Eur J Cancer Clin Oncol, Vol. 19, No. 12, pp. 1711–1716, 1983.
Printed in Great Britain.

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

Screening and Diagnosis in Breast Cancer; Workshop Report

A. B. MILLER* and E. A. VAN SLOOTEN†

NCIC Epidemiology Unit, University of Toronto, Canada and †The Netherlands Cancer Institute, Amsterdam, The Netherlands

INTRODUCTION

THIS workshop was held in two separate sessions, the first devoted to screening and the second to diagnosis.

SCREENING

The underlying general question in this session was "Is it timely to advise population-wide breast cancer screening in the Western world?" By the end of the workshop, it was clear that most participants' answer to this question was "Not yet", but a few indicated their belief that the time for screening was now.

Miller pointed out that we probably know more about screening for breast cancer than for any other cancer. This is primarily because of the one major controlled trial, that of HIP in New York, and a series of demonstration projects that have been conducted, some on a population basis. The findings from these studies have been encouraging and some believe that one should screen appropriately eligible women now and continue to evaluate the outcome in the light of changing population mortality figures. Others point to the continuing need to evaluate the effectiveness of breast cancer screening in population groups. Simultaneously the opportunity should be taken to evaluate the age group of women to be screened, the type of screening tests to be used, the frequency of rescreening and the appropriate means of persuading women to attend for screening.

The first major question addressed was "How reliable are the now-available data in terms of lives (potentially) saved?" Chamberlain described the data required for evaluating screening. A simple count of yield of cases is not sufficient. Obtaining the stage distribution of screen-detected cases is one step further, but there are the problems of potential over-diagnosis and the different staging criteria applied to screen-detected and clinically diagnosed cases. A better approach is to obtain oounts of the absolute numbers of advanoed oases in the total population to which screening has been offered as an indication of possible deaths from breast cancer. Survival suffers from the biases of lead time and length bias, and the only truly valid measure is the assessment of mortality in the total population in which screening has been offered compared to a comparable control population.

The only data currently available to approach an answer to the question are those from the HIP study. Fourteen years after the initiation of the study there was still a significant difference in numbers of deaths between the group offered screening and the control group. This is seen at all ages at entry to the study, though the differential in women under the age of 50 is restricted to those diagnosed over the age of 50.

In the U.K. a largely geographically based trial is underway, though in one of the two screening towns, Edinburgh, there is an inbuilt comparison between screened and control women randomized on the basis of general practices. At 1 yr the rate of detected breast cancer is highest in the two soreening towns, slightly above expectation in the two breast self-examination towns but at the expected level in the four control towns. The rate of advanced cases is essentially the same in all areas. Hendriks presented findings from the Nijmegen study which commenced in 1975, using single-view mammography. Follow-up has now continued for 4 yr, and survival data are expected in 1985. A substantial change in the stage distribution has been found in those offered screening, 45% being stage I and 55% stage II or more, compared to 29 and 71% respectively prior to the initiation of screening.

Burns commented on a 12-yr follow-up of women who had had breast cancer. In this group

there are 6–7 new cancers in the other breast per 1000 per yr. Substantial numbers of those found on mammography are minimal breast cancers, for which a clinical series observed elsewhere suggests a 90% survival. There is a problem in persuading surgeons and others to take appropriate action on diagnosis of an impalpable breast cancer. She believes the time is now for offering screening to women age 40 or more and would recommend annual mammography and physical examination.

Rombach reported on the Utrecht screening project, in which four rounds of screening have now been completed in women aged 50 or more. The size of tumours and the extent of detected metastases has been compared in various subgroups. The refusers and non-attenders have similar distributions as the non-screening participants. The distribution is substantially different, however, in those participating in the first and second screening rounds, with smaller tumours and less metastases.

Tabar presented the current findings from the population-based randomized control trial of single-view mammography being conducted in Kopparberg County since 1977. At 5 yr there is now an indication of reduced absolute numbers of stage II or more cancers in the screened group compared to the control. This is the first indication of the predicted mortality differential.

Frischbier described the mammographic screening programme that commenced in Hamburg in 1971. Screening is offered to asymptomatic women aged 30 yr or more at 2-yr intervals. More than 13,000 women have now been screened, with 122 cancers identified. Their experience indicates that one can reduce the size of tumours and the extent of nodal metastases, and Frischbier was sure that mortality reduction would follow, but time had not progressed long enough to ascertain this.

Miller expressed some surprise that it was not possible to identify reduction in mortality as a result of a screening programme that commenced in 1971. However, he pointed out that there is a possibility that we may all have been mislead by the apparent early mortality differential in the HIP study.

The next question addressed was "Is it possible to predict the compliance of the initially screened population in a lifelong program?" Hendriks indicated that in Nijmegen 76% of those eligible attended the first screening, 66% the second, 56% the third and 52% the fourth. Women of all ages over 40 were invited, but the compliance of those over the age of 70 was less than 25%. Chamberlain indicated that in the two British towns just over 70% attended invitations for screening and approximately 83% of those originally screened reattended. However, attendance for breast self-examination instruction in the other towns was not as good: 50% in one and 30% in the other. Attendance at clinics was better than class attendance. In constrast, Tabar reported that in Sweden 92.5% attended the first screening, 89.1% the second and among younger women 90.2% the third.

A discussion of this issue indicated that there may well be substantial inter-country differences; in Holland, for example, whatever the screening programme offered about 70% attend, but in Sweden a 90% attendance, especially in rural areas, can be anticipated. In Scotland, on the other hand, 67% attend. Attitudes vary in the women who do not come, but many are frightened of breast cancer and others may not regard breast cancer as a major threat to life.

In contrast, Miller reported that in the Canadian breast screening trial only approximately 25% of invited women in Ontario attend. The difficulty here is partly that women are being asked to participate in an individually randomized trial, not invited to attend a specific screening manoeuver. Also, that doctors and gynaecologists are sometimes not in favour of the project, partly because of lack of understanding but also because of some concern over losing hold of their patients. There is also much greater concern over mammography hazards in North America than in Europe.

The next question addressed was "What side-effects can be expected from a large-scale screening programme?" Hendriks indicated the possibility of over-diagnosis in that the incidence of breast cancer in Nijmegen has continued at a higher level than that anticipated from prior rates. He also commented on the need to have a good predictive value of recommendations for biopsy, which in general run at about 45%.

Honig spoke of psychological factors she had identified in Utrecht. As soon as an abnormality is found, the extra attention that women receive immediately results in anxiety, which tends to affect families. Nevertheless, once a diagnosis of a benign abnormality is made, there is no residual concern about the project and women are grateful to have participated. She felt it was important that women are given a great deal of information about the project; for example, if a mammographically indentified abnormality is noted, the mammograms are shown to the women and the potential significance of the abnormality discussed with them. Miller noted the much lower predictive value of recommendations for biopsy in the North American experience which had been noted in the BCDDPs and in the Canadian

Breast Screening Trial. This points to a different philosophy: it is felt important to avoid missing early detectable breast cancer, the general feeling being that it is impossible to distinguish many benign abnormalities from early malignant disease, both on mammography and physical examination, in the absence of a biopsy, which in many instances can be a needle aspiration biopsy. Dr. Chamberlain pointed to the need to assess the predictive value of recommendations arising from soreening tests in the light of the experience in the control group. In their hands a predictive value of recommendations in the control group is of the order of 30%.

Lundgren pointed out that the numbers of breast operations eventually go down in a population with a well-established screening programme. In the long run, as experience increases, the quality of the screening recommendations improves and treatment for breast abnormalities becomes cheaper. In Sweden as a whole now, the breast service is cheaper than before. Hendriks has noted the same changes in Nijmegen, as the use of mammography has introduced more precision in diagnosis.

Habbema discussed the means to assess the implications of these issues in cost–benefit analyses. The favourable effects of screening can be assessed by life years gained, which he felt was better than the numbers of lives saved as mortality is inevitable. Increased quality of life following screening should be assessed if at all possible. There may be other favourable effects of breast cancer screening, e.g. improvements in employment, financial savings from better medical care. The unfavourable effects would include unnecessary treatment and the important consideration of radiation-induced breast cancers, though these can be minimized by quality control of mammography equipment. There are the effects of false-positive tests in terms of anxiety, as well as unnecessary investigations and surgery and the need to compute the investment of women's time and money, with the costs of screening and manpower requirements added in. In making evaluations it was necessary to note that although the HIP study may not seem relevant to current conditions, we will always be faced with this problem, and when the present generation of breast screening trials reach the 15-yr point, once again circumstances may have changed.

The final question addressed in this session was "How can we decide whether it is feasible to start screening in a given population, and if so, what would be the most efficient programme?" Miller pointed out that we need to decide whether we can extrapolate findings from one country to another. Different questions are being addressed in different countries under different socio-political circumstances. In Canada physical examination of the breasts is regarded as an important component of the National Breast Screening Study. The cancers found on physical examination which are invisible on mammography are different from those found on mammography alone. They are often small infiltrating lesions, some with one or two involved nodes but clearly much earlier stage II cancers than would normally be diagnosed with a potentially smaller tumour body burden and possible greater response to adjuvant chemotherapy. He was optimistic that it was going to be possible with the varying approaches being adopted in different programmes to make rational decisions on appropriate policies in the future.

Habbema discussed the means whereby screening strategies could be determined. It was going to be necessary to decide the age range to be screened, the screening test to use and whether one could concentrate on high-risk groups. It was likely that with increasing investment in screening, the marginal additional benefits would diminish and there would be a point where an optimal screening strategy should be adopted. The age range for screening was likely to be dependent on the incidence in different ages of the detectable pre-clinical cancers. The interval between screens was dependent on the distribution of the duration of the screen-detectable stages and on the sensitivity of the screening tests. The magnitude of the favourable effects is influenced by the attendance for screening examinations and the improvement of real prognosis brought about by early detection. Current evidence indicates that modern mammography promises more benefit than the HIP study or, possibly, the same benefit for less screening and therefore less costs. His group are developing a mathematical model that incorporates these features. As an example, using their Miscan analysis, it has been determined that on the basis of the HIP data a benefit from screening is derivable if the life years gained by screening are weighted in terms of 100 examinations; this would involve 11 screens in a lifetime over an age range of 40–70 yr at intervals of 3 yr. However, if the weight is increased to 200 examinations per life year gained, 25 screens in a life are possible over the age range 35–73 yr at intervals of 1.6 yr. The same programme applied to the presumed benefits of modern mammography suggests that it will be possible to screen less frequently for the same benefits.

In the general discussion two issues were raised. The first related to the extent to which there are background changes occurring in the population in incidence and mortality from breast cancer that

may make the potential advantages of breast cancer screening more difficult to assess. In Sweden and possibly the Netherlands, there is increasing incidence but reducing mortality from breast cancer. In Canada there is increasing incidence with stable mortality, interpreted by Miller as suggesting that incidence is rising because of the inclusion of less biologically aggressive breast cancers. Roberts said that in the U.K. there seemed to be little change in locally advanced tumours in national cancer registration, remaining at about the 33% level.

The second question raised was the value of breast self-examination. Tabar suggested that the promotion of breast self-examination by information on the back of a letter informing women that they are free of disease may have helped in the early diagnosis of interval cancers. The use of BSE in the Canadian trial may also be having the same effect. In Trieste an educational programme involving BSE commenced in 1979. There has been a substantial increase in the proportion of stage I breast cancers. Miller commented that Cole suggested that BSE may well be of value under circumstances where other screening tests are not available but may not have a great deal to contribute when more effective screening is being offered to women.

Possibly the most encouraging aspect of this part of the workshop was the multitude of questions that are being addressed in the various screening projects now ongoing. Although some might have wished for more uniformity, the diversity of the questions that are being addressed in varying circumstances on both sides of the Atlantic will eventually enable a future workshop to reach some very definite conclusions in relation to the still unanswered questions which the participants attempted to address.

DIAGNOSIS

The major question addressed in this session of the workshop was the need to improve the initial diagnosis of distant spread before treatment for breast cancer to more precisely define the potentially clinically curable patient. Professor van Slooten pointed out that the bad signs of breast cancer indicating incurability on clinical grounds were well recognized: the presence of inflammation, involved nodes, satellite nodules to the primary tumour and fixation. However, it was still necessary to improve our ability to identify those patients who will develop progressive metastatic breast cancer and thus reduce the extent to which they need to be hospitalized, treated surgically and receive the potential complications of surgery such as mutilation and swollen arms. Our aim should be to make treatment as simple as possible for incurable patients so that we can provide appropriate palliation. To do this it is necessary to make better assessment of lymphatic and haematogenic spread.

There were two broad topics discussed in this session of the workshop. The first related to radiologic approaches to diagnosis of malignant disease and the second to modern immunologic approaches in combination with cytology or radiology for the detection of metastatic disease.

Butzelaar introduced the topic of bone scintigraphy. He pointed out that 85% of patients who die have bone involvement. As a device to detect bone involvement, appreciation of the value of bone scans has swung like a pendulum. Up to about 1977, approximately 26% of the reported papers indicated positive bone scans. However, this applied to many groups where complete staging was not always applied. After 1977, in stage I and II patients, of the order of 1.8–4% of patients only were regarded as being bone-scan-positive. Thus, in general, bone scintigraphy has been rejected as a routine preoperative procedure for patients definitively staged I or II. Doubt has remained, however, on the value of post-operative bone scans as part of routine follow-up. His own view was that they should only be done in patients where there is a definite indication.

Perez, from the Royal Marsden Hospital, described a study of 1116 patients who have had primary treatment for breast cancer since 1976 to evaluate the use of diagnostic tests for bone metastases. Of these, 586 had an abnormal bone scintigram either at primary diagnosis or on subsequent follow-up and 167 developed radiographic bone metastases. Their policy in the last 5 yr has been to do a bone scan on every patient and to X-ray the areas where a hot spot was noted. It was found that only 12% of patients with one or more hot spots identified had radiographic bone metastases. Although the proportion increased with the numbers of hot spots, only 30% of those with three or more hot spots were found to have radiographic bone metastases.

Conversely, of those with radiograph-positive findings, 70% had three or more hot spots. Most sites of bone metastases had approximately 60% with hot spots, but this increased to 84% for the thoracic spine and reduced to 35% of those with metastases in the femur.

The conclusion from this study is that a scan is not particularly good for prediction, diagnosis or detection of radiographically detectable metastases and that X-rays are necessary. However, from an economic point of view it is possible to have a limited X-ray survey, approximately 83% of

radiographic bone metastases being detected on pelvic and lateral thoracic spine views, a proportion which increases to 92% by adding routine skull and chest films. This is the limited survey currently being utilized.

In the discussion Hoefnagel pointed out that there were three different patterns of bone scintigraphy predictive of bone metastases: multiple hot spots; cold lesions, which are often difficult to diagnose; and an increased tracer uptake indicative of extensive metastases that could only be assessed in comparison to other examinations or from the absent kidney sign. Fractures resulted in false-positive lesions; therefore it was essential to have the history of the patient to interpret findings. Perez pointed out that diagnostic confirmation is best done using CT scans and did not recommend a drill biopsy. Burns pointed out the differences in readers in the extent to which reports of scans could be interpretable. A bone scan was usually positive in patients who had symptoms. Perez suggested that in a patient with no symptoms it was not necessary to perform either an X-ray or a scan. There may, however, be some value in using bone scans in a small category of patients with stage III disease. Burns suggested that on occasions baseline examinations in patients with stage II as well as stage III disease was useful.

Hilgers of the Netherlands Cancer Institute gave a general introduction to the topic of monoclonal antibodies. Three groups in Europe were actively evaluating this approach as an aid to the diagnosis of metastatic breast cancer: the ICRF and Ludwig groups in the U.K. and the group in the Netherlands Cancer Institute. Monoclonals have been raised in mice to a number of cell surface membranes of human origin, especially in relation to breast cancer, derived from the milk fat globule membrane. There were separate groups of antigens of different specificity, many of which were not found in tumours. The M3 group of antigens have advantages in that they are not destroyed by formalin fixation and are found in some tumours with a potential value as prognostic indicators. However, the minority of tumours are positive. They are found in many epithelia in the body and a fraction of large bone marrow cells are positive.

The M6 group of antigens have produced three monoclonals, and it is this group which is being used as the major focus for evaluation of localization, though tending to go under different names. They are present in the majority of tumours and form one of the three major proteins in the milk fat globule membrane. They are normally on the outside of the body, but in tumours come in contact with the blood stream and therefore can be used for tumour detection. In many normal tissues they are found on the luminal surface, but in tumours tend to appear away from any surface structure as well.

Hilgers pointed out that the original hope that monoclonals would identify tumour-specific antigens has not yet been realized. The original claims that they were tumour-specific have proven much too optimistic. Although the solid tumour work has tended to lag behind the work on tumours of the lymphoid system, there is now increasing interest in the possibility that the reagents derived can be used to localize tumour metastases *in vivo* by radiolabelling the antibody or can be used to measure the tumour load by determining in a quantitative fashion the amount of antigen in serum. For the future there is hope for therapeutic applications of monoclonal antibodies, once it has been proven possible to attach relevant drugs to them.

Coombes, on behalf of the Ludwig–ICRF–Royal Marsden group, discussed a new approach to detecting micrometastases in bone marrow using an immunocytochemical stain for an epithelial membrane antigen (EMA). This investigation developed out of the recognition that 70% of patients at relapse had positive bone scans, and it was felt important to look at the bone marrow in patients in advance of this situation to attempt to detect micrometastases. A polyclonal antibody was raised in rabbits against human milk fat globule membrane which, though not specific, nevertheless does not result in normal bone marrow staining. Bone marrow aspirates from eight sites in the pelvis, sacrum and lumbar spine are obtained under general anaesthesia as part of routine surgical procedures in women with breast cancer. After appropriate preparation these are stained for EMA. The same approach could be adopted with an appropriate monoclonal antibody.

It was first found that clumps of cancer cells were present in several patients with advanced disease. Subsequently, single cells were found in early cases using the EMA technique. In comparison with other approaches in bone marrow, there was an increased pick-up rate of metastatic cells, with none of those identified by previous approaches missed. In a small group of patients with negative bone scans but at high risk for bone metastases (primarily patients with soft tissue or other metastases) 19 were found to have positive EMA cells. On follow-up there was a far greater probability of the development of bone metastases in those with positive cells than those found negative.

In July 1981 an evaluation of this approach commenced in unselected patients with primary

breast cancer. Currently 110 patients have been admitted to the investigation, of whom 59 are node-positive and 84 oestrogen-receptor-positive. All have been fully screened for metastatic disease and staged. All have had their marrows examined by both conventional cytology as well as the EMA technique. Twenty-eight percent have been found to be EMA-positive, five of these patients having more than 100 carcinoma cells identified and 20 less than 50. There was some correlation between the detection of micro-metastases and the size of the tumour and vascular invasion.

So far 36 patients without breast cancer with normal marrow have been evaluated blind and no apparent carcinoma cells have been found.

Coombes indicated that there will have to be a continued accrual of patients into this study (with the collaboration of the Guy's group, they hope to achieve 400 within a reasonable time period) and long-term follow-up to find out if the test is of predictive value for bone metastases. The test presumably has implications for other cancers. It was his opinion that at present it was undesirable that there should be an extensive use of this technique until its predictive value had been assessed in the current investigation.

Rainsbury, on behalf of the Royal Marsden–Ludwig group, discussed a study of the effect of radioactive label on the localization of human breast carcinoma by a monoclonal antibody. Four monoclonal antibodies have been raised to human milk fat globule membrane. Three were IGM, one IGG-1. Only the latter was found likely to be useful in tumour localization. This they labelled M-8. It is found on the luminal surface of normal cells, with expression increased in human breast carcinoma. After labelling with isotopes of iodine, they found a heterogeneic distribution of antibodies in the tumour as well as specificity of localization in the tumour but not in normal tissue. They were able to obtain some localization with use of the gamma camera. A succession of different radioisotopes were evaluated. After ^{131}I, ^{123}I labelling was evaluated. This, however, was unstable *in vivo*, with 40% of the isotope excreted in 48 hr. A limited number of patients with metastatic breast cancer were studied and in two of five, metastases were demonstrated. ^{125}I labelling resulted in increased specificity, with a tumour to normal breast ratio of 10:1. However, labelling with ^{111} In increased the immunoreactivity, with increased tumour localization in animals. Ten patients with metastatic breast cancer were studied, in eight with known bone metastases, localization was demonstrated. This was specific in the tumour, in contrast to non-specific changes in bone scanning. Currently, therefore, it is felt that the In-labelled antibody has superiority over other approaches and further studies are continuing.

In summing up this session of the workshop, van Slooten felt that real advances are likely in the diagnosis of distant metastases in patients deemed potentially curable in the not too far distant future. It is, however, not yet clear whether micro-metastases in the bone marrow of patients will prove to be a predictor of bone metastases, metastases elsewhere or even just a reflection of propensity for spread with the possibility that under certain circumstances body defense mechanisms or modern adjuvant chemotherapy will make what is currently still a very complicated technique unnecessary. The potential use of radiolabelled monoclonal antibodies has so far only been demonstrated in patients with advanced disease. Their utility will need to be demonstrated in patients with early disease for real application and both techniques will require prolonged patient follow-up in order that they can be fully assessed. Nevertheless, we could be on the threshold of a major improvement in our ability to stage patients and thus achieve the objectives of this session of the workshop.