Breast cancer mortality in Norway after the introduction of mammography screening

Anne Helene Olsen¹, Elsebeth Lynge², Sisse H. Njor², Merethe Kumle³, Marit Waaseth⁴, Tonje Braaten¹ and Eiliv Lund¹

¹ Department of Community Medicine, University of Tromsø, Tromsø, Norway
² Department of Public Health, University of Copenhagen, Copenhagen, Denmark
³ Clinical Research Centre, University Hospital Northern Norway, Norway
⁴ Department of Pharmacy, University of Tromsø, Tromsø, Norway

An organized mammography screening program was gradually implemented in Norway during the period 1996–2004. Norwegian authorities have initiated an evaluation of the program. Our study focused on breast cancer mortality. Using Poisson regression, we compared the change in breast cancer mortality from before to during screening in four counties starting the program early controlling for change in breast cancer mortality during the same time in counties starting the program late. A follow-up model included death in all breast cancers diagnosed during the follow-up period. An evaluation model included only breast cancers diagnosed in ages where screening was offered. The study group had been invited for screening one to three times and followed for on average of 5.9 years. In the follow-up model, 314 breast cancer deaths were observed in the study group, and 523, 404 and 638, respectively, in the four control groups. The ratio between the changes in breast cancer mortality between early and late starting counties was 0.93 (95% confidence interval [CI] 0.77–1.12). In the evaluation model, this ratio was 0.89 (95% CI: 0.71–1.12). In Norway, where 40% of women used regular mammography prior to the program, the implementation of the organized mammography screening program was associated with a statistically nonsignificant decrease in breast cancer mortality of around 11%.

In Norway, piloting of an organized mammography screening program started in four counties in 1996 offering biennial screening to women aged 50–69 years. From 1999 onward, the program was gradually implemented in the rest of Norway, becoming nationwide in 2004.

The Norwegian program is currently being evaluated under the mandate of the Norwegian Research Council. As the purpose of mammography screening is to prevent death from breast cancer, determining the effect of screening on breast cancer mortality is an essential part of the evaluation. Ideally, our study should be carried out comparing women offered screening with women in an unexposed control group. In a nonrandomized setting, where all women are offered screening, this is a complicated task. The obvious possibilities are use of a historical or a regional control group. However, these control groups will also reflect historical trends or regional differences. To take account of this, a study design with three control groups consisting of women in the screening region prior to screening and women in the nonscreening regions both during and prior to screening has been previously used for the evaluation of screening in Denmark.

In addition to the nonrandomized setting, evaluation of the Norwegian program was complicated by the fact that regular mammography was widespread prior to the program and that the implementation of screening coincided with a dramatic increase in hormone use.

In this article, we determined the effect of the organized mammography screening program on breast cancer mortality in the four pilot counties.

Material and methods

Study design

Organized screening started in Rogaland, Oslo, Hordaland and Akershus in 1996. The program was gradually expanded across Norway. The latest counties to be included were Oppland, Møre og Romsdal in 2002, Sogn og Fjordane and Hedmark in 2003, and Vestfold in 2004. Based on the individual data for all women in these counties, we studied change in breast cancer mortality from the period before screening to the period during screening in the early starting counties, controlling for change in breast cancer mortality during the same period in the late starting counties. In this design, the regional control group was exactly similar to the study group except for covering other parts of Norway, and in the same way the historical–regional control group was exactly similar to the historical study group except for covering other parts of Norway.
**Study group**
The study group included women born during the period 1927–1946 and living in Rogaland, Oslo, Hordaland or Akershus on January 1, 1996 or moving in during the study period. Women born during the period 1927–1928 had been invited to screening once, women born during the period 1929–1930 had been invited to screening twice, whereas the remaining birth cohorts had been invited to screening three times. Once included, a woman continued to be in the study group even if she moved to other places in Norway. First invitation date was allocated as the first date of invitation in the woman’s municipality. Women diagnosed with breast cancer prior to first invitation date were excluded. The risk period was from first invitation date for women born during the period 1927–1931 until date of death, emigration, disappearance or December 31, 2008 whichever came first, and for women born during the period 1932–1946 until date of death, emigration, disappearance or December 31, 2001 whichever came first. The latter time restriction was owing to the start of screening in the regional control group (Fig. 1).

**Historical control group**
The historical control group included women born during the period 1921–1940 and living in Rogaland, Oslo, Hordaland or Akershus on January 1, 1990 or moving in during the historical control period. Pseudo first invitation date was allocated as for the study group minus 6 years, and women diagnosed with breast cancer prior to this date were excluded. The risk period was from the pseudo invitation for women born during the period 1921–1925 until date of death, emigration, disappearance or December 31, 2002 whichever came first, and for women born during the period 1926–1940 until the latest December 31, 1995.

**Regional control group**
The regional control group included women born during the period 1927–1946 and living in Oppland, Møre og Romsdal, Sogn og Fjordane, Hedmark or Vestfold on January 1, 1996 or moving in during the study period. Pseudo first invitation date was allocated following a scheme resembling that of the study group, and women diagnosed with breast cancer prior to this date were excluded. The risk period was from the pseudo invitation date for women born during the period 1927–1931 until date of death, emigration, disappearance or December 31, 2008 whichever came first, and for women born during the period 1932–1946 until at the latest December 31, 2001.

**Historical, regional control group**
The historical, regional control group included women born during the period 1921–1940 and living in Oppland, Møre og Romsdal, Sogn og Fjordane, Hedmark or Vestfold on January 1, 1990 or moving in during the historical control period. Pseudo invitation date and exclusion of women with breast cancer were as for the regional control group. The risk period was from the pseudo invitation for women born during the period 1921–1925 until date of death, emigration, disappearance or December 31, 2002 whichever came first, and for women born during the period 1926–1940 until at the latest December 31, 1995.

**Breast cancer deaths**
Cause of death was registered for women where the risk period was terminated by death. The number of women with breast cancer as the cause of death constituted the observed number of cases in the analysis.

**Data sources**
Individual data were retrieved from the Central Population Register on inhabitants, and dates of moving between municipalities, emigration and death. From the Norwegian Cancer Register, data were retrieved on breast cancer diagnosis and from the Norwegian Cause of Death Register data on cause of death. Data were linked based on the personal identification number used in all public Norwegian registers.

**Analysis**
Breast cancer mortality rates in the study and control groups were compared using Poisson regression with the variables age at time of death, period and region. Age was divided into 50–54, 55–59, ... 80–84 years. Period was study versus historical period. Region was Rogaland, Oslo, Hordaland and
Akershus versus Oppland, Møre og Romsdal, Sogn og Fjordane, Hedmark and Vestfold. The method has been reported previously in detail.\(^1\) In brief, the ratio of the breast cancer mortality rate of the study group compared to that of the historical control group was designated RR1, and the ratio of the breast cancer mortality rate of the regional control group compared to that of the historical regional control group was designated RR2. The effect of screening hereafter called the RR and the 95% confidence interval (CI) for this RR were estimated in a Poisson model with 5-year age group, period, region and exposure to screening as parameters. Given similar age distributions, the RR can be considered as the ratio of RR1 and RR2 being an estimate of the change from before to during screening in the study group controlled for the contemporary change in the regional control group. Under the assumption of no interaction between period and region, the RR will estimate the breast cancer mortality in women offered organized screening compared to the breast cancer mortality expected in the absence of organized screening. Following the terminology by Nystrøm et al.,\(^5\) this calculation is referred to as the follow-up model.

As the periods of follow-up extended up to 13 years after end of invitation to screening some deaths from breast cancer diagnosed after women left the program will be included in our study group. To take account of this, we performed an alternative analysis, where only death in breast cancer cases diagnosed when women were invited to screening, and equivalent for the control groups, were included in the analysis. The four groups were still followed up for breast cancer deaths for exactly the same time intervals as used in the follow-up model. Following the terminology by Nystrøm et al.,\(^5\) this is referred to as the evaluation model.

SAS version 9.2 was used for the analysis. The study was approved by the Regional Committee for Medical Research Ethics and the Norwegian Data Inspectorate.

**Results**

In the follow-up model, the study group accumulated 1,182,747 person-years and 314 breast cancer deaths. The study group had been invited for screening one to three times and followed for on average of 5.9 years. The historical control group accumulated slightly fewer person-years (2.5% fewer) and 404 breast cancer deaths (Table 1). Across age groups, the breast cancer mortality rate was consistently lower in the study group than in the historical control group, though being fairly similar in the age group 75–79 years, 79 and 82 per 100,000, respectively. The Poisson regression gave a RR1 of 0.83 (95% CI: 0.72–0.97). The regional control group accumulated 1,576,270 person-years and 523 breast cancer deaths. The historical–regional control group accumulated somewhat more person-years (16.4% more) and 638 breast cancer death. Apart from the age group of 55–59 years, the breast cancer mortality rate was consistently lower in the regional control group than in the historical regional control group. This resulted in a RR2 of 0.89 (95% CI: 0.79–1.00). In the historical period, the breast cancer mortality was lower in the historical control group than in the historical regional control group; age-standardized rates of 35 and 39 per 100,000, respectively. The RR, the ratio of RR1 and RR2, was 0.93 (95% CI: 0.77–1.12) (Fig. 2). This means that the follow-up model showed a 7% statistically nonsignificant reduction in breast cancer mortality in women offered organized screening compared to what would have been expected in the absence of the organized screening.

In the evaluation model, the study group accumulated by definition the same number of person-years as in the follow-up model. The number of breast cancer deaths decreased to 206, a decrease by definition restricted to the ages from 70 years and above. The number of breast cancer deaths in the historical control group decreased to 259. In the evaluation model, the RR1 became 0.82 (95% CI: 0.68–0.98). In the regional control group, the number of breast cancer deaths was 371, and in the historical regional control group, the number was 429. This resulted in a RR2 of 0.91 (95% CI: 0.79–1.04). The RR was 0.89 (95% CI: 0.71–1.12). This means that the evaluation model showed an 11% statistically nonsignificant reduction in breast cancer mortality.

**Discussion**

We found that implementation of organized mammography screening program in Norway was associated with a statistically nonsignificant decrease in breast cancer mortality of 7% in the follow-up model and of 11% in the evaluation model.

**Strengths and limitations**

Our study had several strengths. First, it was based on data linkage where both person-years and breast cancer deaths were counted according to the individual exposure time. Second, our study design was balanced in calendar time where the study group and the regional control group covered the same time period, as did the historical control group and the historical regional control group. In this way, underlying temporal changes in breast cancer mortality were taken into account.

Our study had, however, also limitations. First, both estimates might slightly underestimate the true effect of the program. As stated above, the 7% included breast cancer deaths in women diagnosed after they left the program. This means that we have included some breast cancer cases that cannot have benefited from screening. For the calculation of the 11%, the follow-up period was longer than the accrual period. Owing to the lead time, the analysis will, therefore, include deaths from more breast cancers in the study than in the control groups, as the lead time is not known these numbers cannot be estimated. As all women in a given municipality were allocated the same invitation date, we might furthermore have included some breast cancers diagnosed before first invitation to screening. Second, owing to the short-time interval between program start in study and regional control groups, a large proportion of study group observations, 44% of person-years and 71% of breast cancer deaths, derived
from women offered screening at age 64 or above. Third, for women offered screening below age 64 the maximal follow-up time was 6 years, which might not have allowed for the full effect of screening to materialize. Fourth, our results have to be interpreted in light of the widespread use of regular mammography in Norway prior to implementation of the program, and the dramatic change in hormone use during the implementation phase. Fifth, in our study design the estimated effect of the program could not be separated from a possible interaction between region and period.

Concerning use of regular mammography, we have previously surveyed historical data on mammography activity in Norway. A key finding in this survey was that questionnaire data from the Norwegian Women and Cancer Study (NOWAC) from 1996 showed that at least 40% of Norwegian women aged 50–69 regularly underwent mammography prior to their first invitation to the program. The survey furthermore showed that this percentage increased over time, as 64% of first attendees in the organized program in 1996–2002 reported to have had at least one mammography prior to their first participation in the program. Finally, our survey showed that the self-reported NOWAC data were supported by mammography activity data collected by the Norwegian Radiation Protection Authority. A high preprogram mammography activity in Norway has been noted also by other authors.

Our previously reported data on mammography activity in Norway can be used in the interpretation of the breast cancer mortality data reported in the present study. Given that a substantial proportion of women had mammography prior to their participation in the program, the breast cancer mortality data reported here do not reflect the impact of screening versus no screening. Our evaluation reflects instead the impact of building a program on top of existing widespread regular mammography.

### Table 1. Effect estimates for breast cancer mortality in the Norwegian mammography screening programme by age at death

<table>
<thead>
<tr>
<th>Study group</th>
<th>Age group</th>
<th>No of breast cancer deaths</th>
<th>Person-years</th>
<th>Breast cancer mortality per 100,000 person-years</th>
<th>Relative risk estimates (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50–54</td>
<td>17</td>
<td>247,710</td>
<td>7</td>
<td>RR1: 0.96 (0.47–1.97)</td>
</tr>
<tr>
<td></td>
<td>55–59</td>
<td>40</td>
<td>234,527</td>
<td>17</td>
<td>RR2: 1.00 (0.60–1.65)</td>
</tr>
<tr>
<td></td>
<td>60–64</td>
<td>34</td>
<td>185,615</td>
<td>18</td>
<td>RR3: 1.00 (0.83–1.27)</td>
</tr>
<tr>
<td></td>
<td>65–69</td>
<td>29</td>
<td>178,093</td>
<td>16</td>
<td>RR4: 0.70 (0.50–1.00)</td>
</tr>
<tr>
<td></td>
<td>70–74</td>
<td>80</td>
<td>188,345</td>
<td>42</td>
<td>RR5: 0.69 (0.50–1.00)</td>
</tr>
<tr>
<td></td>
<td>75–79</td>
<td>106</td>
<td>134,642</td>
<td>79</td>
<td>RR6: 0.96 (0.74–1.24)</td>
</tr>
<tr>
<td></td>
<td>80–84</td>
<td>8</td>
<td>13,815</td>
<td>58</td>
<td>RR7: 0.83 (0.72–0.97)</td>
</tr>
<tr>
<td></td>
<td>All age groups</td>
<td>314</td>
<td>1,182,747</td>
<td>27</td>
<td></td>
</tr>
</tbody>
</table>

1Follow-up model. 2Study group (Akershus, Oslo, Rogaland and Hordaland). 3Historical control group (Akershus, Oslo, Rogaland and Hordaland). 4Regional control group (Oppland, Møre og Romsdal, Sogn og fjordane, Hedmark and Vestfold). 5Historical regional control group (Oppland, Møre og Romsdal, Sogn og fjordane, Hedmark and Vestfold).
the possible impact of the preprogram mammography activity on the expected effect of the program on breast cancer mortality. If the Norwegian program was assumed to work in line with the randomized controlled trials on mammography screening, a 25% decrease in breast cancer mortality should be expected, given no preprogram mammography activity. However, with the preprogram mammography activity we found in the survey, only an effect of 11% would be measured under the assumption that mammography within and outside the organized screening worked similarly. The present study showed that the real breast cancer mortality data presented here in fact corresponded very well with the estimate based only on the trial results and the Norwegian mammography activity data. This might not be so surprising as the Norwegian health care system is expected to work fairly similar to the Swedish, from where most randomized controlled trial data derived, and the Danish, from where observational data showed similar results.

Concerning hormone use, sales data from the Norwegian Institute of Public Health on hormones from 1987, 1991, 1995, 1999 and 2002 showed number of defined daily doses per 1,000 women aged 45–64 years per day to be higher in pilot than in control counties. Data after 2002 were not so relevant as women followed past 2002 were all above the age of 70 years. Hormone use could well be expected to have affected breast cancer mortality in Norway during the program implementation, but given the parallel development over time in the pilot and control counties, we do not expect hormone use to have distorted the present study.

Concerning the possible interaction between region and period, it should be considered that breast cancer care units were established in Norway along with the program. One would therefore expect breast cancer treatment, and consequently breast cancer mortality, to have improved more from the historical to the screening period in pilot than in control counties. This would represent an interaction between region and period. Patients diagnosed in the post-program period, but prior to invitation to screening, indeed had a better survival than patients diagnosed in the preprogram period. However, as these data were not available separately for pilot and control counties, it was difficult to draw a firm conclusion on interaction.

Previous studies
A previous study by Kalager et al. of the effect of the Norwegian program to some extent resembled our approach in using three control groups. It differed though by not being balanced in calendar time, by including regions with a very short follow-up time, and by estimating person-years from routine population statistics. For women aged 50–69 years at time of diagnosis, the study found the program to be associated with a 10% reduction in breast cancer mortality. In interpretation of these findings, Kalager et al. emphasized the breast cancer care units introduced along with the program. As breast cancer mortality trends for women diagnosed at age 70–84 years changed in parallel with those for women of screening age, the authors indirectly suggested the mortality reduction to be explained by new treatment modality which, when available, was used for breast cancer patients of all ages. As stated above, we agree that a possible interaction between region and period cannot be separated out from the analysis. However, the mortality in postscreening age
Epidemiology


measured by Kalager et al. was not unaffected by screening. Breast cancer deaths were categorized by age at diagnosis, and lead time will therefore affect the measured mortality for both women aged 50–69 and 70–84 years. The lead time will result in more deaths being allocated to the age group of 50–69 years and less deaths being allocated to the age group of 70–84 years in the screening group as compared to the three control groups. The possible impact of regular mammography prior to the program was not considered by Kalager et al.

In another study, Autier et al. analyzed age-adjusted breast cancer mortality rates from Sweden and Norway. In Sweden, screening was introduced gradually starting in 1986, whereas in Norway gradual implementation started in 1996. From 1989 to 2006, breast cancer mortality decreased by 16.0% in Sweden and by 24.1% in Norway. Based on these data combined with trends from other European countries, the authors concluded that breast cancer mortality trends were not associated with the presence of screening programs. The possible impact of regular mammography prior to the program in Norway was not considered by Autier et al.

Previous studies have shown that the start of the organized screening program left marks on the breast cancer pattern in Norway, and one may ask whether such marks are compatible with the use of regular mammography also before the start of the program. First Zahl et al. showed a prevalence peak in the breast cancer incidence after start of screening in the pilot counties. Based on our data on mammography use, a prevalence peak is expected, as the proportion of women in the four pilot counties reporting prior mammography increased from 47% in 1996 to 73% in 1997–1998. Weedon-Fekjær et al. estimated the incidence increase to be 59% during the Norwegian prevalence peak, where a 78% increase would have been expected in the absence of prior screening. Second, Sørum et al. showed that the incidence of ductal carcinoma in situ (DCIS) increased with the start of the program. The most relevant comparison, taking into account that DCIS detection may also be affected by technological changes, is from a rate of 10 per 100,000 women-years before the program to 30 per 100,000 women-years during the early subsequent invitation rounds. This increase may, to a large extent, although not completely, be explained by the simultaneous increase in proportion of women with previous mammography. In the 11 counties where the program was first implemented, this percentage increased from 43% in 1996 to 92% in 2002.

Methodological considerations

The evaluation of breast cancer mortality after implementa-
tion of the organized mammography screening program in Norway is a prime example of the limitations of observational epidemiology. In the present situation, the disease outcome could be affected both by mammography activity outside the program, hormone use, treatment improvement and by the organized program. These factors could not be fully separated in the observational data. We therefore have to consider also other possibilities to learn about the effect of the program. First, simulation models can incorporate data on mammography activity outside the program, hormone use, treatment improvement and based on this estimate the effect of the program. Simulation models do, however, build on assumptions, and different models may come up with different estimates based on the same data. Second, short-term surrogate indicators can be measured. They reflect characteristics of the screening activity known from the randomized controlled trials to predict breast cancer mortality. Breast cancers detected in the Norwegian program had a favorable tumor size distribution although the rate of advanced tumours remained constant. The detection rate was in line with European recommendations for initial screens and higher for subsequent screen. However, the sensitivity has been relatively low, indicated by a high interval cancer rate as compared to the background incidence rate in the second year of follow-up. If the Norwegian program had been implemented in an unscreened population, one might, based on the short-term indicators roughly, have expected an effect on breast cancer mortality at or slightly below the 25% reduction found in randomized controlled trials.

Conclusion

In conclusion, we found that in Norway where 40% of women used regular mammography prior to the program, the implementation of the organized mammography screening program was associated with a statistically nonsignificant decrease in breast cancer mortality of around 11%. The study group had been invited for screening one to three times and followed for on average 5.9 years.

Acknowledgements

The Research protocol was approved by the Norwegian Research Council. The data were analyzed by Anne Helene Olsen. Some data in this article are from the Cancer Registry of Norway. The Cancer Registry of Norway is not responsible for the analysis or interpretation of the data presented.

References


