PERGAMON

European Journal of Cancer 38 (2002) 1458–1464

European Journal of Cancer

www.ejconline.com

The frequency of breast cancer screening: results from the UKCCCR Randomised Trial

The Breast Screening Frequency Trial Group*,[1]

Received 23 October 2001; accepted 23 October 2001

Abstract

The optimal frequency of breast cancer screening has been a subject of debate since the inception of the UK National Breast Screening Programme (NHSBSP). This paper reports on the only randomised trial directly comparing different screening intervals. 99 389 women aged 50–62 years who had been invited to a prevalent screen were randomly allocated after the scheduled prevalent screen date to the study arm (invited to three further annual screens), or to the control arm (invited to the standard single screen 3 years later). 37 530 women in the study arm and 38 492 in the control arm had attended the prevalent screen. The endpoint was predicted breast cancer deaths. The prediction was based on both the Nottingham Prognostic Index (NPI) and a similar method derived from survival data from a series of tumours in the Swedish Two-County screening trial (2CS). Both indices were based on the size, lymph node status and histological grade of the invasive tumours diagnosed in the two arms of the trial. The pathology of the cases diagnosed was subject to review by two pathologists using standard criteria. The tumours diagnosed in the study arm were significantly smaller than those diagnosed in the control arm ($P=0.05$). The relative risk of death from breast cancer for the annual compared with the 3-yearly group was 0.95 (95% Confidence Interval (CI): 0.83–1.07, $P=0.4$) using the NPI and 0.89 (95% CI: 0.77–1.03, $P=0.09$) using the 2CS. Shortening of the screening interval in this age group is predicted to have a relatively small effect on breast cancer mortality. Improvements to the screening programme would be targeted more productively on areas other than the screening interval, such as improving the screening quality. © 2002 Published by Elsevier Science Ltd.

Keywords: Mammography; Breast screening; Screening interval

1. Introduction

In 1988, population screening for breast cancer was introduced nationally in the UK for women aged 50–64 years in the National Health Service Breast Screening Programme (NHSBSP). The interval between screens was chosen to be 3 years. There are no other trials comparing screening intervals.

The United Kingdom Co-ordinating Committee on Cancer Research (UKCCCR) trial of Breast Screening Frequency was established to investigate whether more frequent screening is advantageous. The trial randomised women to invitation to annual mammographic screening or to the routine NHSBSP first incident screen at 3 years following invitation to a prevalent screen.

A trial comparing two different mammographic screening regimes has to overcome two essential obstacles. First, because both arms are receiving some form of screening, breast cancer deaths will be rare overall. Secondly, such deaths occur over the long term. Thus, a trial with breast cancer death as the end-point would need very large numbers of subjects and very long follow-up. Thus, although the ideal endpoint of such a trial would be mortality from breast cancer, this would involve waiting for at least a further 10 years after the closure of the trial and randomising approximately 200 000 women in order to have adequate power [1].

This trial was therefore, designed to use prognostic features of the tumours diagnosed to predict subsequent mortality from breast cancer [1]. The predicted mortality would be a valid outcome provided that the treatment under trial influenced the true mortality largely via its effect on the prognostic features (which is the case in a trial of diagnostic method) [2], and the prediction of the true mortality was accurate [3]. We therefore designed the trial using as an end-point the predicted

* Corresponding author: Professor R.W. Blamey, City Hospital, Nottingham NG5 1PB, UK. Tel.: +44-115-962-5707; fax: +44-115-962-7765.

[1] The group members are listed at end of paper.

0959-8049/02/$ - see front matter © 2002 Published by Elsevier Science Ltd.
PII: S0959-8049(01)00397-5

deaths from the tumour size, node status and histological grade of the tumours diagnosed. These were combined into the Nottingham Prognostic Index (NPI) [4], which has been shown to predict breast cancer survival in various tumour series, including one of approximately 10 000 cancers [5,6]. The tumour size, node status and grade have also been shown to predict the mortality reduction from screening in various randomised trials of screening [7]. The result based on predicted deaths is available earlier than that based on observed, and has greater statistical power, due to being based on a larger number of events [1].

This paper reports the main endpoint, the predicted survival in the two groups. A later publication will carry full details of the analysis by centre, of attendance and non-attendance at the prevalent screen, of subsequent attendance rates following the prevalence screen, recall rates and numbers of diagnostic operations. There will be further reports from the radiological subgroup on the analysis of 'missed' cancers, particularly at the prevalence screen, on the number of views at each visit and other matters, and from the pathology review group comparing features of the tumours detected in the two groups. A study of tumour biology and growth rates and an economic evaluation will also follow.

2. Patients and methods

Five screening units (Cardiff, Epping, Gateshead, Newcastle and Nottingham) participated in the trial, between 1989 and 1996. The primary end-point of the trial was the expected deaths from all invasive tumours diagnosed from immediately after the prevalence screen up to and including the final trial screen 3 years later. The primary aim of the trial was to establish, for women who had attended a prevalent screen and for whom breast cancer had not been diagnosed, the relative benefits of different intervals before re-invitation to the next screen. Therefore, for purposes of the primary analysis, only women who attended the prevalence screen and in whom no cancer was found at the prevalence screen were included. Prevalence screen detection rates in the centres ranged from 5.6 per 1000 to 8.5 per 1000, closely correlated with the pre-trial incidence in the areas served. Women with breast cancer diagnosed prior to the trial were excluded.

Women aged 50–62 years who had been invited to a prevalent screen in the NHS Breast Screening Programme were randomised immediately after the time of the appointment for the prevalent screen, in ignorance of attendance at the prevalent screen and of its result, to be invited to a conventional incident screen after an interval of 3 years (Control group, $n = 50\,216$) or to three annual screenings (Study group, $n = 49\,173$) (Fig. 1). Randomisation was by month of birth in the first 2 years of the trial, after which the national screening programme software was augmented to allow individual randomisation. Of those invited, 38 492 (77%) in the control group and 37 530 (76%) in the study group had attended the prevalent screen. The size of the trial was based on a requirement of 80% power to detect as significant at the 5% level in a two-tailed test, a reduction, in previous attenders to the prevalent screen, in predicted mortality of 25%, in the study arm compared with the control arm [1].

The trial was designed to use the prognostic attributes of the tumours diagnosed to estimate expected deaths in the Study and Control groups [1,4]. Dedicated data clerks in each unit regularly searched the pathology reports and databases of the local hospitals for cancer cases diagnosed at centres other than the screening and assessment services. This was complemented by searches of regional cancer registries. All cancers, however diagnosed (at screening or as interval cases), *in-situ* and invasive, in the Control and Study groups were recorded with their time of detection related to the time of the prevalent screen. The prognostic factors of the cancers were initially reported by the pathologist in each centre, followed by a pathology review of all available cases by two Consultant Breast Pathologists. Where biological material was not available for review (7% of cases), the unreviewed data from the original pathology laboratory were used. Tumour size on histology, tumour grade [8] and lymph node stage (both number and site of lymph

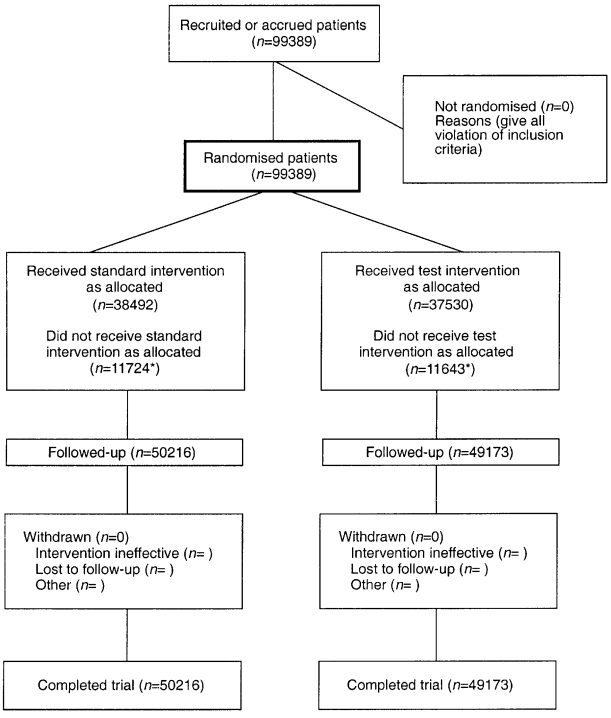

Fig. 1. Flow chart of the progress of patients through the trial (adapted from Ref. [13]). *Did not attend the prevalence screen so, as per protocol, not included in the primary comparison.

nodes involved) were recorded. Cases diagnosed as malignant phyllodes rather than breast carcinoma were excluded (one from each group).

The size, node status and histological grade of the invasive tumours were incorporated into two independent and previously validated prognostic indices, the Nottingham Prognostic Index (NPI) [4] and the Swedish Two-County Study Survival Index (2CS) [7,9]. The NPI is based on all primary operable breast cancers, however detected, is well validated in both intra-centre [5] and inter-centre [6] comparisons (the latter on over 9000 cases), and accurately predicts the percentage of patients who die from breast cancer in three defined prognostic groups (NPI <3.4, NPI between 3.4 and 5.4, NPI >5.4). These groups are associated with 10-year survival rates of 83, 52 and 17%, respectively. Knowledge of the number of Control and Study patients in each prognostic group was used in this study to predict the total numbers in the Study and Control groups dying by 10 years.

In 50 cases in the Study group and 42 in the Control, lymph node biopsy was not undertaken (50% of these were in one centre). For these, a prognostic index based only on size and grade was calculated in the same way as for the NPI, by Cox regression analysis of survival, and using the same dataset as was used for the development and confirmation of the NPI. This gave the index $(0.42\times\text{size in cm})+(0.78\times\text{grade})$. The corresponding three groups yielding the same survival figures as above for the revised index were 2.19 or less, 2.20–3.35 and 3.36 or more.

In addition, a prognostic index based on 2468 tumours (49% of which were screen-detected) diagnosed in a screening trial, the Swedish Two County Study (2CS) [9] was used to estimate the relative risks of death from breast cancer in the Study and Control Groups. This index has been validated on data from other screening trials and shown to be accurate in predicting the breast cancer mortality reductions observed in these trials [7]. If any of grade, node status or size were missing, the prognostic index was modified to estimate survival based only on the available attributes.

Table 1
Invasive cancers by arm and detection mode in prevalence attenders only

Detection mode	Study group (%)	Control group (%)
Interval year 1	22 (9)	19 (9)
First annual screen	46 (20)	0
Interval year 2	13 (6)	34 (16)
Second annual screen	50 (21)	0
Interval year 3	34 (14)	51 (25)
Third annual screen	70 (30)	104 (50)
Total screen-detected	166 (71)	104 (50)
Total interval cancers	69 (29)	104 (50)
Overall total	235	208

Both indices are based on invasive cancers. Thus, in calculation of predicted deaths it has been assumed that there would be no deaths from breast cancer in cases of ductal carcinoma *in situ* (DCIS).

Distributions of tumour attributes were compared between the two arms using Chi-squared tests. The relative risk of predicted death from breast cancer in the Study arm compared with that in the Control arm was calculated using the above prognostic indices. All statistical tests used were two-sided.

3. Results

In the study group, among those women who had attended the prevalence screen, overall attendance rates at the first, second and third annual screens were 78, 78 and 81%, respectively. In the Control group, attendance at the 3-year screen was 85% among those who had attended the prevalence screen.

A total of 643 cancers was diagnosed during the period of the trial. Of these, 535 (443 invasive) were diagnosed in women who attended the prevalent screen, a rate of 2.35 per thousand per annum. In the Study group, 287 cancers (235 invasive) were diagnosed, 2.55 per thousand per annum, and in the Control group, 248 (208 invasive), 2.15 per thousand per annum. These figures do not constitute a significant deviation from the numbers expected for equal incidence in each group; neither did the numbers of ductal carcinoma *in situ* cases differ significantly between the two groups.

Biological material was available for review for 402 (91%) of the 443 invasive cancers, 221 (94%) in the Study group and 181 (87%) in the Control group. For five tumours in the study group and four in the control, no surgery was performed, therefore no prognostic information was available. Review of the notes showed that four were locally advanced and these were ranked as poor prognostic group tumours, one woman was unfit for surgery and was excluded from analysis. Histological grade was not available for six tumours in the Study group and four in the Control. For two tumours in the Study group and five in the Control, histological size was not available. These were not used for the predicted mortality using the NPI, but were used in the prediction from the 2CS.

Table 1 shows the times of detection according to whether the cancers were detected at screening or diagnosed in the first, second or third annual intervals. This is expressed graphically as cumulative incidence in Fig. 2, from which the lead times gained by the Study group can be seen as the difference in time between the cumulative incidences of diagnosed cancers in the two groups. The median additional lead-time in the Study group over the Control is 7 months. Detection rates at the first, second and third annual incident screens in the

Table 2
Invasive cancers by size, node status and histological grade in the Study and Control arms of the trial

Tumour attribute	Study arm number (%)	Control arm number (%)
Size		
1–10 mm	59 (25)	38 (19)
11–20 mm	111 (48)	96 (47)
21–50 mm	59 (25)	64 (32)
50+ mm	4 (2)	5 (2)
Not known	2	5
Total	235	208
P for trend	0.05	
Nodes		
Negative	122 (66)	105 (63)
1–3 positive	43 (23)	46 (28)
4+ positive	20 (11)	15 (9)
Not known	50	42
Total	235	208
P for trend	0.5	
Grade		
1	39 (17)	37 (18)
2	94 (41)	81 (40)
3	96 (42)	86 (42)
Not known	6	4
Total	235	208
P for trend	0.9	

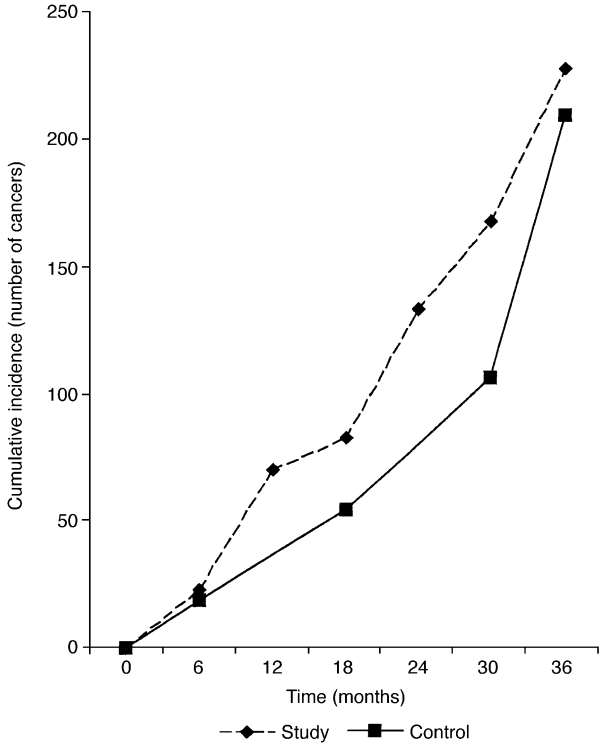

Fig. 2. Cumulative incidence of invasive breast cancer in study and control groups in women attending the prevalent screen.

Study group were 1.6, 1.7 and 2.3 per thousand screened, respectively. The rate at the 3-year incident screen in the Control group was 3.2 per thousand.

Table 2 shows the breakdown of the invasive cancers in each group by prognostic factors: histological size, lymph node status and grade. There was a small, but significant, difference ($P=0.05$) in size at diagnosis in favour of the Study group, but no significant differences were seen in lymph node status or grade.

There was a small, statistically insignificant advantage to the Study group in the proportion lying in the NPI Good Prognostic Group (Table 3). From the previously known outcomes in each Prognostic Group, and the numbers in this study observed to fall in each prognostic group, the numbers of women with cancer in the Study and Control groups predicted to die within 10 years were calculated. From this calculation (Table 3), the total percentages of deaths from breast cancer predicted in the Study and Control Groups were 36 and 38%, respectively—an absolute difference between the two groups of 2%. The relative risks of breast cancer death are shown in Table 4. Using the NPI, there was a relative risk of 0.95 (95% CI: 0.83–1.07, $P=0.4$). Estimation using the 2CS gives a relative risk of 0.89 (95% CI: 0.77–1.03, $P=0.09$).

4. Discussion

In this trial, annual screening conferred an increase over 3-yearly screening of 13% in invasive cancers detected and of 60% in screen-detected cancers, with an

Table 3
NPI categories and predicted 10-year survival by trial arm

NPI category	Predicted 10-year survival (%)	Study group		Control group	
		No. (%)	Predicted number surviving at 10 years	No. (%)	Predicted number surviving at 10 years
Good Prognostic Group	83	113 (49)	94	92 (46)	76
Moderate Prognostic Group	52	96 (42)	50	87 (43)	45
Poor Prognostic Group	13	20 (9)	3	22 (11)	3
Not known		6	–	7	–
Total		235	147 (64%)	208	124 (62%)

NPI, Nottingham Prognostic Index.

Table 4
Predicted survival from prognostic indices and relative risk calculations

Quantity estimated	Study	Control
10-year predicted survival using NPI	88%	87%
RR (95% CI) breast cancer death using NPI	0.95 (0.83–1.07)	1.00 (–)
15-year predicted survival using 2CS	74%	71%
RR (95% CI) breast cancer death using 2CS	0.89 (0.77–1.03)	1.00 (–)

RR, relative risk; 95% CI, 95% confidence interval; NPI, Nottingham Prognostic Index; 2CS, Swedish Two-County Study.

increased total cancer detection rate and a reduction in the size of the tumours, both of which were of borderline statistical significance. Despite these, there was no statistically significant further reduction of predicted mortality from breast cancer for annual as opposed to triennial screening. The small amelioration in prognostic factors from annual screening was not enough to substantially alter the predicted survival. Changing to a 2-year interval would confer an even smaller benefit.

The reasons for the small size of the benefit are also seen in this study. As observed, the median lead time in detection produced is only 7 months as against an average of between 2 and 3 years from a prevalent screen [9]. In addition, only those detected at the first and second annual screens in the study arm could have benefited from the 1-year interval. The results provide empirical confirmation of previous predictions of the effect of changing the interval, from modelling of disease incidence, progression and mortality [10,11]. With prevalence rates of around 3 times the annual incidence, it is clear that the major impact on lead time arises in the prevalent screen.

Whilst screening has its effect by diagnosing cancer at a curable time, lead time bias also artificially increases the survival and might be expected to do so to a greater extent in the Study group than the Control. This can be taken into account by estimating the predicted deaths to a given period after the prevalence screen rather than after diagnosis. Using the Swedish Index, we calculated the relative risk based on deaths up to 15 years after the prevalence screen. This gave a relative risk of 0.90, close to the unadjusted figure of 0.89.

There are a number of qualifications to the interpretation of the results. The results are based on predicted rather than observed deaths, however both prognostic indices used have been thoroughly validated [5–7]. Follow-up for actual mortality is in progress.

It could be argued that our interval estimates in Table 4 are artificially narrow, since they assume that survival probabilities from the NPI and 2CS categories are known without error. These probabilities are, however, based on large numbers of cases (around 10 000 in the case of the NPI and 2468 in the case of the 2CS), and Day and Duffy have shown that in these circumstances, the variation in the survival probabilities contributes only a small amount to the overall variance of the predicted mortality [1].

The predicted reductions in deaths from breast cancer are based on case fatality rather than mortality per person randomised, since there was no significant difference in incidence between the two arms of the trial. Analysis based on expected mortality per person randomised would yield an even smaller estimated benefit, with a larger confidence interval.

It is acknowledged that the results above relate to screening as practised in the early 1990s; since then there have been considerable changes in screening and assessment, notably increased use of two-view mammography at screening and greater use of percutaneous biopsy at assessment. These have led to increased sensitivity in detecting tumours at an earlier stage, so lead time in the study group might be further improved. However, it is difficult to see how such an improvement could be substantial. The greater sensitivity would apply to the prevalence screen and thus reduce the impact of the incidence screens.

It is assumed in the above analysis that no benefit is derived from the higher rate of DCIS in the study arm. There were 12 more cases in the Study arm than in the Control. A proportion of the corresponding occult ductal carcinoma *in situ* cases in the control arm may have progressed to invasive disease or may do so in the future, so one might speculate that there would be a further, albeit small, additional mortality benefit to that which we have estimated. Analyses were repeated with the DCIS cases included and assumed to give rise to no deaths, yielding estimated relative fatalities of 0.91 using the NPI and 0.88 using the 2CS. Thus, inclusion of the DCIS cases did not make a substantial change.

The NPI predicts a higher absolute death rate than the Swedish Index. This is because the survival probabilities associated with the three NPI categories were based on tumours diagnosed between 1973 and 1988, before the use of adjuvant systemic therapy. Later tumour series from Nottingham indicate that the survival within each NPI category has become higher, although the relative survival between prognostic groups remains approximately the same. Use of the more recent survival figures would predict less deaths in both groups and hence reduce the absolute benefit.

The trial was structured so that with equal attendance and perfect sensitivity, the screen at 3 years would equalise the incidence of breast cancer in the two groups. The higher incidence of cancers in the Study arm is not statistically significant, and is not considered to be a real difference. However, if the difference is real, it might be partly due to the increased diagnostic activity in the Study arm, in which the radiologist has potentially three opportunities to observe subtle mammographic features or changes in appearance whereas, in the Control arm, there is only one such opportunity. If the difference is taken as real, from the detection rates at the incidence screens in each arm, we estimate that the extra diagnostic activity was responsible for eight of the additional cancers in the Study arm.

The higher incidence in the Study arm might be interpreted in terms of subsequent breast cancer deaths in two ways. One might conclude that with the higher incidence there would be a larger number of breast cancer deaths in the Study arm. This, however, assumes that the higher incidence in the Study arm is real and systematic, whereas it is in fact compatible with chance variation. In our analysis, using survival rather than mortality, our estimate of relative benefit is the same as the estimate which would result if the difference in incidence were subsequently made up for by incidence in the Control group of cancers with the same distribution of prognostic features as was observed in the Control group during the trial period.

From the incidence rates (Fig. 2), the lead time gained by annual screening was 7 months. This is borne out by a simple approximation using the screen-detected and interval cancers diagnosed in the Study and Control groups, as shown on Table 1. The extra lead time in the Study arm can only arise from the tumours diagnosed at the first and second annual screens (20 and 21% of cancers in the Study group, respectively). Of those cancers occurring in the Control group after the first annual screen in the Study group, 34 (interval year 2) occurred approximately 6 months after the first annual screen, 51 on average 18 months later and 104 on average 2 years later. Thus, the average additional lead time for cancers diagnosed at the first annual screen in the study group is 19 months, i.e. $(34\times6+51\times18+104\times24)/(34+51+104)$. Of the cancers occurring in the control group after the second annual screen of the study group, 51 occur on average 6 months later and 104 on average 12 months later. Thus, the average additional lead time of tumours diagnosed at the second annual screen is $(51\times6+104\times12)/(51+104)=10$ months. Thus, the average additional lead time is approximately 6 months $(0.2\times19+0.21\times10)$, close to the median of 7 months estimated from Fig. 2. It is difficult to see how this could be substantially improved upon without unrealistic expectations of detection rates at the annual screens.

Although a final economic evaluation is not yet published, the national cost of screening yearly would add around £56 million to the breast screening budget in the first year with subsequent additional costs of 35–40 million pounds per annum [12]. Our results suggest that such expenditure would be expected to save few lives and that resources could be better allocated on areas other than the screening interval.

When the study began, the participating screening units were among those most experienced in screening in the UK, but with hindsight, some cancers missed at the prevalent screening have been recognised by the radiology subgroup. The radiology review will deal with this in detail. Since the prevalent screen confers such a large benefit, further resources might be better targeted on improving quality, rather than on reducing the present 3-year interval (e.g. two-view mammography, positioning, radiographic quality, double reading and radiological quality assurance), thus giving a larger benefit from screening every 3 years.

A lesson for trialists in population screening is that the decision by the applicants to restrict the trial to only five centres made population data retrieval and pathological verification possible. There were still logistic and informatic problems to solve, but these would have been considerably greater if there had been more centres.

In conclusion, the main result of the trial is that following a prevalent screen, shortening the screening interval from the current 3 years gave no statistically significant decrease in the predicted mortality from breast cancer.

Acknowledgements

The Medical Research Council, the Cancer Research Campaign, the Imperial Cancer Research Fund and the Department of Health funded this project. We thank the staff of the five screening centres, the surgeons, radiologists, pathologists, radiographers, technical and clerical staff involved in the diagnosis, histopathology and treatment of the cancer cases, and the former project managers at the Medical Research Council (MRC) Biostatistics Unit. We also thank the women who participated in the trial.

Breast Screening Frequency Trial Group

Original applicants: R.W. Blamey (Nottingham City Hospital), N.E. Day (University of Cambridge), J.R. Young (Queen Elizabeth Hospital, Gateshead).

Pathology review: I.O. Ellis, S.E. Pinder (Nottingham City Hospital).

Trial statistician: Stephen W. Duffy (Imperial Cancer Research Fund, London).

Centre principal investigators: Cardiff, Breast Test Wales: K. Lyons, I. Monypenny, E. Roberts.

Epping, St. Margaret's Hospital: M. Letcher, M. Morgan, R.M.L. Warren.

Gateshead, Queen Elizabeth Hospital: M.J. Egan, J.D. Hemming, J.R. Young.

Newcastle, General Hospital: L. McLean, M. McElroy, R.G. Wilson.

Nottingham, City Hospital: R.W. Blamey, C.W. Elston, A.R.M. Wilson.

Project Management at Molecular Research Council (MRC) Biostatistics Unit, Cambridge: N.M. Walker, I.M. Warsi.

Health Economics, Brunel University, Uxbridge: J. Brown, K. Johnston.

References

1. Day NE, Duffy SW. Trial design based on surrogate end points-application to comparison of different breast screening frequencies. *J R Stat Soc A* 1996, **159**, 49–60.
2. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med* 1989, **8**, 431–440.
3. Begg CB, Leung DHY. On the use of surrogate end points in randomized trials. *J R Stat Soc A* 2000, **163**, 15–28.
4. Blamey RW. The design and clinical use of the Nottingham Prognostic Index in breast cancer. *The Breast* 1996, **5**, 156–157.
5. Todd JH, Dowle C, Williams MR, *et al*. Confirmation of a prognostic index in primary breast cancer. *Br J Cancer* 1987, **56**, 489–492.
6. Balslev I, Axelsson CK, Zedeler K, Rasmussen BB, Carstensen B, Mouridsen HT. The Nottingham Prognostic Index applied to 9,149 patients from the studies of the Danish Breast Cancer Cooperative Group (DBCG). *Breast Cancer Res Treat* 1994, **32**, 281–290.
7. Organizing Committee and Collaborators. Falun meeting. Breast cancer screening with mammography in women aged 40–49 years. *Int J Cancer* 1996, **68**, 693–699.
8. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I: The value of histological grade in breast cancer: experience from a large study with long term follow up. *Histopathology* 1991, **19**, 403–410.
9. Tabar L, Fagerberg G, Chen HH, *et al*. Efficacy of breast cancer screening by age: new results from the Swedish two-county trial. *Cancer* 1995, **75**, 2507–2517.
10. Chen HH, Duffy SW, Tabar L, Day NE. Markov chain models for progression of breast cancer, part I: tumour attributes and the preclinical screen-detectable phase. *J Epidemiol Biostat* 1997, **2**, 25–35.
11. Boer R, de Koning H, Threlfall A, *et al*. Cost effectiveness of shortening screening interval or extending age range of NHS breast screening programme: computer simulation study. *Br Med J* 1998, **317**, 376–379.
12. Johnston K, Gerard K, Morton A, Brown J. *NHS Costs for the Breast Screening Frequency and Age Trials. Health Economics Research Group Discussion paper number 16*. Uxbridge, Brunel University, 1996.
13. Begg C, Cho M, Eastwood S, *et al*. Improving the quality of reporting of randomised controlled trials: the CONSORT statement. *JAMA* 1996, **276**, 637–639.