by Terence some 200 yr before the birth of Christ. He wrote "Quot homines tot sententia; suus cuique mos", and obviously had the marker field in mind.

In addition to the markers listed in Table 1, work is still in progress on the relation between serum levels of free and protein-bound oestrogens, free dehydroepiandrosterone, Δ5-androstene-3β-17β-diol, progesterone, prolactin and recurrence rates.

SERIAL ASSAYS AND LEAD-TIME

Some authors have argued that changes in marker levels may be more informative than

Table 1. *Prognostic factors measured*

Casein [9]	Hydroxyproline [14]
CEA [9]	Immunoglobulins [15]
Ferritin [10]	FSH [16]
Tryptophan metabolism [11]	LH [16]
Polyamines	T_3 T_4 TFI [17]
Sialyl transferase [12]	
Urinary androgens [13]	

Ref. [9] to casein: no correlation with recurrence was found in this series. Polyamines: no reference given since methodology not validated.

Table 2. Predictive factors/tumour markers described in the literature

α-FP	Lactic dehydrogenase
CEA	Sebum
Casein	Tryptophan metabolites
Lactalbumin	Polyamines
Lactoferrin	Nucleosides
GCDFP	GP 47
Ferritin	GP 52
HCG	P58
Calcitonin	Anti-P53
PTH	Reaction to PPD
Prolactin	Reaction to varidase
TSH	Reaction to mumps
Fucose	Reaction to candida
Sialic acid	Reaction to DNCB
Haptoglobin	T cell
OH-Proline	B cell
Prostaglandins	PHA
Alkaline phosphatase	Con A
Albumin	IgA
Pre-albumin	IgG
α1-Antitrypsin	IgM
α2-Ceruloplasmin	IgE
α1-Transferrin	Immune complexes
Orosomucoid	Clonogenic assay
C-Reactive protein	Urinary androgens
PAM	Sialyl transferase
γ-Glutamyl transpeptidase	Organ culture
6-Phosphogluconate dehydrogenase	Thymidine labelling index
α-Glycerolphosphate dehydrogenase	
Phosphofructokinase	

single assays at the time of diagnosis and that serial assays are required for monitoring such changes. There can be no doubt that individual patients do sometimes show a dramatic rise in titre before metastases became apparent, but these are the exceptions rather than the rule. Secondly, the lead-time so obtained is generally given as 3–6 months. There is no evidence that the institution of therapy for recurrence at this time will have any effect on outcome. Indeed, it can be predicted that this would be unlikely since adjuvant chemotherapy immediately after mastectomy is largely ineffective (reviewed elsewhere in this volume).

RETROSPECTIVE AND PROSPECTIVE STUDIES

It has to be admitted that this commentary on the present state of the art of the predictive factor and tumour marker field is profoundly pessimistic. We have attempted to resolve some of the difficulties by engaging in a retrospective/prospective study with the Breast Unit at Guy's Hospital, London. In the retrospective phase, blood, urine and tumour samples were collected from some 700 serial patients for whom detailed pathological results was available. Four markers which, on the face of it, appeared to have at least some promise were then measured retrospectively on sub-sets of the population. These were tumour enzymes [18], oestrogen and progesterone receptor sites [19], serum sialyl transferase [12] and urinary steroid metabolites [13]. When a relationship had been established between these factors and recurrence rates, the experiment was switched from a retrospective to a prospective study. In this study the design was such that it was mandatory to send the data to Guy's Hospital within 30 days of mastectomy and all clinical details were withheld from the investigators. So far, a further 500 serial patients have been accrued. The results of the retrospective study, however impressive, are still subject to many doubts: the results of the prospective study are the acid test because they will provide an objective measure of how efficiently the system would work in the real world of the management of patients with breast cancer. It is a comforting thought that if these measurements do not provide sufficiently accurate information on recurrence rates, new work using monoclonal antibodies (e.g. see [20, 21]) may solve what has been an intractable problem.

REFERENCES

1. HALSTEAD WS. The results of operations for the cure of cancer of the breast performed at the John Hopkins Hospital from June 1889 to January 1894. *Ann Surg* 1894, **20**, 497–503.
2. MEYER KK. The 'cure' of cancer of the breast. In: SEGALOFF A, MEYER KK, DEBAKEY S, eds. *Current Concepts in Breast Cancer.* Baltimore, MD, Williams & Williams, 1967, 3–9.
3. MCWHIRTER R. The treatment of cancer of the breast. *Proc R Soc Med* 1948, **41**, 122–136.
4. ATKINS H, BULBROOK RD, FALCONER MA *et al.* Urinary steroids in the prediction of response to adrenalectomy or hypophysectomy. *Lancet* 14 December 1968, 1263–1264.
5. MEAKIN JW, ALLT WEC, BEALE FA *et al.* A preliminary report of two studies of adjuvant treatment of primary breast cancer. In: FORREST APM, KUNKLER PB, eds. *Prognostic Factors in Breast Cancer.* Edinburgh, Churchill Livingstone, 1968, 157–163.
6. FISHER BF, CARBONE P, ECONOMOUS S *et al.* L-Phenylalanine mustard (L-Pam) in the management of primary breast cancer: a report of early findings. *N Engl J Med* 1975, **292**, 117–122.
7. BONADONNA G, BRUSAMOLINO E, VALAGUSSA P *et al.* Combination chemotherapy as an adjuvant treatment in operable breast cancer. *N Engl J Med* 1976, **294**, 405–410.
8. ZELEN M. A hypothesis for the natural time history of breast cancer. *Cancer Res* 1968, **28**, 207–216.
9. WANG DY, BULBROOK RD, HAYWARD JL *et al.* Relationship between plasma carcinoembryonic antigen and prognosis in women with breast cancer. *Eur J Cancer* 1975, **11**, 615–618.
10. JACOBS A, JONES B, RICKETTS C *et al.* Serum ferritin concentration in early breast cancer. *Br J Cancer* 1976, **34**, 286–290.
11. BELL ED, TONG D, MAINWARING WIP *et al.* Tryptophan metabolism and recurrence rates after mastectomy in patients with breast cancer. *Clin Chim Acta* 1972, **42**, 445–447.
12. STEWART J, RUBENS RD, MILLIS RR *et al.* Post-operative serum sialyl transferase levels and prognosis in breast cancer. *Br Cancer Res Tr* 1983, **3**, 225–230.
13. THOMAS BS, BULBROOK RD, HAYWARD JL *et al.* Urinary androgen metabolites and recurrence rates in early breast cancer. *Eur J Cancer Clin Oncol* 1982, **18**, 447–451.
14. GRANT CS, HOARE SA, MILLIS RR *et al.* Urinary hydroxyproline and prognosis in human breast cancer. *Br J Surg* In press.
15. WANG DY, GOODWIN PR, BULBROOK RD *et al.* Plasma immunoglobulin levels in patients with breast cancer. *Cancer* 1977, **39**, 2190–2193.
16. WANG DY, GOODWIN PR, BULBROOK RD *et al.* Plasma FSH and LH in post-menopausal women with breast cancer. *Eur J Cancer* 1976, **12**, 305–311.
17. THOMAS BS, BULBROOK RD, RUSSELL MJ, HAYWARD JL, MILLIS R. Thyroid function in early breast cancer. *Eur J Cancer Clin Oncol* 1983, **19**, 1213–1219.
18. DESHPANDE N, MITCHELL I, MILLIS RR. Tumour enzymes and prognosis in human breast cancer. *Eur J Cancer* 1981, **17**, 443–448.
19. KING RJB. Analysis of estradiol and progesterone receptors in early and advanced breast tumours. *Cancer* 1980, **46**, 2818–2821.
20. TAYLOR-PAPADIMITRIOU J, PETERSON JA, BURCHELL A *et al.* Monoclonal antibodies to epithelium-specific components of the human milk fat globule membrane: production and reaction with cells in culture. *Int J Cancer* 1981, **28**, 17–21.
21. ARKLIE J, TAYLOR-PAPADIMITRIOU J, BODMER W *et al.* Differentiation antigens expressed by epithelial cells in the lactating breast are also detectable in breast cancer. *Int J Cancer* 1981, **28**, 23–29.