Effect of Screening Mammography on Breast-Cancer Mortality in Norway

Mette Kalager, M.D., Marvin Zelen, Ph.D., Frøydis Langmark, M.D., and Hans-Olov Adami, M.D., Ph.D.

ABSTRACT

BACKGROUND
A challenge in quantifying the effect of screening mammography on breast-cancer mortality is to provide valid comparison groups. The use of historical control subjects does not take into account chronologic trends associated with advances in breast-cancer awareness and treatment.

METHODS
The Norwegian breast-cancer screening program was started in 1996 and expanded geographically during the subsequent 9 years. Women between the ages of 50 and 69 years were offered screening mammography every 2 years. We compared the incidence-based rates of death from breast cancer in four groups: two groups of women who from 1996 through 2005 were living in counties with screening (screening group) or without screening (nonscreening group); and two historical-comparison groups that from 1986 through 1995 mirrored the current groups.

RESULTS
We analyzed data from 40,075 women with breast cancer. The rate of death was reduced by 7.2 deaths per 100,000 person-years in the screening group as compared with the historical screening group (rate ratio, 0.72; 95% confidence interval [CI], 0.63 to 0.81) and by 4.8 deaths per 100,000 person-years in the nonscreening group as compared with the historical nonscreening group (rate ratio, 0.82; 95% CI, 0.71 to 0.93; P<0.001 for both comparisons), for a relative reduction in mortality of 10% in the screening group (P=0.13). Thus, the difference in the reduction in mortality between the current and historical groups that could be attributed to screening alone was 2.4 deaths per 100,000 person-years, or a third of the total reduction of 7.2 deaths.

CONCLUSIONS
The availability of screening mammography was associated with a reduction in the rate of death from breast cancer, but the screening itself accounted for only about a third of the total reduction. (Funded by the Cancer Registry of Norway and the Research Council of Norway.)
On the basis of several randomized clinical trials,1-3 the World Health Organization concluded in 2002 that screening mammography for women between the ages of 50 and 69 years reduced the rate of death from breast cancer by 25%.4 Nevertheless, the use of screening mammography is still debated, chiefly because of concern regarding methodologic limitations in some of the randomized trials.5 In addition, the benefit of mammography when implemented in a population-based service program remains poorly quantified. Therefore, continued evaluation of breast-cancer screening programs is warranted.6

The main challenge in quantifying the reduction in mortality from nonrandomized screening programs is to provide valid comparison groups. Although historical, prescreening control groups are often used, such a comparison has important limitations because it does not take into account confounding by chronological trends in breast-cancer mortality, reflecting such factors as advances in breast-cancer awareness and treatment. According to a statistical model based on data regarding breast-cancer mortality in the United States from 1975 through 2000, only half the observed reduction in mortality was causally related to the mammographic intervention itself, whereas the other half was attributable to improved management.7 To establish a valid comparison group, we took advantage of several unique features of the nationwide Breast Cancer Screening Program in Norway, which was implemented by means of gradual geographic expansion over a 9-year period.

**METHODS**

**SCREENING PROGRAM**

Norway, with a total population of 4.8 million, has a public health care system. Patients generally receive treatment in their county of residence, and there is no private primary care for breast cancer.8 The nationwide Cancer Registry of Norway is close to 100% complete.9,10 Patients are identified in the registry by their individually unique national registration number, which includes the date of birth. The registry runs the Breast Cancer Screening Program, which began as a pilot project in 4 of the 19 Norwegian counties in 1996. Two years later, the government decided to expand the program, and over a period of 9 years, the remaining 15 counties were enrolled in a staggered fashion11 (Fig. 1).

The rollout of the program followed no specific geographic pattern. Since 2005, all women in the country between the ages of 50 and 69 years have been invited to participate in screening mammography every 2 years.

Before enrollment in the program, each county was required to establish multidisciplinary breast-cancer management teams and breast units.12 As a result, breast-cancer management became centralized for all residents within each county, and dedicated teams of radiologists, radiologic technologists, pathologists, surgeons, oncologists, and nurses managed the care of all patients, regardless of age.

The screening program is organized with 26 stationary and 4 mobile screening units.13 The Central Population Registry of Norway identifies eligible women on the basis of their national registration number. Invitations are mailed to each eligible woman, suggesting a time for an appointment.14 Overall, 77% of all women who are invited participate in the program.15 In accordance with European guidelines, mammograms are obtained in two views, which are independently read by two radiologists.12

**STUDY GROUPS**

From Statistics Norway we retrieved information on the Norwegian female population, according to county, from January 1, 1986, through December 31, 2005.16 From the Cancer Registry, we retrieved data on all women who had received a diagnosis of invasive breast cancer, including age, tumor stage, date and county of residence at diagnosis, date and cause of death, and information on whether the diagnosis had been made before or after the implementation of the screening program.

By comparing two current groups on the basis of whether screening mammography was available in the county, we would avoid confounding by factors such as improvements in treatment and heightened awareness, temporal changes that may be associated with a reduction in breast-cancer mortality. However, we could not make direct comparisons between these two groups because of the nonconstant risk of death from breast cancer according to the time since diagnosis and differences in rates of death from breast cancer between counties before implementation of the screening program.15 To adjust for such differences and to achieve equal follow-up time in each county, we
established two historical comparison groups that mirrored the implementation of the screening program during the 10-year period preceding the screening program.

Thus, we defined four groups of women, including those in whom a first invasive breast cancer had been diagnosed: two current groups of women who from 1996 through 2005 were living either in counties in which the screening program had been implemented (screening group) or in counties in which the program had not been implemented (nonscreening group), and two historical-comparison groups that from 1986 through 1995 mirrored the county residence of the current groups before the implementation of the screening program (Fig. 1) (see the Supplementary Appendix, available with the full text of this article at NEJM.org).

As pointed out, each county was required to establish multidisciplinary breast-cancer management teams and breast units before enrollment in the national screening program. As a result, the screening program consists of two components: screening mammography and care from multidisciplinary teams. For women between the ages of 50 and 69 years who were invited to participate in the program, the change in mortality after the introduction of the screening program can be related to both the introduction of screening mammography and the establishment of multidisciplinary teams. However, for women who were outside the age range that was eligible for the screening program (i.e., those between the ages of 20 and 49 years and those between the ages of 70 and 84 years) in the counties in which screening was available, the change in mortality could be related only to the establishment of multidisciplinary teams, since these women were not invited to undergo mammography.

STUDY OVERSIGHT

The Norwegian Social Science Data Services approved the study, which was funded by the Cancer Registry of Norway and the Research Council of Norway. The study was conducted in accordance with the protocol, which is available at NEJM.org.

STATISTICAL ANALYSIS

We obtained information on breast cancer as the underlying cause of death through regular linkage between the Cancer Registry and the Cause of Death Registry at Statistics Norway. To isolate the
effect of the breast-cancer screening program, our calculation of mortality in the screening group includes only deaths from breast cancer in women who received the diagnosis after the screening program was implemented (so-called incidence-based mortality). The use of incidence-based mortality avoids the inclusion of breast-cancer deaths that occurred after implementation of the screening program but reflected diagnoses that were made before the program was implemented. So as not to bias our comparisons, we calculated the rate of death in all groups using the incidence-based method. All women in whom breast cancer was diagnosed and who died of breast cancer after implementation of the screening program were included in the screening group, regardless of whether they received the diagnosis at a screening or a diagnostic examination.

On the basis of the date of implementation of the screening program in each county, we grouped the 19 counties into six regions; each county within a given region entered the program at approximately the same time (see the Supplementary Appendix). We compared the rates of death separately for each region. Thus, the regional comparisons have the same follow-up time. This grouping tended to reduce random variation resulting from small numbers and permitted the evaluation of changes in mortality in the same region over a period of time. First, we compared women in the nonscreening group with their historical counterparts to determine the temporal change in mortality that was not attributable to the introduction of the screening program and that was likely to reflect improved treatment and earlier clinical diagnosis. Then, we compared women in the screening group with their historical counterparts to determine the change in mortality after implementation of the screening program. In this second comparison, the difference in the rate of death between the two groups can be attributed both to the screening program and to temporal trends in mortality that were unrelated to the screening program. Thus, the reduction in mortality that was related to the screening program was the difference between the rate ratio for death among women in the screening group as compared with their historical counterparts and the rate ratio for death among women in the nonscreening group as compared with their historical counterparts.

We estimated rates of death from breast cancer in the four study groups according to the age at diagnosis (20 to 49 years, 50 to 69 years, and 70 to 84 years). All tests of statistical significance were one-sided, and a P value of less than 0.05 was considered to indicate statistical significance. (For additional details on the statistical analysis plan, see the Supplementary Appendix.)

RESULTS

SUBJECTS

A total of 40,075 women received a diagnosis of breast cancer between 1986 and 2005. During the follow-up period, 4791 of these women (12%) died from breast cancer. Of the women who died, 423 (9%) had received the diagnosis after the introduction of the screening program. The total follow-up time for the study was 31,613,529 person-years, with an average of 2.2 years and a maximum of 8.9 years of follow-up for women with breast cancer. Among women between the ages of 50 and 69 years, 6967 received a diagnosis of breast cancer between 1986 and 1995, as compared with 12,056 who received the diagnosis between 1996 and 2005. In the latter group, 7975 women (66%) had been invited to participate in screening mammography. In the first screening round, a total of 454,331 women had been invited.

Among women between the ages of 50 and 69 years in the screening group, the rate of death was 18.1 per 100,000 person-years, as compared with 25.3 per 100,000 person-years among their historical counterparts, for a difference of 7.2 deaths per 100,000 person-years (rate ratio, 0.72; 95% confidence interval [CI], 0.63 to 0.81; P<0.001), a relative reduction of 28% (Table 1 and Fig. 2). Among women in the nonscreening group, the rate of death was 21.2 per 100,000 person-years, as compared with 26.0 per 100,000 person-years among their historical counterparts, for a difference of 4.8 deaths per 100,000 person-years (rate ratio, 0.82; 95% CI, 0.71 to 0.93; P<0.001), a relative reduction of 18% (Table 1 and Fig. 2). Given the reduction in mortality among women in the nonscreening group, as compared with their historical counterparts, the relative reduction among women in the screening group was 10% (95% CI, −4 to 24; P=0.13). Since the differences between the current groups and historical groups were 7.2 deaths per 100,000 person-years in the screening group and 4.8 deaths per 100,000 person-years in the nonscreening group, only the overall between-
group difference — 2.4 deaths per 100,000 person-years (95% CI, −1.7 to 6.5) — can be attributed to the screening program alone, representing a third of the total estimated reduction in mortality (2.4 of 7.2).

Among women between the ages of 50 and 69 years in the screening group, those with stage I tumors had a relative reduction in mortality of 16%, as compared with their historical counterparts (rate ratio, 0.84; 95% CI, 0.63 to 1.11); among women in the nonscreening group, the corresponding reduction was 13% (rate ratio, 0.87; 95% CI, 0.62 to 1.23). Among women with stage II tumors, those in the screening group had a marked 29% reduction in mortality, as compared with their historical counterparts (rate ratio, 0.71; 95% CI, 0.58 to 0.86); among women in the nonscreening group, the reduction was 7% (rate ratio, 0.93; 95% CI, 0.76 to 1.12). Among women with stage III or IV tumors, the improvement in prognosis was similar with and without the screening program (rate ratio for death in both groups, 0.70; 95% CI, 0.57 to 0.86 for the screening group and 0.56 to 0.87 for the nonscreening group).

Among women who were not eligible for screening because they were younger than 50 years of age or older than 69 years of age, there was also a significant reduction in the rate of death from breast cancer, as compared with their historical counterparts (Table 1). Women in these age groups who were in the screening group but were not eligible for the screening program had the benefit of the multidisciplinary breast-cancer management teams. Among women under the age of 50 years, there was a nonsignificant relative increase in mortality of 4% (P=1.00) after the introduction of the screening program.
screening program (Table 1). Among women who were 70 years of age or older, the relative reduction in mortality of 8% (P=0.09) could be attributed to the establishment of multidisciplinary teams in the screening program (Table 1 and Fig. 3).

**DISCUSSION**

In our study, the rate of death from breast cancer was reduced by the introduction of a breast-cancer screening program. However, when we took into account temporal trends in breast-cancer mortality caused by other factors, the apparent effect was considerably smaller than expected. Indeed, the take-home message is that breast-cancer screening was associated with an absolute reduction of 10 percentage points in the rate of death from breast cancer. However, the screening program accounted for only one third of the total reduction in mortality among women who were invited to participate in the program. For women between the ages of 50 and 69 years, it was impossible to determine whether the reduction in mortality resulted from earlier diagnoses associated with screening mammography or from the management of treatment by an interdisciplinary team. To our surprise, the reduction in breast-cancer mortality among women between the ages of 70 and 84 years was largely the same as that in the screening group. Although none of the older women were invited to undergo mammography, they were all treated by multidisciplinary teams specializing in breast-cancer care.

The fundamental prerequisite for our analysis was the staggered implementation of the Norwegian Breast Cancer Screening Program. This structure provided the opportunity to identify a non-screening group in order to reduce or perhaps eliminate confounding as a result of temporal changes in breast-cancer mortality attributable to factors other than screening. Additional strengths of our study include its nationwide design, the large size, the high proportion of women participating in the screening program (77%), and the complete follow-up. The incidence-based approach for calculating rates of death also reduced the likelihood that results were obscured by deaths from breast cancers that were diagnosed before the screening program was implemented.

Is it possible that the lead time created a bias in calculating incidence-based mortality? We counted the rate of death from breast cancer only if the death and diagnosis occurred in that group. For example, in the screening group, a death would be attributed to breast cancer only if the disease was diagnosed early by means of screening mammography or if the disease was clinically diagnosed while the woman was in the group. However, for women in whom an early diagnosis was made at screening and who later died of breast cancer, the diagnosis would have been made clinically at an unknown time within the study period. Thus, the lead time plays no role in the calculation of the rate of death, and we believe that the mortality calculations for all groups are free of this bias.

Our study also has limitations. First, the maximum follow-up time of 8.9 years may be too short to show the full potential of the screening program. However, in randomized, controlled trials, there was a reduction in mortality after 4 years, with an increasing effect up to 10 years. In our study, the reduction in mortality was seen mainly in the first 4 years of follow-up (data not shown). Second, since the screening program was implemented gradually in the counties, diagnoses were...
made more recently in the screening group than in the nonscreening group (Fig. 1) and there may be an overestimation of the mortality benefit associated with the screening program. Third, some of the women in the nonscreening group may have actually undergone mammography (opportunistic screening), potentially resulting in an underestimation of the benefit of screening. Unfortunately, we have no precise information about the numbers of such examinations. However, several circumstances provide reassuring evidence against contamination by opportunistic screening as an important source of bias. Before the implementation of the screening program, access to mammography was limited, especially in the predominantly rural areas of the country, and the reduction in mortality was of similar magnitude in urban and rural areas (data not shown). Also, the public health care system provides no financial incentives for offering screening mammography. Finally, the organized screening mammography entailed a substantial increase in diagnosed cases of breast cancer, with no similar trends in counties before they joined the program.

Our finding that only about one third of the reduction in mortality can be directly attributed to breast-cancer screening is in line with evidence from the National Health Service screening program in the United Kingdom. Other studies have shown a relative reduction in the rate of death from breast cancer of 6.4 to 25% with follow-up periods of 10 years or less. However, most of these studies have compared current breast-cancer mortality with mortality in a period preceding the introduction of screening mammography, with no ability to account for the confounding effect of temporal trends. As our data show, such confounding may entail a considerable overestimation of the mortality benefit of mammography.

The implementation of multidisciplinary breast-cancer management teams was intended to provide comprehensive and integrated optimization of breast-cancer care. As a corollary, it is not possible to attribute the reduction in mortality to any specific component of such a change in health care, although increased breast-cancer awareness, higher sensitivity of diagnostic techniques, and improvements in treatment can all be conducive to a lower rate of death. The greatest reduction in the death rate associated with mammography was observed among women with stage II tumors. This finding might be explained by selective stage migration among screening participants as a result of more sensitive staging techniques (including the use of sentinel-node biopsy, which increased from virtually no use in 1998 to a 65% rate of use in 2004) and improvements in treatment.

We conclude that our results support the evidence that screening mammography reduces the rate of death from breast cancer. However, the magnitude of this benefit seems modest in the high-attendance, nationwide screening program we evaluated. Most important, the apparent benefit conveyed by optimized patient care may be missed unless breast-cancer screening is integrated into a well-functioning health care system that is available to the entire population.

Figure 3. Incidence-Based Rate Ratios for Death from Breast Cancer, According to Age Group.

Shown are the differences in breast-cancer mortality among women living in counties in which breast-cancer screening had been implemented, as compared with their historical counterparts, and corresponding values for women living in counties in which screening had not been implemented, as compared with their historical counterparts. Only women between the ages of 50 and 69 years were invited to participate in mammographic screening.
REFERENCES


