14 years of follow-up from the Edinbugh randomised trial of breast-cancer screening

Citation for published version:

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published in:
The Lancet

Publisher Rights Statement:
© 1999 Elsevier Ltd

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening


Summary

Background The Edinburgh randomised trial of breast-cancer screening recruited women aged 45–64 years from 1978 to 1981 (cohort 1), and those aged 45–49 years during 1982–85 (cohorts 2 and 3). Results based on 14 years of follow-up and 270 000 woman-years of observation are reported.

Methods Breast-cancer mortality rates in the intervention group (28 628 women offered screening) were compared with those in the control group (26 026) with adjustment for socioeconomic status (SES) of general medical practices. Rate ratios were derived by means of logistic regression for the total trial population and for women first offered screening while younger than 50 years. Analyses were by intention to treat.

Findings Initial unadjusted results showed a difference of just 13% in breast-cancer mortality rates between the intervention and control groups (156 deaths [5·18 per 10 000] vs 167 [6·04 per 10 000]; rate ratio 0·87 [95% CI 0·70–1·06]), but the results were influenced by differences in SES by trial group. After adjustment for SES, the rate ratio was 0·79 (95% CI 0·60–1·02). When deaths after diagnosis more than 3 years after the end of the study were censored the rate ratio became 0·71 (0·53–0·95). There was no evidence of heterogeneity by age at entry and no evidence that younger entrants had smaller or delayed benefit (rate ratio 0·70 [0·41–1·20]). No breast-cancer mortality benefit was observed for women whose breast cancers were diagnosed when they were younger than 50 years. Otherwise mortality rates did not differ by trial group when adjusted for SES.

Interpretation Our findings confirm results from randomised trials in Sweden and the USA that screening for breast cancer lowers breast-cancer mortality. Similar results are reported by the UK geographical comparison, UK Trial of Early Detection of Breast Cancer. The results for younger women suggest benefit from introduction of screening before 50 years of age.

Lancet 1999; 353: 1903–08
See Commentary page 1896

Introduction

The Edinburgh randomised trial of breast-cancer screening was started in 1978. Between 1978 and 1981, the trial recruited 44 288 women aged 45–64 years of age from 87 general practices in Edinburgh to form the initial cohort. Subsequently, further eligible women who became patients of these practices and existing patients who reached 45 years of age were recruited in two further cohorts, 4867 in 1982–83 (cohort 2) and 5499 in 1984–85 (cohort 3). Participating practices were randomly assigned to intervention and control groups, women taking their status from their general medical practice at their time of entry. Women in the intervention group were invited to participate in a screening programme, which included clinical examination every year and two-view mammography every 2 years. Control-group women received only normal medical care. The prospectively defined hypothesis was that breast-cancer mortality would be lower in the intervention group than in the control group after 7 years and longer periods of follow-up. At that time, there was no reason to believe that the effect might differ by age of entry to the trial and no subgroup analyses were planned. Subsequent evidence from other randomised controlled trials of mammographic screening have placed importance on two age-groups (<50 years, ≥50 years at entry).

There have been two reports of mortality from breast-cancer relating to experience of women in the initial cohort followed up for 7 and 10 years, respectively. The second report also included the later cohorts over a shorter follow-up period. For both these follow-up periods, we report that breast-cancer mortality was 17–18% in the screening group; these differences were not significant. The 10-year analysis reported that breast-cancer mortality was 22% lower in the screening group for younger women aged 45–59 years at entry.

As a result of the cluster randomisation, there was bias between the two groups, women in the control group having higher all-cause mortality rates and lower socioeconomic status (SES) than those randomly assigned to intervention. Attempts to adjust for these differences by quantifying SES in samples from each practice were unsuccessful. An improved method of quantifying SES has now been developed, which we believe adjusts for this bias. We include in this report the effect of this adjustment on breast-cancer mortality rates after 14 years and separately consider the effect of screening for younger women.

With longer periods of follow-up, the inclusion of women whose diagnosis could not have been influenced by screening becomes difficult. The consequences of...
different policies on censoring of deaths by their date of diagnosis were considered in the Swedish overview analysis and the 14-year report of the Health Insurance Plan (HIP) trial and are now considered for the first time for the Edinburgh randomised trial. The HIP investigators were the first to note that for entrants younger than 50 years to the intervention group benefit was restricted to women whose breast cancers were diagnosed when they were in their fifties. Indeed, breast-cancer mortality for women whose breast cancers were diagnosed when they were younger than 50 years was higher in the intervention than in the control group. A corresponding analysis is now done for the Edinburgh trial. Women in the intervention group of the Edinburgh trial formed the Edinburgh component of the UK Trial of Early Detection of Breast Cancer (TEDBC). The latest results of TEDBC are also reported in this issue.

Methods

Study participants and design

The geographical base for this trial was 87 general medical practices within the city of Edinburgh. The only practices not included were those that had participated in a pilot study or refused to take part. Every woman aged 45–64 years registered at one of these practices and without a previous diagnosis of breast cancer was eligible for entry to the trial. Practices were individually randomised after stratification by number of partners (to balance the numbers in the two groups) and entered sequentially between 1978 and 1981. All women aged 45–64 years from each practice were entered as the practice was recruited (with cluster randomisation of women) to form an initial cohort (cohort 1). During two 2-year periods (1982–83 and 1984–85), newly eligible women registered with participating practices were also recruited to form cohorts 2 and 3. We included only women aged 45–49 years at entry in these two cohorts. Older entrants to these cohorts were mainly women who had moved to Edinburgh and were not representative of the population. The only exclusion criterion applied at trial onset was a previous diagnosis of breast cancer, but this was normally ascertained after randomisation (figure 1). Informed consent was obtained for the screening process, but not for trial entry. The trial was approved by the Lothian Research Ethics Committee. The primary outcome measure was the rate ratio of breast-cancer mortality rates in the two groups; the trial was designed to have 80% statistical power to detect a 30% lower mortality rate in the intervention group after 7 years of follow-up (rate ratio <0.70), by a one-sided test. Power calculations ignored the cluster randomisation.

The study continued until 1988, with women in the intervention group offered a maximum of four mammographic screenings every 2 years for cohort 1, three screenings for cohort 2, and two screenings for cohort 3. During the study period these women were also offered clinical breast examinations once a year. With the start of service screening in Edinburgh in June, 1988, all trial women younger than 65 years were eligible to be screened. Those who had participated in regular screening during the trial were invited for their first service screening 3 years after their last trial screening. Controls were invited for their first service screening at the corresponding time: year 10 from entry for the initial cohort, and years 8 and 6 for cohorts 2 and 3, respectively.

The records of all women in the trial were flagged by the General Registry Office in Edinburgh in 1985. This standard UK procedure allows follow-up information on cancer incidence, breast-cancer mortality, and other-cause mortality to be sent to the trial administrators wherever individual participants are living. This report is restricted to the 98% of women whose records were successfully flagged. Failure of flagging is therefore a second exclusion criterion; most of the women whose records were not flagged had died or left Edinburgh before the trial began, although their records remained on their general practice list.

Data analysis

Mortality statistics were based on death certification accessed through the flagging process. Breast-cancer deaths are those in which breast cancer was mentioned on the death certificate as the underlying cause, according to WHO rules. All such deaths were checked against the trial database and, if necessary, additional sources. They were included only after confirmation that the date at which breast cancer was diagnosed was later than that of randomisation. All deaths in which breast cancer was not the underlying cause were recorded as other-cause mortality. The follow-up period is 14 years for cohort 1, 12 years for cohort 2, and 10 years for cohort 3, from date of randomisation. To counteract the effect of the inclusion of deaths from breast cancer that could not have been influenced by screening, results are presented not only for all deaths during follow-up, but also with death censored if diagnosis occurred after a specified date (1987, 1989, 1991) and if the diagnosis occurred after a specified number of years in the study. Additional analyses censored deaths after diagnoses before or after age 50 years to allow comparisons with results from the HIP trial.

Breast-cancer and other-cause mortality rates were calculated as rates per 10 000 woman-years at risk for the two groups in the trial, and the risk ratio calculated with a modified logistic regression procedure, as in previous studies, incorporating adjustment for extrabinomial variation by the method of Williams to respect the cluster randomisation. These analyses were done in SAS Proc Logistic version 6.12 and stratified by age at randomisation (45–49 years, 50–54 years, 55–59 years, 60–64 years). When the later cohorts were included further stratification by cohort was added. Age-specific results were derived from fitting of interaction terms for age with trial group, but all women were included in the modelling process. Cumulative mortality curves were expressed as rates per 10 000 woman-years at risk. Our method of analysis differs somewhat from that used for TEDBC; that trial used Poisson rather than logistic regression, but we used the TEDBC method to confirm that the differences in statistical methods do not explain the divergence between their results and ours. Analyses are by intention to treat.

The trial database, during the study phase, held and still holds the best (most accurate) address for each woman. These addresses did not include postal codes. When we noted in 1983 that there was an imbalance in all-cause mortality between the...
Table 1: **Trial population and woman-years of follow-up**

Two trial groups, addresses of 20% of women from each general practice were given postal codes manually and 1981 census data derived from these postal codes were used to allocate an SES score to each practice. From this score, practices were classified into three groups (high, medium, and low SES). This classification system (SES-1) has the strength that it was derived before any inspection of breast-cancer mortality but it did not succeed in eliminating the difference in other-cause mortality. Postal codes have now been allocated by computer to the addresses of almost all women in the trial, allowing retrieval of the Carstairs index of deprivation for the area of residence of each woman. The mean value of Carstairs index for general practices was used to derive a continuous variable (Carstairs score) for general practices. The Carstairs score was used to assign practices into three groups (system SES-2) as before.

All analyses were repeated with adjustment for each of SES-1 (as used previously), and SES-2 and Carstairs score in the logistic regression model. In these adjusted analyses, we fitted terms for SES using the total trial population, even to report age-specific results. This approach assumes that the effect of SES is independent of age entry to the trial. That assumption has been verified; the interaction between SES and age at entry did not approach conventional levels of significance.

**Results**

Table 1 shows the number of women and woman-years at risk in cohort 1 by age at entry and for those aged 45–49 years at entry in cohorts 2 and 3. Women in cohorts 2 and 3 were generally younger at age of entry than those in the 45–49-year age-group in cohort 1 (mean age cohort 1, 47.4 years; cohort 2, 46.1 years; cohort 3, 45.8 years). When the cohorts are combined, they allow a direct comparison of 22,746 women who were younger than 50 years at entry (11,479 intervention and 10,267 control).

Women in the intervention group were offered an initial screening and up to six further screenings (three mammographic) for the initial cohort, four (with two mammographic) for cohort 2, and two (one mammographic) for cohort 3. 61.3% of women in the intervention group accepted the first invitation to screening, but attendance rates fell with time and were just over 50% during the final year of the study. Attendance at screening for women in the intervention group was strongly associated with SES (table 2). Younger women (45–49 years at entry) began trial screening under age 50 years but many continued to receive trial screening after 50 years. Of all trial screenings for these younger women 46–0% for the initial cohort, 79.0% for cohort 2, and 97.5% for cohort 3 were done when they were under 50 years. Little screening was available in Edinburgh during the study period (1978–88) and we believe that very few women in the control group arranged screening for themselves, although we have no way of confirming this assumption.

In cohort 1 followed up for 14 years there were 323 deaths in which breast cancer was the underlying cause. Analysis without adjustment for SES or with the previous method of adjustment (SES-1) showed differences of 13–16% between the intervention groups, which were not significant (table 3). Analysis of the unadjusted data by the TEDBC method gave almost identical point estimates.

However, adjustment by SES-2, in three categories or as a continuous variable (Carstairs score), gave point estimates for differences in breast-cancer mortality rates of 21–22%, which are of borderline significance (p=0.055 and 0.05, respectively). There was no significant interaction between age-group at entry and trial group (p=0.6), but we show results by age of entry within cohort 1 in table 3. All 95% CI are wide, but point estimates for all groups, except those of 50–54 years, are similar to the overall results; the difference in breast-cancer mortality between intervention and control groups for the youngest women (45–49 years) in the initial cohort was 30%.

When the late entrants (cohorts 2 and 3) were included in the 45–49-year age-group, the point estimate of the difference in breast-cancer mortality rates was 25%. Although the 95% CI was narrower, this difference in mortality is not significant. Cumulative breast-cancer mortality rates for the initial cohort at all ages and for women in all three cohorts entered at ages 45–49 years by trial group are shown in figure 2.

### Table 2: Proportions of women attending for screening according to SES-2

<table>
<thead>
<tr>
<th>SES-2</th>
<th>Number of women from intervention group</th>
<th>Number of first-time attenders attending last screen</th>
<th>Number of woman-years of follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>5777 (79%)</td>
<td>78 761</td>
<td>5594 (70%)</td>
</tr>
<tr>
<td>Medium</td>
<td>5878 (78%)</td>
<td>78 838</td>
<td>5168 (71%)</td>
</tr>
<tr>
<td>Low</td>
<td>6109 (75%)</td>
<td>79 500</td>
<td>5749 (73%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60–64</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5162 (78%)</td>
<td>64 065</td>
<td>4831 (78%)</td>
</tr>
<tr>
<td></td>
<td>22 926</td>
<td>301 155</td>
<td>21 342 (78%)</td>
</tr>
</tbody>
</table>

*Intervention group versus control group.

### Table 3: Breast-cancer mortality during 14 years of follow-up
When general practices were grouped by SES-2 there was a strong trend in the control group of mortality rates in cohort 1 and a weaker trend in the intervention group towards higher breast-cancer rates in those of higher SES (table 4). The benefit of screening seems to be concentrated in women of higher SES, but neither the interactions with SES-2 nor those with Carstairs score were significant.

Other-cause mortality rates were significantly lower in the intervention group than in the control group (rate ratio 0·84 [95% CI 0·79–0·90]) but after adjustment for Carstairs score (the continuous variable) this difference was no longer apparent (0·98 [0·92–1·04]).

In general, the results were not affected by censoring by study year of diagnosis. When censoring was applied to cases diagnosed after study year 10, the observed benefit in the intervention group was 29% (rate ratio 0·71 [0·53–0·95] if deaths after diagnoses after study year 10 are excluded). Figure 2 shows breast-cancer mortality with this censoring applied. Results were similar for the other censoring imposed for diagnosis around 3–4 years after the end of the study (study years 9–12, calendar years 1990–92).

If deaths in the 45–49-year age-group were censored according to age at diagnosis, no benefit in the intervention group was observed when analyses were restricted to deaths of women whose breast cancers were diagnosed before they reached age 50 years (rate ratio 2·36 [0·79–7·08]); substantial benefit was seen when these deaths were excluded (0·47 [0·25–0·89]). The question of benefit from screening offered to women before they reach age 50 years (in addition to that from routine screening in their fifties) can be addressed, but with limited power, by comparison of breast-cancer mortality by trial group for younger women (entrants 45–46 years) in cohorts 2 and 3; for these women the rate ratios were 0·83 and 0·50, respectively (with wide 95% CI, 0·32–2·18 and 0·11–2·23).

Discussion

This report of the Edinburgh randomised trial of breast-cancer screening, with “before randomisation” consent and no offer of screening to the control group, includes nearly 70,000 woman-years of follow-up. Trials in Canada included a volunteer population all of whom, including the control group, received some form of intervention at entry to the trial or throughout the study. Because of recruitment of two later cohorts of women aged 45–49 years, the Edinburgh trial provides a substantial contribution to the total information available on the screening of younger women (more than 270,000 woman-years of follow-up). A limitation of this trial has been the imbalance between the intervention and control groups for other-cause mortality rates. New computer software has enabled us to access postal codes for all women, so we have been able to derive an improved estimate of general practice SES (SES-2). The variation in SES explains the difference in other-cause mortality rates and thus allows best estimates of effects on breast-cancer mortality.

Application of SES-2 gave an overall point estimate for the difference between intervention and control groups in breast-cancer mortality rates in the Edinburgh trial at 14 years of follow-up as 21%; this estimate is of borderline significance (rate ratio 0·79 [95% CI 0·60–1·02]). If SES-2 is applied to our previously reported results at 10 years of follow-up, the difference in breast-cancer mortality rates is 24%, compared with the 18% first reported. This difference approaches significance (rate ratio 0·76 [0·55–1·06]). These benefits do show some variation by age, with the largest benefits in women first screened in their late fifties and none in those first screened in their early fifties. Although this latter observation is consistent with results from some other trials, we emphasise that the formal test of heterogeneity by age-group did not approach significance.

Breast cancers diagnosed after a suitable period from the end of the study clearly cannot have been influenced by trial screening, and the inclusion of deaths after these later diagnoses dilutes the comparison by trial group. In the HIP study, screening seemed to have no impact on diagnoses 3–5 years after the end of the study. The first comparisons of the Swedish overview, based on

<table>
<thead>
<tr>
<th>SES-2</th>
<th>Intervention</th>
<th>Control</th>
<th>Rate ratio†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n  Rate/104</td>
<td>n  Rate/104</td>
<td>(95% CI)</td>
</tr>
<tr>
<td>High</td>
<td>87 5·47</td>
<td>31 8·11</td>
<td>0·68 [0·44–1·03]</td>
</tr>
<tr>
<td>Medium</td>
<td>44 5·16</td>
<td>60 6·08</td>
<td>0·86 [0·58–1·27]</td>
</tr>
<tr>
<td>Low</td>
<td>25 4·39</td>
<td>76 5·45</td>
<td>0·78 [0·49–1·44]</td>
</tr>
</tbody>
</table>

*Breast cancer was the underlying cause of death. Rates are those per 10,000 woman-years at risk.
†Adjusted for age at survey entry, and, where appropriate, for cohort.

Table 4: Breast-cancer mortality by SES-2 in cohort 1 (45–64 years)
short follow-up periods (mean 10 years) found only slight differences between follow-up (without censoring) and evaluation (with censoring) models, but later comparisons have shown increased estimates of benefit when censoring is imposed.\(^\text{19}\) We found that estimated benefit was slightly larger when deaths from diagnoses more than 3–4 years after the end of the study were censored, with no evidence for attenuation of benefit at 14 years of follow-up. The HIP study\(^\text{20}\) did show attenuation with, for example, deaths from diagnoses up to study year 7 showing benefits of 34.7%, 29.3%, and 22.2% at 7 years, 10 years, and 14 years of follow-up.

In view of the controversy over age at which screening should start, the results for women aged 45–49 years at entry are particularly important. Since there is no statistical evidence of heterogeneity between breast-cancer mortality benefit and age at entry, there is no reason to suppose that benefit is less for women first screened under 50 years than for older women. Furthermore, when age-specific results were derived, the estimated difference in breast-cancer mortality rates for those aged 45–49 years, although not significant by itself, is no less than in older entrants. In addition, there is no evidence that the benefit emerged later in these younger women. The 25% difference in mortality rate at 14 years of follow-up in women aged 45–49 years at entry agrees closely with most meta-analyses of randomised trials.\(^\text{21}\)

There are the two critical components to the decision whether population screening for breast cancer should be available to women younger than 50 years. First, do women first screened when younger than 50 have lower breast-cancer mortality than those not so screened; and can the same benefit be achieved by screening from the age of 50 years? Even after a US National Institutes of Health consensus conference these issues remain controversial.\(^\text{16–20}\) Although our numbers for such women are small, the Edinburgh trial findings make an important contribution to the first question.

The second question is more complex. Deaths classified by age at diagnosis require careful interpretation; for purposes of comparison with the HIP trial\(^\text{7}\) we have presented results for younger entrants to the trial with diagnoses at 50 years or older and diagnoses at less than 50 years censored. Our results agree with those of the HIP trial that the benefit in women entering the trial before the age of 50 years is evident only in deaths occurring in their fifties and from cancers diagnosed after age 50 years. These observations could be artefacts with lead time advancing age at diagnosis to under 50 years in some women in the intervention group who, if unscreened, would have been diagnosed in their fifties. A proportion of the benefit (possibly 70%)\(^\text{22}\) will be attributable to screening of women in their fifties. The design of the Edinburgh trial means that its analyses, based on randomisation, can address the second question for younger entrants to cohorts 2 and 3 (since almost all trial screening was done before age 50 years and women in both trial groups were eligible for Forrest screening), but the numbers available for analysis are small and the results equivocal. An observational study\(^\text{23}\) on these cohorts and entrants at ages 45–49 years and 50–52 years to the initial cohort of the intervention group reported lower mortality rates for women in their fifties screened under 50 years.

The improved method of analysis and longer follow-up period have now shown a reduction of 21% in breast-cancer mortality for women aged 45–64 years at entry to the intervention group of the Edinburgh trial; this difference is of borderline significance. Consideration of the results at follow-up periods of 7 years and 10 years, and 14 years follow-up censored by study year of diagnosis indicates a benefit of 25–29% in a population offered regular screening (the steady state). The reduction is close to but smaller than the 30% expected when available data were restricted to HIP trial, the Swedish two-county trial, and case-control analyses of screening.\(^\text{24}\) Compared with the Swedish two-county trial\(^\text{11}\) pathological characteristics of cancers in screened women in Edinburgh showed that screening in Edinburgh (which used current mammographic technology) has not advanced the diagnosis sufficiently to influence histological grade despite reducing size and frequency of node involvement.\(^\text{25}\) Although there is no reason to believe this finding is not typical of the UK, it may explain why mortality benefit has not been larger. Mammographic standards have certainly improved since the Edinburgh study but our data come from a research setting with a 2-year interval between screening. These data may not be applicable without reduction in benefit to screening done as part of routine health service with a 3-year interval. Only cautious optimism is appropriate. Subgroup analyses presented here and by the Swedish investigators,\(^\text{19}\) and one Swedish trial of younger women,\(^\text{26}\) are very promising for women younger than 50 years when first screened, but we believe that decisions on service screening in this age-group should await the results of specifically designed randomised trials (UK Age Trial and EUROTRIAL).

Contributors
Freda Alexander was responsible for leading the trial during its later stages, advised on the statistical analysis and study design, and drafted the paper; Tom Anderson was responsible for pathological aspects and assisted in the preparation of the paper; Helen Brown did most of the statistical analyses; Patrick Forrest provided advice on treatment and causes of death, and assisted in drafting the paper; Wilma Hepburn constructed the population register from which the trial population and the random allocation of women was derived, and was responsible for the flagging procedures; Alistair Kirkpatrick and Berenicé Muir were responsible for radiology; Robin Prescott advised on statistical analyses and assisted in drafting the paper; Alice Smith was data manager.

Acknowledgments
We thank other members of the project committee for advice and encouragement during the study: M M Andrew, J J K Best, C Brough, W Forbes, R Graer, A Huggins, I J Kinlen, N B Loudon, W Lutz, U MacLean, M M Roberts, J Duncan, and I Sutherland. We wish to record the contribution by M Roberts, who was director of the Edinburgh breast screening clinic from 1979 until her death in 1989. We also thank M A Miller and B Mour for help in coordinating the funding; J Warner and other members of the staff of the Information and Statistics Division of the Scottish Health Service Common Services Agency, the staff at the General Register Office, Scotland, the Lothian Health Board Primary Health Care Division, and general practitioners in Edinburgh for their assistance; our colleagues in the screening clinic and the University Departments of Public Health Sciences, Pathology, and Surgery; and P Bisset for typing the paper.

The Edinburgh randomised trial was funded by the Cancer Research Campaign and the Chief Scientist’s Office (CSO) of the Scottish Home and Health Department.

References