Organised mammography screening reduces breast cancer mortality: A cohort study from Finland

Tytti Sarkeala1, Sirpa Heinävaara1,2 and Ahti Anttila1
1Finnish Cancer Registry, Helsinki, Finland
2Department of Public Health, University of Helsinki, Helsinki, Finland

We evaluated the effectiveness and the efficacy of population-based mammography programme in Finland, and explored associations between the screening performance and the screening efficacy. The main outcome, incidence-based mortality from breast cancer, was estimated by invitation, participation, age at death, and screening centres categorised by recall rates. The study was based on an individual followup of screening invitees and participants from 1992 to 2003. The coverage of screening invitations was 95% among 50–59 years old women, and 20–40% among women aged 60–69 years. We compared observed deaths from breast cancer to expected breast cancer deaths without screening in ages 50–69 at death. The observed deaths were obtained from a cohort of individual invitees ($n = 361,848$). The expected deaths were defined by modelling breast cancer mortality from 1974 to 1981 and from 1982 to 2003 at population level. The population data were derived from the same municipalities ($n = 260$) that were incorporated into the cohort. The breast cancer mortality among the invited women was reduced by 22% (relative risk 0.78, 95% confidence interval 0.70–0.87). After adjusting for the self-selection, the efficacy among the participants was 28% (0.72, 0.56–0.88). No clear association between the recall rates and the screening efficacy was observed. The organised mammography screening in Finland is effective. The relationship between the estimates of process and outcome of mammography is not yet straightforward: effectiveness and efficacy remain the best estimates for evaluating the success of mammography screening.

Key words: effectiveness; mammography; screening; breast cancer; epidemiology

Several randomised trials have reported effectiveness of screening for breast cancer. The combined results showed 25% reduction in breast cancer mortality among women aged 50–69 years at randomisation.1 The cohort studies on service screening from Sweden and Denmark have reported effects at similar level.2–4

In Finland, a pilot study on breast cancer screening began in 1982.2 The nationwide mammography programme started in 1987 and it was implemented gradually. The effectiveness for the first 5 years of the programme was assessed using randomised birth cohorts.6 The study demonstrated a nonsignificant, 24% reduction in breast cancer mortality associated with invitation to screening.

The adaptation of results from randomised trials to routine screening is not self-evident. The performance and validity of mammography screening within and between the European programmes have varied widely.5–7 suggesting differences also in the effectiveness and adverse effects of screening. There has also been variation in the screening policies.1

The effectiveness of mammography screening has been debated during the last years.10–12 and the impact of population-based screening on breast cancer mortality has been analysed in many European countries.2–4,6,13–15 The main aim of the current study was to analyse the effectiveness of mammography service screening in Finland. The additional aims were: (i) to explore the efficacy of screening among the screening participants with appropriate adjustment for self-selection bias, and (ii) to report associations between the screening performance and the screening efficacy.

Material and methods

The randomised implementation period within the Finnish mammography programme ended in 1991. Since then, the actual coverage of screening invitations has been over 95% among women aged 50–59 years. Among 60–64 years old women the invitational coverage has been ~40%, and among women aged 65–69 years 20%.7 The screening interval is 2 years.

We analysed the effect of mammography programme on breast cancer mortality in 1992–2003 by comparing observed breast cancer deaths among screening invitees, participants and nonparticipants with expected breast cancer deaths without screening. The observed deaths and the person-years at risk were obtained from a cohort of individual screening invitees. As all the women from the target age group (50–59) were invited to mammography screening in the study period, no noninvited controls were available. The expected mortality rates without screening were thus estimated by modelling, using population data at municipal level. The data were derived from the Mass Screening Registry, the Finnish Cancer Registry and the National Population Registry.

Defining observed breast cancer deaths and person-years at risk, cohort followup

The cohort included women aged 50 years or more, who in 1992–2003 had been invited at least once to any of the 10 screening centres of the Cancer Society of Finland (CSF). The CSF centres covered 50–60% of the organised mammography in Finland, and were the only screening providers sending service screening information to the Mass Screening Registry regularly throughout the study period.

As individual municipalities in Finland are entitled to conduct the mammography program, and may change their screening provider annually, we excluded 55 municipalities (96,001 women, 21.0%) from the study due to a limited duration of their service screening (<9 years) within the CSF centres. The final number of women in the study was 361,848 and the number of invitations in the study period is 1,166,331. The median number of invitations was 3 (ranging from 1 to 8), and the median followup time 9.8 years.

The individual followup of the invitees was initially started from the date of first invitation in 1992–2003 in ages 50–69. However, the expected breast cancer mortality rates were available at the population level only, where the followup was started at January 1st. To obtain comparability with the population data, we set the individual cohort entry date to January 1st in the first invitation year in 1992–2003. The followup was ended at death, emigration or December 31st 2003. Women responding positively to their first invitation in 1992–2003 were classified as participants, the other invitees were nonparticipants. In our previous study, we discovered that over 90% of first-screen participants reattended at least once after subsequent invitation.7

To study the performance of screening, we categorised the 10 CSF centres into 3 groups. The categorisation reflected variation in the screening specificity, and was based on information on centre-specific recall rates at the subsequent screens in
In the group “Low recall rate” (3 centres) the range in recall rates was 0.9–1.9%, in the group “Intermediate recall rate” (4 centres) 2.3–2.7%, and in the group “High recall rate” (3 centres) 2.8–3.5%. The corresponding attendance rates were 92.1, 91.5 and 93.7%, and the detection rates for the screen-detected cancers (/1000 screened) 3.44, 3.60 and 3.87. The recall rates correlated inversely with the incidence of interval cancers.8

We linked the individual screening data with the National Population Registry and with the Cancer Registry using the unique personal identifier as a key. Person-years at risk and incidence-based (refined) breast cancer deaths were calculated by participation, 5-year age groups at death, and 3 centre categories. Deaths from breast cancers diagnosed during the followup in 1992–2003 represented the observed, refined breast cancer deaths.

Estimating expected mortality rates without screening, population data

Population data consisted of age-specific numbers of women and refined breast cancer deaths by calendar year and by municipality from 1974 to 1985 and 1992 to 2003. The municipalities (n = 260) were the same as in the cohort followup. The period 1974–1985 represented the latest prescreening era of equal length to the study period 1992–2003: some municipalities engaged in mammography screening in 1986, prior to the launch of the national mammography programme. The refined breast cancer deaths were defined as follows: breast cancers diagnosed in 1974–1985 or in 1992–2003 among 50–69 years old women formed the basis. Deaths from these breast cancers within the corresponding 2 periods represented the refined breast cancer deaths.

Population data were used to estimate the expected, refined breast cancer mortality rates without screening. The rates were modelled by Poisson regression with a logarithmic link function. In the model, 5-year age groups at death, centre categories (low, intermediate and high recall rates), period before (1974–1985) and with screening (1992–2003), calendar year within the 2 periods (1, 2, ..., 12 years), and interaction between the calendar year and the age at death were used as explanatory variables. The first 3 variables were categorical, and the 4th variable was numerical. The average difference in the refined breast cancer mortality between the two 12-year periods represented the overall screening effect. Birth cohorts were also studied, but they were excluded from the model, because they were strongly correlated with the estimated screening effect. Likelihood ratio statistics and descriptive evaluation between the model-based and observed rates were used as decision criteria in formulation of the model.

The fitted mortality values of the model were calculated with the screening effect excluded, and they represent the expected, refined breast cancer mortality rates without screening in 1992–2003.

Table 1 contains the person-years and the observed numbers of refined deaths from breast cancer in the prescreening (1974–1985) and the screening (1992–2003) periods by age at death and by centre categories. The observed rates from 1974 to 1985 and the expected rates without screening from 1992 to 2003 are also shown. The accumulation of breast cancer deaths by 5-year age groups at death in 1974–1985 and 1992–2003 are illustrated in Figure 1.

### Formulating effect estimates

We studied the ratio between the observed and expected refined breast cancer deaths within the cohort followup in ages 50–69 at death. The observed deaths were divided by the corresponding number of expected deaths. The expected breast cancer deaths without screening were calculated by multiplying the expected mortality rates derived from the population data with the corresponding person years derived from the cohort followup. The confidence intervals were corrected with an over-dispersion constant produced by the model (1.36). The bias due to self-selection was adjusted by method described by Cuzick et al.16

We further analysed, whether the variation in the refined breast cancer mortality by centre category was consistent with the previous association between rising recall rates and incidence of interval cancers.8 We restricted this examination to deaths at 50–64 years of age, as this invitational age range was covered in all the 3 centre categories.

### Results

In 1992–2003, altogether 2,731,268 person-years were accumulated to the cohort in ages 50–69. The number of refined breast cancer deaths was 617 (Table II). Among all invitees, the reduction in the refined breast cancer mortality in ages 50–69 at death was 22% (relative risk 0.78, 95% confidence interval 0.70–0.87). The mortality reduction was highest, 29%, in women aged 60–64 years at death (0.71, 0.58–0.87). In ages 50–54 and 55–59 years at death, the relative risks were 0.74 (0.55–0.96) and 0.84 (0.69–1.01).

The proportion of person-years among participants out of all person-years in the cohort was 87% in ages 50–69 (Table III). The
ratio between the observed and expected breast cancer deaths among the participants was 0.66 (0.58–0.75). The corresponding relative risk among the nonparticipants was more than 2 times higher, 1.56 (1.25–1.91). After adjusting for the self-selection in attendance, the relative risk among the participants became 0.72 (0.56–0.88).

In the 3 centre categories grouped by ascending order of recall rates (low, intermediate, high), the risk ratios among the screening invitees were 0.83 (0.67–1.01), 0.71 (0.58–0.85) and 0.79 (0.60–1.02) (Table IV). Among participants, the person-years at risk were 626,171 (87.5% of invitees), 953,664 (87.7%), and 465,540 (88.3%). After adjusting for the self-selection, no clear association between the screening efficacy and the recall rates could be found (Fig. 2).

Discussion

We compared the observed deaths from breast cancer with screening to the expected breast cancer deaths without screening in 1992–2003. The observed deaths were obtained from a cohort of individual invitees, participants and nonparticipants. The expected deaths were derived by modelling mortality rates at population level. Concurrent, noninvited controls were not available.

There was a significant, 22% reduction in the incidence-based breast cancer mortality among the screening invitees. The reduction was seen only among the screening participants, while the relative risk among the nonparticipants was significantly elevated. The large difference in the relative risks between the participants and the nonparticipants remained significant after adjustment for selection-bias. No clear association between the performance and the efficacy of mammography screening was observed.

Data and methods

Women invited at least once to any of the 10 centres of CSF in 1992–2003 were members of the cohort until emigration, death or end of the followup irrespective of possible migration to municipalities outside the study area. The fixed accumulation of person-years in the cohort thus slightly differed from the dynamic accu-
mulation of person-years in the population data. These differences were small, however, and are not likely to affect the comparability between the individual and the population level data.

During the study period, an increase in breast cancer incidence and a decrease in breast cancer mortality were reported in the Finnish female population aged 50–69. The decrease in mortality was consistent only since 1995, after a persistent increase since the early 1970s. The increase in breast cancer incidence refers to increase in the background risk, and suggests similar pattern also for the background trends of the refined breast cancer mortality in the current study. The decrease in breast cancer mortality indicates improvements in breast cancer diagnostics and treatment.

Our model on the expected rates used information from 1974 to 1985 and 1992 to 2003, and was constructed to attain both the increase in the background risk and the decrease in the breast cancer mortality. Because of contradicting developments, the eventual difference between the observed rates in 1974–1985 and the expected rates in 1992–2003 remained small: the expected rates in ages 50–69 at death were only 1.3% lower than the corresponding observed rates in the prescreening period (see Table I).

When the reduction of breast cancer mortality within service screening is examined, the contributions of screening, other diagnostic services, and treatment are difficult to separate. In the United States, the effect of screening and adjuvant therapy on breast cancer mortality was estimated to be similar, but the variation in the modelling approaches reflected considerable uncertainty. In East Anglia, an analysis of survival by tumour size was used to separate screening effects from other effects. The authors estimated that 60% of the improved survival was due to earlier diagnosis by screening.

We studied the effectiveness of organised mammography screening in a cohort of individual participants and nonparticipants. The information on screening attendance elaborates the assessment, because health care practices independent of screening are assumed to be similar to all. The nonparticipants, however, may be persons with breast cancers at the time invitations are generated, or persons with unhealthy behaviour. We therefore adjusted the effect estimates of the participants for self-selection. The corrected estimates were well below one referring that selection did not explain the current result. Variability in treatment or in the access of other diagnostic services between the participants and the nonparticipants may have modified the screening effect, however. This issue is beyond the scope of the current study.

The firm impact of screening on breast cancer mortality in Finland is supported by a recent study in 2 Finnish cities employing different screening policies. Compared to a city with no screening (Helsinki), the breast cancer mortality among studied birth cohorts was over 40% lower in a city providing regular screening (Turku). The mortality rates from breast cancer increased between the prescreening and screening periods in Helsinki, while in Turku the rates reduced significantly by 36%

Corrected refers to effect estimates with adjustment for selection bias.

<table>
<thead>
<tr>
<th>TABLE III – PERSON-YEARS, OBSERVED AND EXPECTED NUMBERS OF BREAST CANCER DEATHS, AND EFFECT ESTIMATES WITH 95% CONFIDENCE INTERVALS IN 1992-2003 BY 5-YEAR AGE GROUPS AT DEATH AMONG PARTICIPANTS AND NONPARTICIPANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>50–54</td>
</tr>
<tr>
<td><strong>Nonparticipants</strong></td>
</tr>
<tr>
<td>Person-years</td>
</tr>
<tr>
<td>Observed</td>
</tr>
<tr>
<td>Expected</td>
</tr>
<tr>
<td>Observed/expected</td>
</tr>
<tr>
<td><strong>Participants</strong></td>
</tr>
<tr>
<td>Person-years</td>
</tr>
<tr>
<td>Observed</td>
</tr>
<tr>
<td>Expected</td>
</tr>
<tr>
<td>Observed/expected</td>
</tr>
<tr>
<td>Corrected</td>
</tr>
<tr>
<td>Corrected (95% CI)</td>
</tr>
</tbody>
</table>

Corrected refers to effect estimates with adjustment for selection bias.

<table>
<thead>
<tr>
<th>TABLE IV – PERSON-YEARS, OBSERVED AND EXPECTED NUMBERS OF BREAST CANCER DEATHS, AND EFFECT ESTIMATES WITH 95% CONFIDENCE INTERVALS IN 1992-2003 BY CENTRE CATEGORIES (RECALL RATES, %, IN PARENTHESES), INDIVIDUAL FOLLOWUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (0.9–1.9%)</td>
</tr>
<tr>
<td>Person-years</td>
</tr>
<tr>
<td>Observed</td>
</tr>
<tr>
<td>Expected</td>
</tr>
<tr>
<td>Observed/expected</td>
</tr>
<tr>
<td>(95% CI)</td>
</tr>
</tbody>
</table>

Corrected refers to effect estimates with adjustment for selection bias.

**Figure 2** – Effect estimates with 95% confidence intervals in 3 centre categories by participants and nonparticipants (recall rates, %, in parentheses). Corrected refers to effect estimates with adjustment for selection bias.
A reduction of 24% in the refined breast cancer mortality was demonstrated among 49–64 years old invitees by the group-randomised design in 1987–1991 in Finland. In 1986–1997, a decrease of 19% in the refined mortality was reported by a dynamic study on screening among Finnish women aged 50–59 years at entry. In another dynamic cohort study, the breast cancer mortality in 1987–1997 was compared to mortality during the prescreening period of 1976–1986. The organised mammography was associated with reduced breast cancer mortality, particularly among elderly. Design, geographical coverage and main results of the Finnish studies on the effectiveness of mammography screening are summarised in Table V.

The cohort studies on service screening from Sweden have reported similar or even higher overall effect estimates than the Swedish randomised controlled trials. In 9 counties, representing 45% of the Swedish women, the reduction in breast cancer mortality was significant, 27%, among the 40–69 years old invitees. After adjustment for self-selection, contemporaneous changes in incidence and changes in mortality independent of screening, the reduction among the participants became 39%.

In the Netherlands, the screening programme started in the late 1980s, and by 1997 all the women aged 50–69 were covered. The breast cancer mortality rates in the screening period were compared to those in the prescreening period. From 1997 onwards the reduction was significant. In 2001, the rates were 20% lower than in 1986–1988 among 55–74 years old women.

In Denmark, the effect of organised programme on breast cancer mortality was studied among 50–69 years old invitees in Copenhagen. Historical, national and historical-national groups were used as controls. During the first 10 years of the programme (1991–2000), the refined breast cancer mortality at ages 50–79 years reduced significantly by 25%. After adjusting for the self-selection bias, the reduction among the participants became 37%. The relative risks were 0.57 (0.25–1.30) (9 breast cancer deaths) and 1.08 (0.55–2.10) (34 breast cancer deaths) in ages 50–54 and 55–59 at death in the Copenhagen study.

In the current study, the overall reduction in the refined breast cancer mortality in ages 50–69 was at similar level as reported from Denmark, the Netherlands, and Sweden. The attendance rate in Finland was higher, and the efficacy estimates therefore lower than in Denmark and in Sweden, where the organised screening was continued until 69 years of age. In Finland, the considerable accumulation of postscreening followup in ages 60–69 reduced the mortality reduction in this age group. In ages 50–59, the reduction in Finland was greater than in Denmark. The remarkable decrease in the refined breast cancer mortality in ages 50–54 was unexpected, but appeared consistent both in the individual and population level analyses.

No earlier reports on breast cancer mortality by screening performance have been published. In the current study, no difference in the screening efficacy between the centre categories (low, intermediate, high recall rate) was observed in ages 50–64 (p = 0.4105). Our previous study from 1991 to 2003 reported significant increase in the incidence of interval cancers by ascending recall rate category. The detection rates of nonlocalised breast cancers (interval and screen-detected combined) were similar in all the 3 centre categories, however. The current result may at least partially indicate differences in diagnostic and treatment practices within and/or outside the organised screening. Noteworthy, studies on organised mammography from the Netherlands have reported low recall rates (0.7–1.3%), reduced rates of advanced cancers, and similar impact on the breast cancer mortality as the current study. Thus it is possible that indicators of screening performance may be of limited use in predicting the quantity of mortality reduction. Further studies are warranted.

### Conclusions

The organised breast cancer screening in Finland has been effective. The number of prevented breast cancer deaths could probably be increased, if service screening were uniformly extended to 60–69 years old women. The relationship between the estimates of process and outcome of mammography is not yet straightforward: effectiveness and efficacy remain the best estimates for evaluating the success of mammography screening.