The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 SEPTEMBER 23, 2010 VOL. 363 NO. 13

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.)

From the Cancer Registry of Norway, Oslo (M.K., F.L., H.-O.A.); the Departments of Epidemiology (M.K., H.-O.A.) and Biostatistics (M.Z.), Harvard School of Public Health; and the Dana–Farber Cancer Institute and Harvard Medical School (M.Z., H.-O.A.) — all in Boston; and the Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm (H.-O.A.). Address reprint requests to Dr. Kalager at Oslo University Hospital, Department of Surgery, Montebello, 0310 Oslo, Norway, or at mkalager@hsph.harvard.edu.

N Engl J Med 2010;363:1203-10.

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 fashion[11] (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

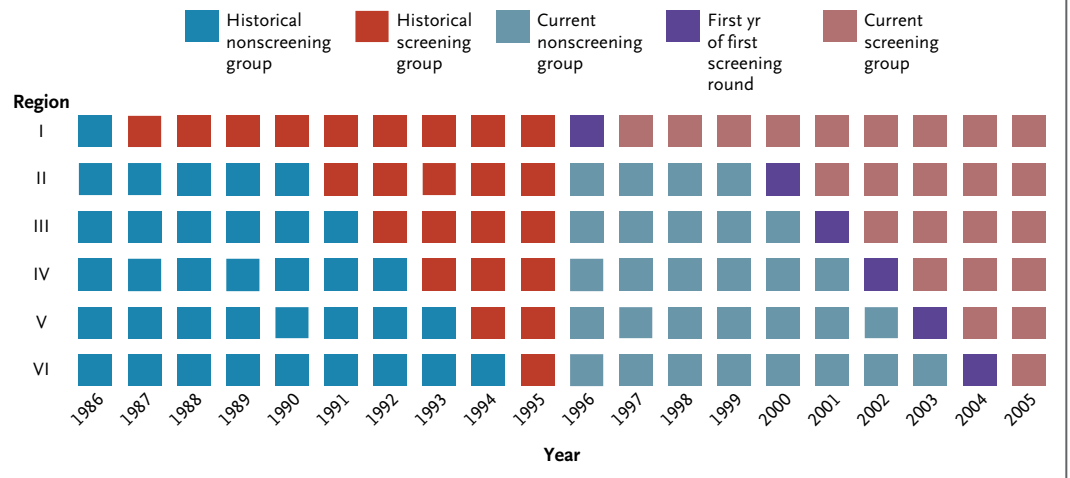

Figure 1. The Four Study Groups, According to Region and Year.

The 19 counties were grouped into six regions according to the date of introduction of the screening program, which was implemented throughout the country in a staggered fashion, starting in 1996. The screening group consisted of women who received a diagnosis of breast cancer after the introduction of the screening program. The nonscreening group consisted of women living in regions where screening was not offered in the same calendar period that screening was offered in other regions. The historical study groups consisted of women residing in the 19 counties in the 10-year period before screening was offered. A screening round lasted for 2 years, and the first year of the first round was included in both the screening and nonscreening groups (purple).

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).[17-19] 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-

Table 1. Rates of Death from Breast Cancer, According to Study Group and Age.*

Age Group and Mortality Data	Nonscreening Groups		Screening Groups		Difference		
	Historical Group	Current Group	Historical Group	Current Group	Nonscreening Groups†	Screening Groups‡	Nonscreening Groups vs. Screening Groups§
50–69 Yr							
No. of deaths	494	396	555	423			
No. of person-yr	1,898,989	1,866,741	2,197,469	2,337,323			
No. of deaths/100,000 person-yr	26.0	21.2	25.3	18.1	4.8	7.2	2.4±4.1
Rate ratio for death (95% CI)					0.82 (0.71–0.93)	0.72 (0.63–0.81)	0.10
20–49 Yr							
No. of deaths	238	183	332	267			
No. of person-yr	3,842,740	4,030,443	5,134,212	5,357,163			
No. of deaths/100,000 person-yr	6.2	4.5	6.5	5.0	1.7	1.5	–0.2±4.4
Rate ratio for death (95% CI)					0.73 (0.63–0.92)	0.77 (0.65–0.90)	–0.04
70–84 Yr							
No. of deaths	429	386	623	465			
No. of person-yr	1,101,019	1,173,624	1,349,967	1,318,004			
No. of deaths/100,000 person-yr	39.0	32.9	46.1	35.3	6.1	10.8	4.7±6.9
Rate ratio for death (95% CI)					0.84 (0.74–0.97)	0.76 (0.68–0.86)	0.08

* Only women between the ages of 50 and 69 years were invited to participate in screening mammography. All women in this group were also eligible for treatment by the multidisciplinary teams that are part of the screening program.

† For the nonscreening groups, the value shown is the difference between the rate of death in the historical group and that in the current group. This difference represents changes in mortality over time as a result of increased breast-cancer awareness, improved therapy, and more sensitive diagnostic tools.

‡ For the screening groups, the value shown is the difference between the rate of death in the historical group and that in the current group. This difference represents changes in mortality both over time and after introduction of the breast-cancer screening program.

§ For the comparison of the nonscreening groups with the screening groups, the value shown is the difference between the two rate-of-death differences. This value represents the effect of introducing the breast-cancer screening program. Plus–minus values are 95% confidence intervals.

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

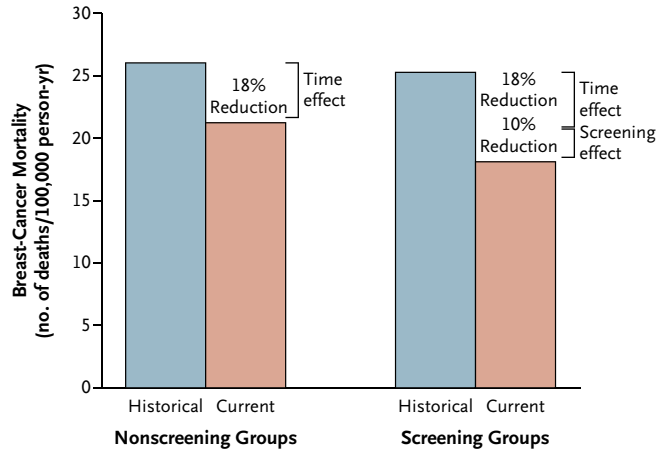

Figure 2. Rates of Death among Women between the Ages of 50 and 69 Years in the Four Study Groups.

Among women in the nonscreening group, there was an 18% reduction in the rate of death from breast cancer, as compared with the preceding 10-year period, presumably as a result of increased breast-cancer awareness, improved therapy, and the use of more sensitive diagnostic tools. Among women in the screening group, there was a 28% reduction in mortality from breast cancer during the same period. Thus, the relative reduction in mortality that was causally related to the screening program alone was 10%.

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 nonscreening 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.[20] 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.[21] 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.[18,19,21-25] 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.[18,21,23-25] As our data show, such confounding may entail a considerable overestimation of the mortality benefit of mammography.[23-25]

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[26] 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[15]) and improvements in treatment.

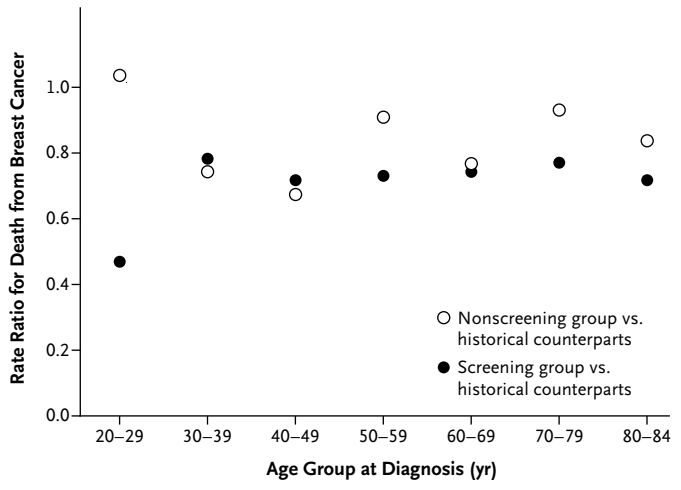

Figure 3. Incidence-Based Rate Ratios for Death from Breast Cancer, According to Age Group.

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.

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.

Supported by the Cancer Registry of Norway and the Research Council of Norway.

No potential conflict of interest relevant to this article was reported.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Michael Bretthauer, M.D., Ph.D., of Oslo University Hospital for valuable help with editing and scientific discussion during the preparation of the manuscript.

REFERENCES

1. Chu KC, Smart CR, Tarone RE. Analysis of breast cancer mortality and stage distribution by age for the Health Insurance Plan clinical trial. J Natl Cancer Inst 1988;80:1125-32.
2. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. CMAJ 1992;147:1477-88. [Erratum, CMAJ 1993; 148:718.]
3. Nyström L, Rutqvist LE, Walls S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. Lancet 1993;341:973-8. [Erratum, Lancet 1993;342:1372.]
4. Vainio H, Bianchini F, eds. IARC handbook of cancer prevention. Vol 7. Breast cancer screening. Lyon, France: IARC Press, 2002.
5. Gøtzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2006;4: CD001877.
6. Nyström L. How effective is screening for breast cancer? BMJ 2000;321:647-8.
7. Berry DA, Corin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med 2005;353:1784-92.
8. Consumer report 4/2002. Oslo: The Consumer Council of Norway. (In Norwegian.) (Accessed August 30, 2010, at http://forbrukerportalen.no/filearchive/4_2002_s_1_48.pdf.)
9. Larsen IK, Småstuen M, Johannesen TB, et al. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. Eur J Cancer 2009;45:1218-31.
10. Tingulstad S, Haldorsen T, Norstein J, Hagen B, Skjeldestad FE. Completeness and accuracy of registration of ovarian cancer in the Cancer Registry of Norway. Int J Cancer 2002;98:907-11.
11. Kalager M, Haldorsen T, Bretthauer M, Hoff G, Thoresen SO, Adami HO. Improved breast cancer survival following introduction of an organized mammography screening program among both screened and unscreened women: a population-based cohort study. Breast Cancer Res 2009;11:R44.
12. Perry N, Broeders M, de Wolf C. European guidelines for quality assurance in breast cancer screening and diagnosis. 2nd ed. Luxembourg: Office for Official Publications of the European Communities, 1996.
13. Erzaas A. Quality assurance manual of the Norwegian Breast Cancer Screening Program. Oslo: The Cancer Registry of Norway, 2003. (In Norwegian.) (Available at http://www.kreftregisteret.no.)
14. Wang H, Kåresen R, Hervik A, Thoresen S. Mammography screening in Norway: results from the first screening round in four counties and cost effectiveness of a modeled nationwide screening. Cancer Causes Control 2001;12:39-45.
15. Kalager M, Kåresen R, Wist E. Survival after breast cancer — differences between Norwegian counties. Tidsskr Nor Laegeforen 2009;129:2595-600. (In Norwegian.)
16. Statistics Norway home page. (Accessed August 30, 2010, at http://statbank.ssb.no/statistikkbanken/Default_FR.asp?PXSid=0&nvl=true&PLanguage=0&tilside=selecttable/MenuSelS.asp&SubjectCode=02.)
17. The Swedish Organised Service Screening Evaluation Group. Reduction in breast cancer mortality from organized service screening with mammography: 1. Further confirmation with extended data. Cancer Epidemiol Biomarkers Prev 2006;15:45-51.
18. Tabar L, Duffy SW, Yen MF, et al. All-cause mortality among breast cancer patients in a screening trial: support for breast cancer mortality as an end point. J Med Screen 2002;9:159-62.
19. Hakama M, Pukkala E, Heikkilä M, Kallio M. Effectiveness of the public health policy for breast cancer screening in Finland: population based cohort study. BMJ 1997;314:864-7.
20. Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomized trials. Lancet 2002;359:909-19. [Erratum, Lancet 2002;360:724.]
21. Blanks RG, Moss SM, McGahan CE, Quinn MJ, Babb PJ. Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990-8: comparison of observed with predicted mortality. BMJ 2000;321:665-9.
22. Olsen AH, Njor SH, Vejborg I, et al. Breast cancer mortality in Copenhagen after introduction of mammography screening: cohort study. BMJ 2005;330:220.
23. Tabar L, Yen MF, Vitak B, Chen HHT, Smith R, Duffy SW. Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. Lancet 2003; 361:1405-10.
24. Otto SJ, Fracherboud J, Looman CWN, et al. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. Lancet 2003;361:1411-7.
25. Jonsson H, Nyström L, Törnberg S, Lenner P. Service screening with mammography of women aged 50-69 years in Sweden: effects on mortality of breast cancer. J Med Screen 2001;8:152-60.
26. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon: stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. N Engl J Med 1985;312: 1604-8.

EDITORIALS

Screening Mammography — A Long Run for a Short Slide?

H. Gilbert Welch, M.D., M.P.H.

No screening test has ever been more carefully studied than screening mammography. In the past 50 years, more than 600,000 women have participated in 10 randomized trials, each involving approximately 10 years of follow-up. Given this extraordinary research effort, it is ironic that screening mammography continues to be one of the most contentious issues within the medical community.

The juxtaposition of such a charged medical debate in the face of such an exhaustive scientific investigation is in itself instructive. For context, one trial involving fewer than 150 men who were followed for less than 2 years was sufficient to convince physicians of the value of treating severe hypertension.[1] That physicians are still debating the relative merits of screening mammography despite the wealth of data suggests that the test is surely a close call, a delicate balance between modest benefit and modest harm.

In this issue of the *Journal*, Kalager et al.[2] provide additional data that the benefit of mammography is modest. Making use of the opportunity provided by the staggered implementation of a national screening program in Norway, the investigators were able to isolate the benefit of the screening program from other factors that may have changed over time, including increased breast-cancer awareness and improvements in treatment. They report that the benefit of the Norwegian screening program was disappointingly small: a 10% reduction in breast-cancer mortality among women between the ages of 50 and 69 years.

Moreover, this reduction in mortality reflected the combined effect of the two interventions that make up the Norwegian screening program: screening mammography and multidisciplinary teams instituted to better treat breast cancer. Kalager et al. provide data that the latter may be the more important of the two factors, since women over the age of 70 years, who were exposed to the program's multidisciplinary teams but were not invited to undergo mammography, had an 8% reduction in breast-cancer mortality. Thus, the relative reduction in mortality due to screening mammography alone could be as low as 2%.

Clinicians who follow the mammography debate will reasonably wonder why the benefit estimated by Kalager et al. is so much smaller than the reduction in mortality of 15 to 23% estimated by the U.S. Preventive Services Task Force.[3] The easiest explanation would be that the Kalager estimate is wrong. Although the task force uses data from randomized trials, the Norwegian data are observational — and as with all observational data, the primary threat to validity is the comparability of the comparison groups.

But the staggered cohort design that was used by Kalager et al. mitigates the concern that the women in the four study groups are somehow different, since many of the women in the study actually contributed data to each group at different points in their life. Contamination is a more relevant concern. If the women in the nonscreening groups were exposed to opportunistic mammography screening or began to benefit from the multidisciplinary teams, which had to be in place before the screening program was initiated, then the background effect of time may have been overestimated. This would have led to an underestimation of the benefit of the screening program. Furthermore, the follow-up period may be too short to fully capture the benefits of screening. The authors argue that these effects are small.

So another explanation must be considered:

the estimates of both the task force and Kalager et al. are correct. But where the randomized trials reflect the world before 1990, the observational data reflect the world after 1990. It is quite plausible that screening mammography was more effective in the past than it is now. If women with new breast lumps now present earlier for evaluation, the benefit of screening will be less. If treatment of clinically detected breast cancer (i.e., tumors that are detected by means other than screening) has now improved, the benefit of screening will be less. Thus, the increased awareness about the importance of promptly seeking care for overt breast abnormalities (there is no debate about diagnostic mammography) and the widespread use of adjuvant therapy have probably combined to make screening now less important.[4,5]

Nevertheless, the public widely perceives screening mammography to be one of the most important services provided by modern medicine. The perception is largely the product of well-crafted public health messaging, such as the American Cancer Society's print campaign in the 1980s that featured the headline "If you haven't had a mammogram, you need more than your breasts examined." Given current data, such messaging must become more balanced.

If we assume that mammography screening is associated with a 10% reduction in the rate of death from breast cancer (making the optimistic assumption that all the benefit comes from screening mammograms), the 10-year risk of breast-cancer death for a 50-year-old woman in the United States is now about 4 per 1000 women.[6] If we assume that this risk already incorporates the benefit of screening mammography, the risk estimate without mammography would be about 4.4 per 1000 women.

Because we are all subject to framing effects, it is important to consider the reverse frame. The number of women who will not die from breast cancer rises from 995.6 to 996 per 1000 women with the addition of screening mammography. Although readers may each respond differently to these frames, both reflect the same absolute benefit: 0.4 per 1000 women. In other words, 2500 women would need to be screened over a 10-year period for 1 to avoid death from breast cancer (Table 1).

What happens to the other 2499 women who had to undergo screening to achieve this benefit is also relevant. Estimates of harm vary considerably. In the United States, more than 1000 women would be expected to have at least one false positive result,[7] a number that would be considerably lower in Europe.[8] Less frequent but more worrisome is the problem of overdiagnosis. Somewhere between 5 and 15 women would be expected to be needlessly treated for a condition that was never going to bother them, with all the accompanying harms.[9,10]

Table 1. Estimated Benefits and Harms Associated with a 10-Year Course of Screening Mammography for 2500 Women Who Are 50 Years of Age.*

Benefit	Harm
One woman will avoid dying from breast cancer.	Up to 1000 women will have at least one "false alarm," about half of whom will undergo biopsy.
	Breast cancer will be overdiagnosed in 5 to 15 women, who will be treated needlessly with surgery, radiation, chemotherapy, or a combination.

* The assumed benefit of screening mammography is a reduction of 10% in the rate of death from breast cancer, as reported by Kalagar et al.[2]

Screening mammography has become one of the most prominent measures of health care performance. Since the inception of health care report cards, such evaluations have focused on ensuring that all women undergo the test.[11] There were practical reasons for this: it was easily measured, easy to understand, and hard to argue against. But by highlighting that the mortality benefit is modest, Kalager et al. help confirm that the decision about whether to undergo screening mammography is, in fact, a close call. Many observers will argue that because it is a delicate decision — involving trade-offs among noncomparable outcomes — it must be left to informed individuals to decide. Others will argue that physicians should continue to persuade women to undergo screening and that the modest benefit is worth the associated harms.

But no one can argue that screening mammography is one of the most important services we provide in medicine. The time has come for it to stop being used as an indicator of the quality of our health care system.

Disclosure forms provided by the author are available with the full text of this article at NEJM.org.

From the Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth Medical School, Lebanon, NH.

1. Effects of treatment on morbidity in hypertension: results in patients with diastolic blood pressures averaging 115 through 129 mm Hg. JAMA 1967;202:1028-34.
2. Kalager M, Zelen M, Langmark F, Adami H-O. Effect of

screening mammography on breast-cancer mortality in Norway. N Engl J Med 2010;363:1203-10.
3. Mandelblatt JS, Cronin KA, Bailey S, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Intern Med 2009;151:738-47.
4. Rostgaard K, Vaeth M, Rootzén H, Lynge E. Why did the breast cancer lymph node status distribution improve in Denmark in the pre-mammography screening period of 1978-1994? Acta Oncol 2010;49:313-21.
5. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687-717.
6. The risk of developing/dying from cancer based on November 2009 SEER data submission. Bethesda, MD: National Cancer Institute. (Available at http://seer.cancer.gov/faststats/.)
7. Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. N Engl J Med 1998;338:1089-96.
8. Breast cancer screening. In: IARC handbooks of cancer prevention. Vol. 7. Lyon, France: International Agency for Research on Cancer, 2002:135-44.
9. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. BMJ 2006;332:689-92.
10. Jørgensen KJ, Gøtzsche PC. Content of invitations for publicly funded screening mammography. BMJ 2006;332:538-41.
11. McIntyre D, Rogers L, Heier EJ. Overview, history, and objectives of performance measurement. Health Care Financ Rev 2001;22(3):7-21.

Superficial Phlebitis and Phase 3.5 Trials

Lee Goldman, M.D., and Jeffrey Ginsberg, M.D.

In this issue of the *Journal*, Decousus et al.[1] report on the efficacy and safety of fondaparinux for the treatment of superficial-vein thrombosis in the legs. The results of their carefully conducted, placebo-controlled trial show that treatment with fondaparinux, at a dose of 2.5 mg once daily for 45 days, as compared with placebo, reduced the probability that superficial-vein thrombosis in the legs would progress to deep-vein thrombosis or pulmonary embolism (1.3% with placebo vs. 0.2% with fondaparinux), without an increase in bleeding or other serious adverse events. The probability that patients would undergo surgery for superficial-vein thrombosis was reduced from 3.8% to 0.7%. Two patients in the fondaparinux group and one in the placebo group died, but none of the deaths were apparently the result of a pulmonary embolism. This study adds to previous work describing the natural history of superficial-vein thrombosis,[2-5] although it did not address which patients might be at an increased risk because of previously undiagnosed thrombophilia.[4,5]

To put the rates of deep-vein thrombosis and pulmonary embolism — the most important outcomes — into perspective, it is useful to consider the generally "acceptable" failure rates in strategies to diagnose venous thromboembolism. In the study by Decousus et al., the rate at which symptomatic deep-vein thrombosis or pulmonary embolism developed in untreated patients during follow-up (1.3%) was similar to the rate with widely accepted strategies for diagnosing deep-vein thrombosis and pulmonary embolism. For example, among patients who are evaluated for suspected deep-vein thrombosis but have normal results on a contrast venogram[6] or duplex ultrasonography,[7] about 1.3% and 0.6% of patients, respectively, will return with symptomatic deep-vein thrombosis or pulmonary embolism over the course of long-term follow-up. Similarly, among patients who have a suspected pulmonary embolism but then have normal results on a conventional pulmonary angiogram[8] or a computed tomographic pulmonary angiogram,[9] about 1.7% and 1.2%, respectively, will return with symptomatic deep-vein thrombosis or pulmonary embolism. These historical comparisons and the extremely low mortality among untreated patients with superficial-vein thrombosis support an initial "no anticoagulant treatment" approach, unless conservative measures fail to resolve symptoms or deep-vein thrombosis develops. It is also clear from the stringent inclusion and exclusion criteria in the study by Decousus et al. that treatment with fondaparinux for 45 days is clinically reasonable for patients with severe symptoms, thrombosis in the proximal saphenous vein, or recurrent disease.

Agents such as fondaparinux, low-molecular-weight heparins, and perhaps oral direct factor Xa inhibitors (apixaban, rivaroxaban) and thrombin inhibitors (dabigatran) have better risk profiles than do unfractionated heparin and warfarin, and the favorable risk-to-benefit ratio associated with them could lead to an extension

BMJ 2014;348:g3701 doi: 10.1136/bmj.g3701 (Published 17 June 2014)

RESEARCH

Modern mammography screening and breast cancer mortality: population study

Harald Weedon-Fekjær *researcher*[1 2 3], Pål R Romundstad *professor of epidemiology*[1], Lars J Vatten *professor of epidemiology*[1 4]

[1]Department of Public Health, Norwegian University of Science and Technology, 7491 Trondheim, Norway; [2]Oslo Center for Biostatistics and Epidemiology, Department of Biostatistics, University of Oslo, Oslo, Norway; [3]Oslo Center for Biostatistics and Epidemiology, Research Support Services, Oslo University Hospital, Oslo, Norway; [4]Harvard 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.[1] More than 10 years ago, an overview by the World Health Organization indicated that mammography screening may reduce mortality from breast cancer by 25%.[2] 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.[3 4] Recent advances in modern chemotherapy and adjuvant treatment have improved the survival of women with breast cancer,[5 6] and progress in treatment has led some investigators to question the need for early detection of breast cancer by mammography screening.[7]

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.[8-16] Norway provides an ideal setting to study the effects of mammography screening,[17 18] 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.[8 11]

Correspondence to: H Weedon-Fekjær harald.weedon-fekjar@medisin.uio.no

Extra material supplied by the author (see http://www.bmj.com/content/348/bmj.g3701?tab=related#datasupp)

Figure showing breast cancer mortality rates in Norway from 1986 through 2009
Appendices a-f

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.[18] 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 colinearity. 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.[8 11]

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.[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[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⇓).

After the invitations to screening had ended (at 70 years of age), we found that the benefit for breast cancer mortality persisted (table 2⇓), 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⇓). 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,[5-20] 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.[23] 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.[24 25] 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%.[12] In a recent Norwegian study, attendance was associated with a mortality benefit of 43%.[26] 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.[27] 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.[8 11] 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,[8] 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',[8] 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,[11] 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,[13] two (from Denmark and Finland) that used incidence based mortality were identified as particularly reliable.[9 28] According to those studies, the mammography screening programme in Copenhagen was associated with a 25% reduction in breast cancer mortality,[8] 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.[28]

It has been questioned whether the evidence from the original screening trials is still relevant within the context of modern treatment for breast cancer,[5 6] and with generally greater awareness of the disease among women. Our findings, as well as the results from the Danish and Finnish studies,[9 28] suggest that the relative effectiveness of mammography screening is comparable to the efficacy reported from some of the randomised screening trials.[2 29]

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.[5-30] 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.[31] 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.[10 12] In these studies, breast cancers that were diagnosed before screening cannot be reliably distinguished from screening detected cancers. In 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.[32] Therefore, incidence based mortality and detailed screening status are necessary requirements for an appropriate assessment of the effectiveness of mammography screening.[32]

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. PRR 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 some 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

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 at www.icmje.org/coi_disclosure.pdf 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 <due@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.

1 Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002;359:909-19.
2 World Health Organization. IARC handbooks of cancer prevention: handbook 7: breast cancer screening. IARC Press; 2001.
3 Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;355:129-34.
4 Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2013;6:CD001877.
5 Effects of adjuvant tamoxifen and of cytotoxic therapy on mortality in early breast cancer. An overview of 61 randomized trials among 28,896 women. Early Breast Cancer Trialists' Collaborative Group. *N Engl J Med* 1988;319:1681-92.
6 Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. *Lancet* 1998;351:1451-67.
7 Welch HG. Screening mammography—a long run for a short slide? *N Engl J Med* 2010;363:1276-8.
8 Kalager M, Zelen M, Langmark F, Adami HO. Effect of screening mammography on breast-cancer mortality in Norway. *N Engl J Med* 2010;363:1203-10.
9 Olsen AH, Njor SH, Vejborg I, Schwartz W, Dalgaard P, Jensen MB, et al. Breast cancer mortality in Copenhagen after introduction of mammography screening: cohort study. *BMJ* 2005;330:220.
10 Jorgensen KJ, Zahl PH, Gotzsche PC. Breast cancer mortality in organised mammography screening in Denmark: comparative study. *BMJ* 2010;340:c1241.
11 Olsen AH, Lynge E, Njor SH, Kumle M, Waaseth M, Braaten T, et al. Breast cancer mortality in Norway after the introduction of mammography screening. *Int J Cancer* 2013;132:208-14.
12 Autier P, Boniol M, Gavin A, Vatten LJ. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. *BMJ* 2011;343:d4411.
13 Njor S, Nyström L, Moss S, Paci E, Broeders M, Segnan N, et al. Breast cancer mortality in mammographic screening in Europe: a review of incidence-based mortality studies. *J Med Screen* 2012;19(Suppl 1):33-41.
14 Njor S, Nyström L, Moss S, Paci E, Broeders M, Segnan N, Lynge E; Euroscreen Working Group. The impact of mammographic screening on breast cancer mortality in Europe: a review of trend studies. *J Med Screen* 2012;19(Suppl 1):26-32.
15 Paci E, Giorgi D, Bianchi S, Vezzosi V, Zappa M, Crocetti E, et al. Assessment of the early impact of the population-based breast cancer screening programme in Florence (Italy) using mortality and surrogate measures. *Eur J Cancer* 2002;38:568-73.
16 Ascunce EN, Moreno-Iribas C, Barcos Urtiaga A, Ardanaz E, Ederra Sanz M, Castilla J, et al. Changes in breast cancer mortality in Navarre (Spain) after introduction of a screening programme. *J Med Screen* 2007;14:14-20.
17 Hofvind S, Vacek PM, Skelly J, Weaver DL, Geller BM. Comparing screening mammography for early breast cancer detection in Vermont and Norway. *J Natl Cancer Inst* 2008;100:1082-91.
18 Larsen IK, Småstuen M, Johannesen TB, Langmark F, Parkin DM, Bray F, et al. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. *Eur J Cancer* 2009;45:1218-31.
19 Plevritis SK, Sigal BM, Salzman P, Rosenberg J, Glynn P. A stochastic simulation model of U.S. breast cancer mortality trends from 1975 to 2000. *J Natl Cancer Inst Monogr* 2006;(36):86-95.
20 Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* 2005;353:1784-92.
21 CISNET. Cancer Intervention and Surveillance Modeling Network. 2012. http://cisnet.cancer.gov/.
22 R. The R Project for Statistical Computing. www.r-project.org.
23 Lynge E, Braaten T, Njor SH, Olsen AH, Kumle M, Waaseth M, et al. Mammography activity in Norway 1983 to 2008. *Acta Oncol* 2011;50:1062-7.
24 Sorum R, Hofvind S, Skaane P, Haldorsen T. Trends in incidence of ductal carcinoma in situ: the effect of a population-based screening programme. *Breast* 2010;19:499-505.
25 Weedon-Fekjaer H, Bakken K, Vatten LJ, Tretli S. Understanding recent trends in incidence of invasive breast cancer in Norway: age-period-cohort analysis based on registry data on mammography screening and hormone treatment use. *BMJ* 2012;344:e299.
26 Hofvind S, Ursin G, Tretli S, Sebuodegard S, Moller B. Breast cancer mortality in participants of the Norwegian Breast Cancer Screening Program. *Cancer* 2013;119:3106-12.
27 Falk RS, Hofvind S, Skaane P, Haldorsen T. Overdiagnosis among women attending a population-based mammography screening program. *Int J Cancer* 2013;133:705-12.
28 Hakama M, Pukkala E, Heikkila M, Kallio M. Effectiveness of the public health policy for breast cancer screening in Finland: population based cohort study. *BMJ* 1997;314:864-7.
29 Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet* 2012;380:1778-86.
30 Engholm G, Ferlay J, Christensen N, Bray F, Gjerstorff ML, Klint A, et al. NORDCAN—a Nordic tool for cancer information, planning, quality control and research. *Acta Oncol* 2010;49:725-36.
31 Paci E, Group EW. Summary of the evidence of breast cancer service screening outcomes in Europe and first estimate of the benefit and harm balance sheet. J Med Screen 2012;19(Suppl 1):5-13.
32 Olsen AH, Njor SH, Lynge E. Estimating the benefits of mammography screening: the impact of study design. *Epidemiology* 2007;18:487-92.

Accepted: 19 May 2014

Cite this as: *BMJ* 2014;348:g3701

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

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

*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 2| **Breast cancer mortality rate ratios associated with invitations to mammography screening programme in relation to screening period**

Screening period	Mortality rate ratio (95% CI)
During active (biennial) screening period of programme (age 50-69)	0.70 (0.62 to 0.78)
During first five years after invitations to screening ended	0.77 (0.64 to 0.89)
5-10 years after invitations to screening ended	0.79 (0.57 to 1.01)

Table 3| Breast cancer mortality rate ratios associated with invitations to mammography screening programme in alternative (sensitivity) analyses under different statistical assumptions

Statistical assumptions	Mortality rate ratio (95% CI)
Main estimate (from table 1)	0.72 (0.64 to 0.79)
Alternative analyses:	
Screening effect weighted by time since first or last screening*	0.72 (0.65 to 0. 80)
Model without birth cohort effects	0.75 (0.67 to 0.82)
Model without smoothing of period effects	0.72 (0.64 to 0.79)
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	0.75 (0.67 to 0.80)

*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.