Modern mammography screening and breast cancer mortality: population study

Harald Weedon-Fekjær researcher¹²³, Pål R Romundstad professor of epidemiology¹, Lars J Vatten professor of epidemiology¹⁴

¹Department of Public Health, Norwegian University of Science and Technology, 7491 Trondheim, Norway; ²Oslo Center for Biostatistics and Epidemiology, Department of Biostatistics, University of Oslo, Oslo, Norway; ³Oslo Center for Biostatistics and Epidemiology, Research Support Services, Oslo University Hospital, Oslo, Norway; ⁴Harvard School of Public Health, Department of Epidemiology, Boston, MA, USA

Abstract

Objective To evaluate the effectiveness of contemporary mammography screening using individual information about screening history and breast cancer mortality from public screening programmes.

Design Prospective cohort study of Norwegian women who were followed between 1986 and 2009. Within that period (1995-2005), a national mammography screening programme was gradually implemented, with biennial invitations sent to women aged 50-69 years.

Participants All Norwegian women aged 50-79 between 1986 and 2009.

Main outcome measures Multiple Poisson regression analysis was used to estimate breast cancer mortality rate ratios comparing women who were invited to screening (intention to screen) with women who were not invited, with a clear distinction between cases of breast cancer diagnosed before (without potential for screening effect) and after (with potential for screening effect) the first invitation for screening. We took competing causes of death into account by censoring women from further follow-up who died from other causes. Based on the observed mortality reduction combined with the all cause and breast cancer specific mortality in Norway in 2009, we used the CISNET (Cancer Intervention and Surveillance Modeling Network) Stanford simulation model to estimate how many women would need to be invited to biennial mammography screening in the age group 50-69 years to prevent one breast cancer death during their lifetime.

Results During 15 193 034 person years of observation (1986-2009), deaths from breast cancer occurred in 1175 women with a diagnosis after being invited to screening and 8996 women who had not been invited before diagnosis. After adjustment for age, birth cohort, county of residence, and national trends in deaths from breast cancer, the mortality rate ratio associated with being invited to mammography screening was 0.72 (95% confidence interval 0.64 to 0.79). To prevent one death from breast cancer, 368 (95% confidence interval 266 to 508) women would need to be invited to screening.

Conclusion Invitation to modern mammography screening may reduce deaths from breast cancer by about 28%.

Introduction

The efficacy of mammography screening was tested in randomised trials in the 1970s and 1980s. More than 10 years ago, an overview by the World Health Organization indicated that mammography screening may reduce mortality from breast cancer by 25%. However, the methods used by some of the original trials have been criticised, and a report from the Cochrane Collaboration considered the estimates of mortality benefit from many of those trials to be invalid. Recent advances in modern chemotherapy and adjuvant treatment have improved the survival of women with breast cancer, and progress in treatment has led some investigators to question the need for early detection of breast cancer by mammography screening.

Updated studies are clearly needed, but new randomised trials are not realistic and evaluations of modern screening require accurate information about screening history compared with the timing of breast cancer diagnosis, as well as precise and long term follow-up of mortality. Many observational studies have assessed breast cancer mortality associated with mammography screening, but results have been inconsistent, ranging from no effect to improved mortality benefits than those obtained in the original screening trials. Norway provides an ideal setting to study the effects of mammography screening, but in two previous Norwegian studies that used an incidence based mortality approach, only fractions of the available and potentially important data were included in the analyses.
We analysed data from all women in Norway who were aged 50 to 79 during 1986 to 2009, the period during which the Norwegian mammography screening programme was gradually implemented (1995-2005). We compared the rates of deaths from breast cancer among those who were invited to screening (with a potential for screening effect) with those who had not been invited to screening before breast cancer was diagnosed (without a potential screening effect).

Methods

The Norwegian breast cancer screening programme

The Norwegian breast cancer screening programme was initiated by the Norwegian government in 1995 and introduced in four counties in November of that year. The programme was gradually implemented in the remaining 15 counties, with complete national coverage achieved in 2005. The screening programme is administered by the Norwegian Cancer Registry, and all women aged 50-69 are invited to screening every two years. Two view screening mammograms are taken in breast diagnostic centres exclusively dedicated to the diagnosis and treatment of breast diseases. Two readers independently evaluate the mammograms, and women whose mammograms require further consideration are referred for diagnostic mammography, and, if necessary, for additional clinical evaluation. Attendance for screening has been relatively stable, at approximately 76%.

The reporting of cancer to the Norwegian cancer registry is mandatory, and diagnostic information is obtained separately from clinicians, pathologists, and death certificates, with 0.2% of all cancers ascertained only from death certificates. The unique 11 digit personal identification number of each citizen allows follow-up for cause specific mortality, which is provided by Statistics Norway. We used data used on individual dates of screening invitations, dates of breast cancer diagnoses, and dates of breast cancer deaths.

Study participants

We included all Norwegian women aged 50 to 79 years between 1986 and 2009. The dynamic nature of inclusions and exclusions to the cohort by age means that women contributed person years of observation from the age when they were eligible to be observed until they were censored from further observation, either because of death (from breast cancer or other causes), they had reached 80 years of age, or they had reached the end of follow-up (31 December 2009). The actual number of participating women in dynamic cohorts will vary for each given year, but in 2000 a total of 638 238 women were under observation, and the study included 15 193 034 person years of observation.

The first invitation to take part in the Norwegian mammography screening programme depended on the woman’s county of residence and her birth cohort, but from 1995 to 2005 all women in the country aged between 50 and 69 were gradually invited to participate. The supplementary figure shows the mortality rate of breast cancer in Norway (1986-2009) among women aged 50-79 and the period during which the mammography screening programme was implemented in Norwegian counties (1995-2005).

Statistical analysis

In the analysis, we regarded women with a diagnosis of breast cancer after the invitation date to mammography screening as being exposed to screening, and women with a diagnosis of breast cancer before the invitation date as being unexposed to screening. To assess the effect of invitation to screening we compared incidence based breast cancer mortality among women invited to screening (intention to screen) with those not invited, under the counterfactual assumption that if invited women had not been invited, their risk of death from breast cancer would be similar to that of women who had not (yet) been invited.

To account for differences in age and effects of birth cohort and calendar time, we used a multivariable Poisson regression model. To achieve optimal flexibility in the statistical adjustments, we used natural splines to allow for non-linear variations in age, period, and cohort effects (see R code in supplementary appendix d). In sensitivity analyses, we also tested the statistical models without smoothing of period and cohort effects, and we used age and period models without the birth cohort variable to limit the potential for collinearity. In addition because the rates for breast cancer mortality differed slightly between counties, we adjusted for county of residence. In the Poisson regression analysis we took competing causes of death into account by censoring from further follow-up those women who died from causes other than breast cancer.

The time interval from diagnosis until death from breast cancer varies from a few months to many years, and therefore we carefully separated breast cancers diagnosed in women before invitation to first screening from those diagnosed after invitation to first screening to avoid misclassification of breast cancer deaths according to exposure status (invited or not invited before diagnosis). At the beginning of the implementation period in each county almost all deaths from breast cancer occurred among women with a diagnosis before screening invitations started. Over time a gradually higher proportion of breast cancer deaths could be attributed to breast cancers diagnosed after women had been invited to screening. We accounted for this dynamic change by estimating the proportion of the observed breast cancer mortality that was expected to be due to cancers diagnosed after the first screening invitation, assuming that invitations to screening had no effect on breast cancer mortality. In the estimation we used the interval from diagnosis until death from breast cancer among women (in 10 year age groups) who had not yet been invited. Thus we avoided the lead time bias that would have occurred if we had used the interval from diagnosis until breast cancer death among invited women. As an offset in the statistical modelling we added to the model the estimated proportion of breast cancer deaths that was attributed to breast cancers diagnosed after screening invitation, thus adjusting the expected breast cancer mortality for each group according to invitation status (see supplementary appendix for formulas and implementation).

The individual data were precisely split according to exposure status, with separation of invited and not yet invited women within each age-period-county combination during the implementation period of mammography screening in each county. Thus the analysis compares two groups, using detailed information, with adjustment for differences by age, period, cohort, and county. Using this dynamic modelling approach we could utilise all the available individual data in the analysis, without the limitation of selected comparison groups, as in previous studies using data from Norway.

To account for all random statistical uncertainty, we used bootstrap replications and calculated 95% confidence intervals for the estimated effects associated with invitation to mammography screening. To test the robustness of the results, we repeated the analyses under a broad range of statistical assumptions, including a pure age-period-county model, different smoothing of age and period effects, different choice
of reference period and reference age groups, and varying the effect of screening invitation by calendar year.

Since screening effects are likely to vary by age and time since screening, these variables may not be balanced between comparison groups. In a separate sensitivity analysis we therefore weighted the screening variable based on the simulated screening effects by age and time since screening provided by the CISNET (Cancer Intervention and Surveillance Modeling Network) Stanford simulation model.\textsuperscript{19-21} We also calculated the number of women who need to be invited to screening to prevent one death from breast cancer. The number relates to Norwegian women in the age group 50-69 years in 2009. Firstly, we assumed an effect of screening invitations corresponding to the reduction in breast cancer mortality that we observed in our data. Secondly, we used the observed breast cancer mortality in Norway in 2009 and adjusted for the observed reduction in mortality associated with invitation to mammography screening. Thus we could estimate the likely breast cancer mortality in the absence of screening. Thirdly, we used the observed all cause mortality in Norway in 2009 and calculated the probability that women who were first invited at 50 years of age were alive at a given age (51, 52, 53, and so on up to 79 years of age). Effects of screening are likely to vary by age and by time since screening, but these effects are difficult to estimate empirically owing to a limited number of observations. Therefore we applied the CISNET Stanford model scaled to the observed Norwegian breast cancer mortality reduction to estimate the likely screening effects by age and time since screening. In the CISNET Stanford model, smaller tumour size and lower clinical stage at diagnosis resulting from an earlier diagnosis is assumed to explain potential reductions in breast cancer mortality. By combining the breast cancer mortality rates in Norway in 2009, the estimated reduction in breast cancer mortality, and the CISNET Stanford simulation model, we calculated the absolute reduction in breast cancer mortality that could be attributed to screening within each age group. After combining the estimated reduction in breast cancer mortality with the probability of reaching a certain age, given the observed all cause mortality in 2009, we could summarise the data and estimate the probability that one death from breast cancer could be avoided by being invited to mammography screening. Thus the inverse of that probability yielded the number of women aged 50-69 who need to be invited to screening to prevent one death from breast cancer during their lifetime. (See the spreadsheet in the supplementary appendix for further details.)

All statistical analyses were conducted using the R statistical package\textsuperscript{22} (see the supplementary appendix for details of the calculation).

Results

During 15 193 034 person years of observation, breast cancer deaths occurred in 1175 of the women invited to mammography screening and in 8996 of the women who were not invited. After adjustment for age, birth cohort, county of residence, and underlying national trends in breast cancer mortality, the mortality rate ratio associated with being invited to screening was 0.72 (95% confidence interval 0.64 to 0.79), indicating a 28% lower risk of death from breast cancer in women who were invited for screening compared with women who were not invited (table 1).\textsuperscript{19-21} After the invitations to screening had ended (at 70 years of age), we found that the benefit for breast cancer mortality persisted (table 2),\textsuperscript{19-21} but with a possible gradual decline by time since screening (P for trend 0.35). Thus, between five and 10 years after the invitations to screening had ended, the adjusted mortality rate ratio was 0.79 (95% confidence interval 0.57 to 1.01).

To test the robustness of the findings we repeated the analyses under different statistical assumptions (sensitivity analyses), including leaving out the cohort effect, using non-smoothed period effects, and weighting the screening effect by age and time since screening (table 3).\textsuperscript{19-21} However, these additional procedures did not substantially influence the estimated effect and yielded mortality rate ratios ranging from 0.71 to 0.75. By introducing a period dependent screening effect, the results suggested a possible increasing reduction in breast cancer mortality by calendar year, but that analysis had limited statistical power (P=0.29).

We also estimated how many women between 50 and 69 years of age would need to be invited to mammography screening to prevent one death from breast cancer, based on the estimated effect on breast cancer mortality that we found in this study and the observed all cause and breast cancer specific mortality in Norway in 2009. Overall, 368 (95% confidence interval 266 to 508) women in the age group 50-69 years would need to be invited to biennial mammography screening to prevent one death from breast cancer during their lifetime (see supplementary appendix table for calculation).

Based on the estimated effect of screening invitations (table 1), we also estimated the effect of mammography screening among women who actually attended (approximately 76% of invited women). Thus attendance may be associated with a 37% reduction in breast cancer mortality (0.28/0.76=0.37), and 280 women would need to attend screening to prevent one death from breast cancer (368×0.76=280).

Discussion

In this study, based on more than 15 million person years of observation, we estimated that invitation to mammography screening was associated with a 28% reduced risk of death from breast cancer compared with not being invited to screening, and that 368 women need to be invited to screening to prevent one death from breast cancer. The screening effect persisted but seemed to be gradually reduced after invitations to screening had ended. The large population and long follow-up of mortality provided precise estimates and suggests that chance is unlikely to explain the main findings of the study.

Strengths and limitations of this study

Modern treatment has reduced the number of deaths from breast cancer,\textsuperscript{5 6 20} and in the analysis we took into account the effect of changes in nationwide treatment by adjusting for trends in national breast cancer mortality. To improve and standardise breast cancer treatment across Norway, clinical guidelines were implemented before mammography screening became established. Although some differences in treatment may remain, such differences are unlikely to be systematically related to mammography screening status (invited or not invited).

However, breast diagnostic centres were established in parallel with the Norwegian mammography screening programme and resulted in centralisation of care for women with breast cancer. We cannot exclude the possibility that organisational aspects of care related to these centres may have contributed to some of the decrease in breast cancer mortality that we observed after invitations to screening.
Before the national screening programme, mammography screening was available at private radiology institutions, and many women had mammograms for clinical or screening purposes. Assuming that screening activity was highly frequent, an increase in breast cancer incidence and some increase in ductal carcinoma in situ would be expected to precede the implementation of the screening programme. However, in contrast with this expectation, no clear increase in incidence was observed before the national mammography screening programme was established. Therefore it seems unlikely that screening activity before the national programme could have substantially influenced and attenuated the results of the present study.

**Comparison with other studies**

In some studies, women who attended for mammography screening were compared with women who did not attend. In a review of studies that compared breast cancer mortality in women who did and did not attend for screening programmes in Europe, attendance was estimated to be associated with a breast cancer mortality benefit of 31%. In a recent Norwegian study, attendance was associated with a mortality benefit of 43%. Attendance does, however, imply an active choice, and women who choose to attend may differ from those who choose not to attend in ways that may lead to biased estimates of the screening effect. To prevent such a bias we analysed the data according to whether women were invited or not invited to screening (intention to screen).

Two previous prospective studies in Norway also used incidence based mortality to assess the potential benefits of mammography screening. In contrast with the present study, those studies restricted the analyses to selected comparison groups (birth cohorts or counties) and reported moderate mortality benefits (10% and 11%, respectively) with low precision (wide confidence intervals). In the study by Kalager and colleagues, the low precision was due to a short follow-up of mortality, which ended in 2005. Another limitation was that instead of using detailed information about the actual age of the women and date of screening invitations in each county, the investigators used broad categories that probably resulted in some misclassification of exposure (screening or not, in relation to diagnosis). Also, the investigators included breast cancer deaths based on time of diagnosis and not on the actual time of death. Therefore, women with an earlier diagnosis as a result of screening were more likely to be included as invited cases (deaths) than were unscreened women, whose diagnosis was not forwarded by the screening facility. As a consequence, the association of screening invitation with breast cancer mortality is likely to be diluted in that study. In a separate analysis, we limited our data to more closely match that of Kalager and colleagues, and found a reduction in breast cancer mortality of 14% associated with an invitation to screening, which is slightly stronger than the effect reported by the investigators using even fewer detailed data. In the study by Olsen and colleagues, effects of mammography screening were only assessed for selected birth cohorts and only in the four counties where the screening programme was first introduced. Therefore the investigators missed any effect in the remaining birth cohorts, as well as in the other 15 Norwegian counties.

In a recent comprehensive review of European studies, two (from Denmark and Finland) that used incidence based mortality were identified as particularly reliable. According to those studies, the mammography screening programme in Copenhagen was associated with a 25% reduction in breast cancer mortality, and in Finland, a reduction of 24% was attributed to the recently established mammography screening programme. The Finnish study, however, was associated with substantial statistical uncertainty.

It has been questioned whether the evidence from the original screening trials is still relevant within the context of modern treatment for breast cancer, and with generally greater awareness of the disease among women. Our findings, as well as the results from the Danish and Finnish studies, suggest that the relative effectiveness of mammography screening is comparable to the efficacy reported from some of the randomised screening trials.

In our study the estimated benefit for breast cancer mortality (28%) associated with invitation to mammography screening indicates a substantial effect, but evolving improvements in treatment will probably lead to a gradual reduction in the absolute benefit of screening. Based on breast cancer mortality data from 1980, the Euroscreen Working Group estimated that 111 to 143 women would need to be screened to prevent one death from breast cancer. Using breast cancer mortality data from 2009, we estimated that 368 women in the age group 50-69 years would need to be invited to screening to prevent one death from breast cancer during their lifetime. Our higher number is partly attributable to different assumptions about the duration of the effect of screening and partly attributable to lower breast cancer mortality in the absence of screening. The secular decline in breast cancer mortality caused by progress in treatment is substantial, and one consequence of further improvements in treatment is that increasingly more women will need to be invited to mammography screening to prevent one death from breast cancer.

Instead of using individual screening information (incidence based analysis), other researchers have related the timing of introducing mammography screening to time trends in breast cancer mortality. In these studies, breast cancers that were diagnosed before screening cannot be reliably distinguished from screening detected cancers. A separate analysis of our data, we disregarded individual information about the time of diagnosis, and similar to studies using mortality trend analysis, we also found no association of the time that mammography screening was implemented with breast cancer mortality (data not shown). This illustrates how important it is to properly separate breast cancers according to screening status at diagnosis, otherwise any effect of screening will be diluted and cannot be attributed to screening. Therefore, incidence based mortality and detailed screening status are necessary requirements for an appropriate assessment of the effectiveness of mammography screening.

To avoid bias by subjective modelling, we developed a detailed analysis protocol and submitted it to the Norwegian Research Council before data delivery from the Norwegian Cancer Registry. This study is based on data from the Cancer Registry of Norway. The interpretation and reporting of these data are the sole responsibility of the authors, and no endorsement by the Cancer Registry of Norway is intended nor should be inferred. We thank Sylvia Plevritis and Diego Munoz for providing inputs needed to derive the number of woman needed to screen to avoid one breast cancer death; these inputs were taken from the Stanford breast cancer screening model, funded by the National Cancer Institute CISNET programme U01CA159256, and were generated for this study.

Contributors: HWF designed the study, collected and analysed the data, and wrote the report. PRF critically reviewed the analyses, interpreted the results, and contributed to the writing of the report. LJV participated in the design, analyses, and interpretation of the results, and wrote the report. HWF and LJV are guarantors of the study.
What is already known on this topic

Randomised trials from the 1970s and 80s suggested that mammography screening prevents deaths from breast cancer.

The methods used by most of the original studies have been criticised, and this has raised doubts about the validity of the findings.

New trials on screening are unrealistic, and updated observational studies are needed to reliably compare the effects on breast cancer mortality among screened and unscreened women.

What this study adds

Women invited to screening in the Norwegian mammography screening programme were at a 28% lower risk of death from breast cancer than women who had not (yet) been invited.

Attendance was associated with a 37% lower risk of death from breast cancer than women who had not (yet) been invited.

368 women aged 50-69 would need to be invited to biennial mammography screening to prevent one death from breast cancer during their lifetime.

Funding: This study was supported by the Norwegian Research Council (reference No 189503).

Competing interests: All authors have completed the ICMJE uniform disclosure form and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: This study was approved by the regional committee for ethics in medical research.

Data sharing: The data are available for research projects from the legal administrator of the data, the Norwegian Cancer Registry. For data requests use duse@kreftregisteret.no.

Transparency: The lead author (the manuscript’s guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

### Tables

**Table 1** Mortality rate ratio of breast cancer among women aged 50-79 who were invited or not invited (reference) to the Norwegian mammography screening programme, 1986-2009

<table>
<thead>
<tr>
<th>Screening status</th>
<th>Deaths from breast cancer</th>
<th>Person years*</th>
<th>Rate* (per 100 000)</th>
<th>Unadjusted rate ratio*</th>
<th>Age adjusted rate ratio*</th>
<th>Adjusted† mortality rate ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not invited</td>
<td>8996</td>
<td>12 785 325</td>
<td>70.4</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
</tr>
<tr>
<td>Invited</td>
<td>1175</td>
<td>2 407 709</td>
<td>48.8</td>
<td>0.69</td>
<td>0.71</td>
<td>0.72 (0.64 to 0.79)</td>
</tr>
</tbody>
</table>

*Using incidence based mortality with separation of breast cancer cases (and corresponding person years at risk) diagnosed before and after invitation to the screening programme.

†Adjusted for age, birth cohort, national breast cancer mortality trends, and county of residence.
<table>
<thead>
<tr>
<th>Screening period</th>
<th>Mortality rate ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>During active (biennial) screening period of programme (age 50-69)</td>
<td>0.70 (0.62 to 0.78)</td>
</tr>
<tr>
<td>During first five years after invitations to screening ended</td>
<td>0.77 (0.64 to 0.89)</td>
</tr>
<tr>
<td>5-10 years after invitations to screening ended</td>
<td>0.79 (0.57 to 1.01)</td>
</tr>
</tbody>
</table>
### Table 3: Breast cancer mortality rate ratios associated with invitations to mammography screening programme in alternative (sensitivity) analyses under different statistical assumptions

<table>
<thead>
<tr>
<th>Statistical assumptions</th>
<th>Mortality rate ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main estimate (from table 1)</td>
<td>0.72 (0.64 to 0.79)</td>
</tr>
<tr>
<td>Alternative analyses:</td>
<td></td>
</tr>
<tr>
<td>Screening effect weighted by time since first or last screening*</td>
<td>0.72 (0.65 to 0.80)</td>
</tr>
<tr>
<td>Model without birth cohort effects</td>
<td>0.75 (0.67 to 0.82)</td>
</tr>
<tr>
<td>Model without smoothing of period effects</td>
<td>0.72 (0.64 to 0.79)</td>
</tr>
<tr>
<td>Including broader groups (age 40-89 during 1961-2009), screening effect weighted by time since first or last screening,* and applying incidence based mortality based on pre-screening data</td>
<td>0.75 (0.67 to 0.80)</td>
</tr>
</tbody>
</table>

*According to Cancer Intervention and Surveillance Modeling Network Stanford model, and scaled equal to a constant screening effect between 50 and 74 years of age.