Acta Oncologica, 2005; 44: 120–128

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

A model for determining the effect of mammography service screening

ANNE HELENE OLSEN[1], SISSE HELLE NJOR[1], ILSE VEJBORG[2], WALTER SCHWARTZ[3], PETER DALGAARD[1], MAJ-BRITT JENSEN[4], ULLA BRIX TANGE[5], MOGENS BLICHERT-TOFT[4], FRITZ RANK[6], HENNING MOURIDSEN[5] & ELSEBETH LYNGE[1]

[1]*Institute of Public Health, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen N,* [2]*Centre of Diagnostic Imaging, University Hospital Copenhagen, Copenhagen Ø,* [3]*Mammography Screening Clinic, University Hospital Odense, Odense,* [4]*Danish Breast Cancer Cooperative Group, University Hospital Copenhagen, Copenhagen N,* [5]*Department of Oncology, University Hospital Copenhagen, Copenhagen Ø,* [6]*Department of Pathology, University Hospital Copenhagen, Copenhagen Ø, Denmark*

Abstract

In an overview, Swedish mammography screening trials demonstrated a 29% reduction in breast cancer mortality in women aged 50–69 and no effect on total mortality. Three Danish regions introduced mammography screening in the 1990s. The authors developed a method to evaluate the effect of mammography service screening on breast cancer mortality and validated it applying it to total mortality, where no effect was expected. Study groups and historical and national control groups were analysed using Poisson regression. Total mortality in the screening regions and periods was close to that expected in the absence of screening. A small (3%) excess risk, also seen in a non-screening area, probably results from regional differences changing over time. The effect of this is inseparable from the screening effect. The design is adequate for analysing the effect of mammography service screening on breast cancer mortality.

Introduction

Nine randomized controlled trials have evaluated the effect of mammography screening on breast cancer mortality. The results of these trials are not fully coherent, but most studies indicate a reduction in breast cancer mortality.

Based on these results, organized population-based mammography service screening was introduced in three regions in Denmark in 1991, 1993, and 1994. The first programme started in the municipality of Copenhagen, the second in the county of Fyn, and the third in the small municipality of Frederiksberg. Adaptation of trial results to routine healthcare is not straightforward, and it is therefore important to examine whether these service screening programmes achieve a reduction in breast cancer mortality similar to that seen in the randomized trials.

In the overview of the five Swedish randomized trials, a reduction of 29% was found in breast cancer mortality in women aged 50–69 at randomization after a follow-up of on average nine years [1], but no effect was seen on total mortality [2]. This is not surprising, as breast cancer deaths contributed a relatively small proportion of deaths in these age groups. In 1990, 6% of deaths among Danish women aged 50–79 were due to breast cancer. With an effect of 29% on breast cancer mortality, we would therefore expect a decrease of 1–2% in total mortality. Such a small decrease is difficult to detect unless very large populations are studied.

Taking advantage of the unique registers of population and health in Denmark, we have developed a method to evaluate the effect of mammography service screening on breast cancer mortality. As no detectable effect of screening is expected on total mortality, we have validated the evaluation model by applying it to this end point. The point is to see how well the model predicts the expected lack of a detectable effect of screening on total mortality,

Correspondence: Anne Helene Olsen, Department of Epidemiology, Institute of Public Health, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark. Fax (+45) 35 32 73 83. E-mail a.h.olsen@pubhealth.ku.dk

(Received 19 May 2004; accepted 16 December 2004)

ISSN 0284-186X print/ISSN 1651-226X online
DOI: 10.1080/02841860510007576

and to discuss, on this basis, whether the design is adequate for an analysis of the effect of screening on breast cancer mortality.

We present here our method for determining the effect of mammography service screening on breast cancer mortality, and we validate the method by applying it to total mortality as the end point.

Material and methods

Study and control groups

A study group was constructed for each of the three mammography screening programmes in Copenhagen, Fyn, and Frederiksberg. For each study group a historical control group, a national control group, and a historical national control group was constructed.

Study groups

Figure 1 shows the study groups in Lexis diagrams for all three screening regions. The analyses were performed according to the intention to treat principle. Therefore, once a woman was invited to screening in one region, she remained in the study group for that region even if she moved to another region in Denmark.

All women were followed up from their first date of invitation until death, emigration, or 30 June 2002. Women with prevalent breast cancer before their first invitation date were excluded.

Copenhagen. In the municipality of Copenhagen, organized mammography screening started on 1 April 1991. The screening interval was two years. The approximately 40 000 women aged 50–69 at the start of each invitation round were targeted by the programme. The first invitation round targeted women aged 50–69 on 1 April 1991. The second invitation round targeted women aged 50–71 on 1 April 1993. This practice of inviting women above the age of 69 did not continue in the subsequent rounds.

Women were invited according to their date of birth, that is, women born on 1 January were invited first, subsequently women born on 2 January were invited etc., irrespective of their year of birth. Women moving to Copenhagen were invited shortly after their arrival, unless their date of birth was scheduled for invitation later in the round. Women were not invited if they moved out of Copenhagen before their date of birth was invited. The practice of inviting women aged 50–69 at the beginning of each invitation round means that women turning 70 before their invitation date were not excluded,

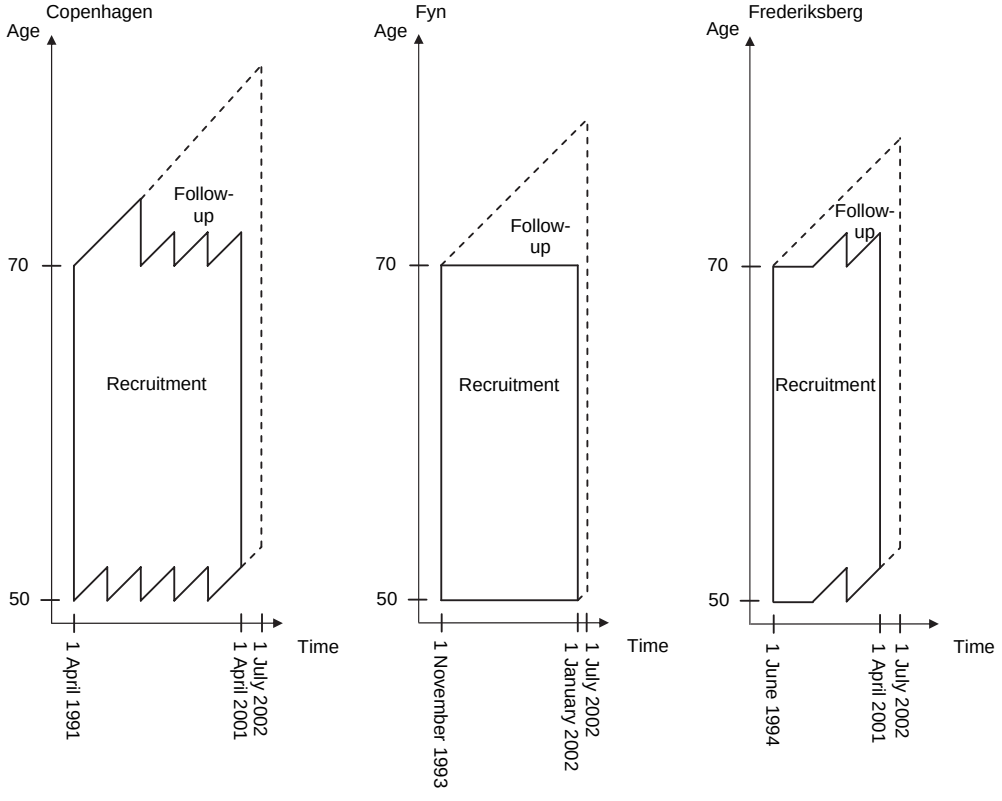

Figure 1. Recruitment and follow-up for the study groups in Lexis diagrams.

whereas women turning 50 after the start of the invitation round were not included.

Women in the study group for Copenhagen were recruited from the first five invitation rounds undertaken from 1 April 1991 to 31 March 2001.

Fyn. In the county of Fyn, organized mammography screening started on 1 November 1993. The screening interval was two years. The approximately 50 000 women aged 50–69 and living in the county of Fyn on their invitation date were targeted by the programme. Invitations were issued centrally but the date depended on the address of the general practitioner with whom the woman was registered. Women in the study group for Fyn were recruited from the first four invitation rounds performed in the period from 1 November 1993 to 31 December 2001.

Frederiksberg. In the municipality of Frederiksberg, organized mammography screening started on 1 June 1994. The screening interval was two years with some irregularities in the first three invitation rounds due to the merge with the Copenhagen programme in 1996. In the first invitation round, women born between 1 July 1924 and 31 December 1945 were invited, if they were between 50 and 69 on their invitation date. From the second invitation round onwards, the target group was the approximately 10 000 women aged 50–69 at the start of each invitation round.

As in Copenhagen, women were invited according to their date of birth. In the first invitation round, however, women moving out of Frederiksberg or turning 70 before their invitation date, and women moving to Frederiksberg or turning 50 after their invitation date, were not invited.

Women in the study group for Frederiksberg were recruited from the first three invitation rounds performed from 1 June 1994 to 31 March 2001.

Historical control groups (Table I)

The historical control groups consisted of women living in the screening regions before the start of the screening programmes. Pseudo-invitation rounds of the same length as the invitation rounds in the study groups were constructed, and pseudo-invitation dates were allocated.

Table I. Study and control groups.

	Screening region	Denmark outside screening regions
Pre-screening period	Historical control group	Historical national control group
Screening period	Study group	National control group

Women were included in the historical control group for one region only, defined as the first region with a pseudo-invitation date in the historical control group. They remained in it even if they moved to another region in Denmark. This was done to ensure an exit and entry system between the historical control groups similar to that between the study groups. In- and exclusion in the historical control groups was as for the study groups.

Women were followed up from their first pseudo-invitation date until death, emigration, or 31 March 1991. Women with prevalent breast cancer before their first pseudo-invitation date were excluded.

Copenhagen. For Copenhagen five pseudo-invitation rounds were constructed starting on 31 December 1979 and ending on 30 December 1989. The invitation system in Copenhagen could be directly reconstructed for the historical control group.

Fyn. Four pseudo-invitation rounds were constructed starting on 2 August 1982 and ending on 1 October 1990. Since the invitation dates for the study group in Fyn were allocated by general practitioner, it was not possible to reconstruct this system for the control group. Instead, women in the control groups were assigned a pseudo-invitation date the first time they were eligible for an invitation round. They were regarded as eligible if at some point in the invitation round they were living in Fyn and were between age 50 and 69. They were then assigned a pseudo-invitation date, selected randomly from all days in the pseudo-invitation round. Women turning 70 or moving out of Fyn before their pseudo-invitation date were excluded from the invitation round, and women turning 50 or moving into Fyn after their pseudo-invitation date were not included in the invitation round.

It would have been possible to end the follow-up on 31 October 1993 but we chose not to do so, since this would have resulted in exclusion from the historical control group of women entering the Copenhagen programme that started on 1 April 1991.

Frederiksberg. Three pseudo-invitation rounds were constructed for Frederiksberg starting on 2 March 1983 and ending on 30 December 1989. The invitation system in Frederiksberg was reconstructed for the historical control group.

National control groups (Table I)

The national control groups consisted of women living in Denmark outside the three screening regions. Pseudo-invitation rounds were constructed for the same periods as the invitation rounds in the study groups. The invitation systems for the national control groups were constructed in the same way as for the historical control groups. The women were followed up from their first pseudo-invitation date until death, emigration, or 30 June 2002. Women with prevalent breast cancer before their first pseudo-invitation date were excluded.

The national control groups were merged with the study groups for Copenhagen, Fyn, and Frederiksberg, and women were excluded from the national control groups from their first invitation date in one of the study groups.

Historical national control groups (Table I)

The historical national control groups consisted of women living in Denmark outside the three screening regions in the period before start of the screening programmes. Pseudo-invitation rounds were constructed for the same periods as the pseudo-invitation rounds in the historical control groups. The invitation systems for the historical national control group were constructed in the same way as for the historical control groups. The women were followed up from their first pseudo-invitation date until death, emigration, or 31 March 1991. Women with prevalent breast cancer before their first pseudo-invitation date were excluded.

The historical national control groups were merged with the historical control groups for Copenhagen, Fyn, and Frederiksberg, and women were excluded from the historical national control groups from their first pseudo-invitation date in one of the historical control groups. This was done to ensure a similar exclusion system in the historical national control group as in the national control group.

Data

Data on invited women in Copenhagen and Frederiksberg were retrieved from the mammography screening register at Kommunedata. Data on invited women in Fyn were retrieved from the registers of the mammography screening clinic for Fyn. Data on women in the control groups were retrieved from the Central Population Register.

The Central Population Register was used for follow-up for death and emigration. Retrieval of data from the register was performed on 29 November 2002. To secure that the data used were as complete as possible, we chose to end follow-up on 30 June 2002.

Women with prevalent breast cancer at their invitation date or pseudo-invitation date were identified from the Danish Cancer Register and the register of the Danish Breast Cancer Cooperative Group.

Linkage between registers was performed by use of the personal identification number issued to all residents of Denmark.

Model

To analyse the effect of screening, we compared total mortality rates in the study groups with rates in the control groups, adjusting for age, period, and region.

Analyses were performed using Poisson regression. Relative risks were adjusted for age by including five-year age group but this was, for simplicity, not included in the model below.

In a full Poisson regression model with the variables exposure (invitation to screening) e, period p and region r, the mortality rate λ_{epr} can be written as a product of main and interaction factors:

$$\lambda_{epr} = \theta \times (\theta_{eksp})^{e} \times (\theta_{per})^{p} \times (\theta_{reg})^{r} \times (\theta_{eksp,\,per})^{ep} \times (\theta_{per,\,reg})^{pr} \times (\theta_{eksp,\,reg})^{er} \times (\theta_{eksp,\,per,\,reg})^{epr}$$

where the θs indicate the main and interaction factors of *eksp*, exposure, *per*, period, and *reg*, region, and with

$e = 1$ for exposure, otherwise $e = 0$,

$p = 1$ for the period before screening, otherwise $p = 0$,

$r = 1$ for the regions outside the three screening regions, otherwise $r = 0$.

In our case, we have screening only in one region in one period. The estimate of interest is the effect of exposure in the screening region in the screening period compared with no exposure in the screening region in the screening period, that is,

$$\frac{\lambda_{e=1,\,p=0,\,r=0}}{\lambda_{e=0,\,p=0,\,r=0}} = \theta_{eksp}.$$

An estimate of $\lambda_{e=0,p=0,r=0}$ is not directly available. It is, however, possible to estimate directly:

The mortality rates of the study groups:

$$\lambda_{e=1,\,p=0,\,r=0},$$

the mortality rates of the national control groups:

$$\lambda_{e=0,\,p=0,\,r=1},$$

the mortality rates of the historical control groups:

$$\lambda_{e=0,\,p=1,\,r=0},$$

and the mortality rates of the historical national control groups:

$$\lambda_{e=0,\,p=1,\,r=1}.$$

These give the following relative risks:
Study group rate/national control group rate =

$$\frac{\lambda_{e=1,\,p=0,\,r=0}}{\lambda_{e=0,\,p=0,\,r=1}} = \frac{\theta_{eksp}}{\theta_{reg}},\ \text{and,}$$

Historical control group rate/historical national control group rate =

$$\frac{\lambda_{e=0,\,p=1,\,r=0}}{\lambda_{e=0,\,p=1,\,r=1}} = \frac{1}{\theta_{reg} \times \theta_{per,\,reg}},$$

and, therefore, the ratio of the two relative risks is

$$\theta_{eksp} \times \theta_{per,reg}.$$

It is not possible to estimate these two factors independently. The estimated exposure effect in this model is therefore confounded by the interaction term between region and period.

Estimates were calculated using a Poisson regression model with the parameters five-year age group a, exposure e, period p, and region r.

Since trends in total mortality were stable and similar in Denmark and Copenhagen in the periods observed [3], we did not take mortality trends within the two periods into account.

Results

In Copenhagen, the study group had a higher mortality than the national control group, in total relative risk 1.26 (95% confidence interval: 1.23, 1.29), although there were minor variations by age group. The same was true for the period before screening started, relative risk 1.22 (95% confidence interval: 1.20, 1.25), whereas mortality rates in the study group compared with the historical control group, and in the national control group compared with the historical national control group, were similar, with relative risks of 1.02 (95% confidence interval: 1.00, 1.05) and 0.99 (95% confidence interval: 0.98, 1.00) respectively. When estimating the effect of the combination of screening and the interaction term between period and region adjusted for age, period, and region the relative risk was 1.03 (95% confidence interval: 1.00, 1.06) (exact lower limit 1.00036, borderline p-value 0.0473) (Table II). In Fyn, the study group had a mortality equal to that of the national control group, relative risk 0.98 (95% confidence interval: 0.95, 1.01), whereas Fyn was slightly below the national control group in the period before screening, relative risk 0.95 (95% confidence interval: 0.92, 0.98). When estimating the effect of the combination of screening and the interaction term between period and region adjusted for age, period, and region the relative risk was 1.03 (95% confidence interval: 0.99, 1.08) (Table III). The results for the small municipality of Frederiksberg resembled those of Copenhagen, and estimating the effect of the combination of screening and the interaction term between period and region adjusted for age, period, and region resulted in a relative risk of 1.07 (95% confidence interval: 0.98, 1.17) (Table IV).

Discussion

Our model showed that the total mortality in the screening regions was close to what we would have expected in the absence of screening. We would expect the mammography screening to have no detectable effect on the overall mortality, as the Swedish randomized trials had no detectable effect on total mortality [1,2]. The fact that we found a small increased effect on overall mortality, borderline significant for Copenhagen, non-significant for Fyn and Frederiksberg, is most probably a result of the interaction between period and region. It should be remembered, however, that our study is an observational one, although with a quasi-experimental design. We therefore have to consider also the possible biases.

First, we excluded women with prevalent breast cancer before their first date of invitation/pseudo-invitation. This is important in an analysis of the effect of screening on breast cancer mortality, since the course of breast cancers diagnosed before screening cannot be affected by screening. This exclusion is not, however, necessary or even reasonable when the end point is total mortality. Nevertheless we made this exclusion because (1) we are developing an evaluation model for breast cancer mortality, and (2) the service programmes have, to some extent, excluded prevalent breast cancers from their invitations. In Copenhagen this practice was, however, abandoned after a short while. We therefore performed the total mortality analysis for Copenhagen without excluding prevalent breast cancers (data not shown). This did not change the estimate of the combined effect of exposure and interaction. It was not possible to perform a similar analysis for Fyn, since the majority of women with prevalent breast cancer were not invited.

Second, women can be transferred from the national control group to the study group and not vice versa. This could introduce a potential difference in the age distribution of the person years in the national control group compared with the study group. Despite our adjustments for five-year age groups, it could affect the age distribution within each five-year age group. Also, migrants tend to be healthier. This would, however, also affect the comparison between the historical and the historical national control groups.

Table II. Effect estimates for total mortality in the Copenhagen mammography screening program.

	50–54	55–59	60–64	65–69	70–74	75–79	80–84	All age groups
Study group (SG)								
No. of deaths	426	906	1 368	2 139	2 856	1 737	96	9 528
Person years	72 002	109 374	96 112	98 184	87 924	36 587	1 442	501 625
Mortality rate per 100 000 person years	592	828	1 423	2 179	3 248	4 748	6 657	1 899
National control group (NCG)[1]								
No. of deaths	3 351	7 652	11 163	16 588	19 395	11 424	559	70 132
Person years	859 934	1 255 569	1 045 410	942 245	737 579	284 150	10 498	5 135 385
Mortality rate per 100 000 person years	390	609	1 068	1 760	2 630	4 020	5 325	1 366
Historical control group (HCG)								
No. of deaths	493	1 310	2 396	3 685	4 605	2 796	131	15 416
Person years	64 576	133 883	165 170	182 951	154 428	60 640	2 165	763 813
Mortality rate per 100 000 person years	763	978	1 451	2 014	2 982	4 611	6 051	2 018
Historical national control group (HNCG)[2]								
No. of deaths	3 114	7 798	11 346	15 593	18 122	10 911	532	67 416
Person years	625 488	1 062 062	1 039 971	949 755	709 851	269 605	9 570	4 666 302
Mortality rate per 100 000 person years	498	734	1 091	1 642	2 553	4 047	5 559	1 445
Relative risk estimates (95% confidence interval)								
SG[3] compared with NCG[1]	1.52	1.36	1.33	1.24	1.24	1.18	1.25	1.26 (1.23 1.29)
HCG[4] compared with HNCG[2]	1.53	1.33	1.33	1.23	1.17	1.14	1.09	1.22 (1.20 1.25)
SG[3] compared with HCG[4]	0.77	0.85	0.98	1.08	1.09	1.03	1.10	1.02 (1.00 1.05)
NCG[1] compared with HNCG[2]	0.78	0.83	0.98	1.07	1.03	0.99	0.96	0.99 (0.98 1.00)
Combined effect of invitation to screening and interaction between region and period	0.99	1.02	1.00	1.01	1.06	1.04	1.15	1.03 (1.00 1.06)

[1]National control group (Denmark except Copenhagen Fyn and Frederiksberg); [2]historical national control group (Denmark except Copenhagen Fyn and Frederiksberg); [3]study group; [4]historical control group.

A relative risk larger than 1 due to interaction between period and region reflects that the mortality trend in the region observed was less advantageous than the mortality trend in the rest of Denmark. The mortality may have increased relatively more or decreased less in the region observed than in the rest of Denmark.

Confounding by the interaction between period and region may explain the result for Copenhagen, if the total mortality in Copenhagen compared with that of the rest of Denmark would have been higher in the screening period than in the period before, given that screening had not been introduced. Such an outcome is reasonable if the Copenhagen population had become relatively more vulnerable in the 1990s than it was in the 1980s. In the age group 50–69, the Copenhagen population decreased from constituting 11% of the Danish population in the 1980s to 7% in the 1990s and was therefore potentially more selected. The interaction term for Fyn would indicate that Fyn had become relatively more vulnerable in the 1990s than in the 1980s.

To examine whether the size of the interaction term was reasonable, we performed the analyses for one of the counties with no screening programme. The county of Vejle was chosen because it has a population fairly similar to that of Fyn. The estimated overall interaction between period and region was 1.06 (95% confidence interval: 1.02, 1.10) although the age group-specific effects varied. We therefore consider the interaction between region and period to be the most likely explanation for the small effects on the overall mortality found in our model. This type of problem is unavoidable in observational studies.

A number of earlier studies have looked at the effect of mammography service screening on breast cancer mortality using various methodologies.

Otto et al. compared mortality trends in the screening period with those in the pre-screening period and evaluated the turning point of the trends [4]. They were not able to exclude breast cancers diagnosed before start of screening, and were aware that this diluted their estimated effect of screening. The turning point was observed around the introduction of screening, and they argued that adjuvant systemic therapy was unlikely to be the cause, since mortality rates continued to rise up to one year after

Table III. Effect estimates for total mortality in the Fyn mammography screening program.

	50–54	55–59	60–64	65–69	70–74	75–79	All age groups
Study group (SG)							
No. of deaths	347	628	891	1 335	1 180	195	4 576
Person years	98 524	103 381	86 836	78 675	48 365	5 754	421 535
Mortality rate per 100 000 person years	352	607	1 026	1 697	2 440	3 389	1 086
National control group (NCG)[1]							
No. of deaths	3 331	5 707	8 051	12 142	10 836	1 944	42 011
Person years	909 667	971 886	787 903	693 899	427 138	54 289	3 844 782
Mortality rate per 100 000 person years	366	587	1 022	1 750	2 537	3 581	1 093
Historical control group (HCG)							
No. of deaths	315	692	983	1 353	1 205	232	4 780
Person years	70 491	89 194	90 727	87 832	53 550	6 675	398 469
Mortality rate per 100 000 person years	447	776	1 083	1 540	2 250	3 476	1 200
Historical national control group (HNCG)[2]							
No. of deaths	3 079	5 858	8 502	11 773	10 454	1 975	41 641
Person years	645 776	785 876	775 558	715 267	417 427	52 542	3 392 446
Mortality rate per 100 000 person years	4770	745	1 096	1 646	2 504	3 759	1 227
Relative risk estimates (95% confidence interval)							
SG[3] compared with NCG[1]	0.96	1.03	1.00	0.97	0.96	0.95	0.98(0.95 1.01)
HCG[4] compared with HNCG[2]	0.94	1.04	0.99	0.94	0.90	0.92	0.95(0.92 0.98)
SG[3] compared with HCG[4]	0.79	0.78	0.95	1.10	1.08	0.98	0.99(0.95 1.03)
NCG[1] compared with HNCG[2]	0.77	0.79	0.93	1.06	1.01	0.95	0.95(0.94 0.97)
Combined effect of invitation to screening and interaction between region and period	1.03	0.99	1.02	1.04	1.07	1.02	1.03(0.99 1.08)

[1]National control group (Denmark except Copenhagen Fyn and Frederiksberg); [2]historical national control group (Denmark except Copenhagen Fyn and Frederiksberg); [3]study group; [4]historical control group.

implementation in regions where screening began after 1995. Other changes that may have affected breast cancer mortality in the recent period could not, however, be taken into account.

Blanks et al. evaluated the effect of the NHS breast screening programme in England and Wales on breast cancer mortality by comparing observed mortality in the screening period with that predicted for a situation without screening [5]. They were also unable to exclude breast cancers diagnosed before start of screening, and they pointed out that this diluted their estimated effect of screening. They also acknowledged that the estimates included effects of screening as well as of other factors such as changes in treatment. They pointed out that their predictions were impossible to test.

Duffy et al. estimated the effect on breast cancer mortality of invitation to screening and of participation in screening in seven Swedish counties [6]. They selected a starting date for each county with as little screening as possible taking place before it and high coverage as soon as possible after that date. Breast cancers diagnosed before that date were excluded. The screening period was compared with a pre-screening period of the same length. The influence of other factors than screening was taken into account in several ways. One was to take into account the changes in mortality rates over time in the pre-screening period. Duffy et al. acknowledged the limitations in this, since it may not reflect factors that could, more recently, have affected breast cancer mortality rates. They took account of this by also comparing the pre-screening period with the non-attenders, correcting for self-selection bias estimated for the randomized trials, and pointed out that this depended on an assumption that the self-selection bias was of the same magnitude in the service-screening programmes.

Jonsson et al. compared counties in Sweden which started screening in 1986–87 with counties which started in 1993 or later [7]. They too were unable to avoid some contamination of breast cancers diagnosed between start of the screening programmes and the individual invitations, but estimated the magnitude of the problem and adjusted their estimates accordingly. To adjust for geographical differences between the study and control groups, a reference period was used. This was based on an assumption of multiplicative effects between regions and periods (no interaction).

We were able to reconstruct the invitation system of the study groups in the control groups. This made it possible to exclude women with prevalent breast cancer before their individual first dates of invitation/pseudo-invitation. Both the deaths and person years

Table IV. Effect estimates for total mortality in the Frederiksberg mammography screening program.

	50–54	55–59	60–64	65–69	70–74	75–79	All age groups
Study group (SG)							
No. of deaths	68	111	199	263	249	33	923
Person years	14 917	19 091	14 881	13 850	9 216	915	72 870
Mortality rate per 100 000 person years	456	581	1 337	1 899	2 702	3 607	1 267
National control group (NCG)[1]							
No. of deaths	2 629	5 255	7 383	11 093	9 549	1 201	37 110
Person years	715 751	906 488	729 877	638 780	376 796	34 393	3 402 085
Mortality rate per 100 000 person years	367	580	1 012	1 737	2 534	3 492	1 091
Historical control group (HCG)							
No. of deaths	57	131	239	357	381	45	1 210
Person years	10 211	16 798	19 874	22 555	13 955	1 236	84 629
Mortality rate per 100 000 person years	558	780	1 203	1 583	2 730	3 641	1 430
Historical national control group (HNCG)[2]							
No. of deaths	2 424	5 368	7 712	10 895	9 054	1 156	36 609
Person years	488 119	719 434	710 108	659 465	363 360	31 358	2 971 844
Mortality rate per 100 000 person years	497	746	1 086	1 652	2 492	3 686	1 232
Relative risk estimates (95% confidence interval)							
SG[3] compared with NCG[1]	1.24	1.00	1.32	1.09	1.07	1.03	1.12(1.05 1.20)
HCG[4] compared with HNCG[2]	1.12	1.05	1.11	0.96	1.10	0.99	1.05(0.99 1.11)
SG[3] compared with HCG[4]	0.82	0.75	1.11	1.20	0.99	0.99	1.03(0.94 1.12)
NCG[1] compared with HNCG[2]	0.74	0.78	0.93	1.05	1.02	0.95	0.95(0.93 0.96)
Combined effect of invitation to screening and interaction between region and period	1.10	0.96	1.19	1.14	0.97	1.05	1.07(0.98 1.17)

[1]National control group (Denmark except Copenhagen Fyn and Frederiksberg); [2]historical national control group (Denmark except Copenhagen Fyn and Frederiksberg); [3]study group; [4]historical control group.

for the women with prevalent breast cancer were excluded.

We were also able to take both time trends and regional differences into account in the estimation of the exposure effect. In only one of the earlier studies mentioned above were both regional and historical control groups used [7] but the study assumed no interaction between period and region. In contrast, we were able to evaluate the magnitude of the interaction between period and region.

The potential bias of different follow-up periods was avoided in our model. By the reconstruction of invitation systems in the control groups, women without prevalent breast cancer in the study and control groups were followed up for a similar period of time. This is important, since the breast cancer mortality rate is not constant over time in a population without prevalent breast cancer on a given date. In the first few years, there will be fewer deaths from breast cancer, since all prevalent cases have been excluded.

The reconstruction of the invitation systems in the control groups also ensured that the study design in itself did not cause different follow-up of incident breast cancers (cancers diagnosed within the follow-up period). As pointed out by Duffy et al. [6], this is important as the risk of death is higher in the first few years after breast cancer diagnosis.

This study is, as expected, not large enough to detect the expected small effect of screening on total mortality. Also, since the expected effect is small, a 1–2% decrease, the interaction between region and period makes it difficult to estimate the effect of screening on total mortality. However, the small magnitude of the interaction indicates that in the period studied there have not been large unsynchronized changes in factors affecting total mortality in the different regions, such as for example socio-economic factors that could also affect breast cancer mortality. The effect of screening on breast cancer mortality is expected to be much larger than that on total mortality. If a reduction of 29% is achieved, as seen in the Swedish overview [1], a small interaction of the magnitude seen in the total mortality study will not result in a misleading estimate of the effect. The small interaction for total mortality is therefore an indication of a potentially well-functioning model for estimating the effect of screening on breast cancer mortality. It does, however, not necessarily imply a small interaction in an analysis of breast cancer mortality. When interpreting the interaction term in the breast cancer analysis, we will of course consider factors specific for breast cancer, such as treatment. It should be noted that changes over time in breast cancer mortality are adjusted for in the model, so only changes unsynchronized between

regions will affect the estimate of the screening effect.

In conclusion, our study design solves some of the biases in earlier studies of the effect of mammography service screening on breast cancer mortality. Mammography service screening introduced in the total population over a two-year period can be evaluated only in a quasi-experimental design, and such a design cannot be completely balanced. The excess total mortality in our study groups is small and is most probably due to an interaction between region and period. We therefore consider the study design to be adequate for an analysis of the effect of mammography service screening programmes on breast cancer mortality.

For further validation, application of this model to screening data from randomized controlled trials, with control groups constructed similarly to those presented here, could be considered.

Acknowledgements

This study was financially supported by the Danish Medical Research Council, the Centre for Evaluation and Medical Technology Assessment in the Danish National Board of Health, and the EU-Commission, DG Sanco.

References

[1] Nyström L, Rutqvist LE, Wall S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. Lancet 1993;341:973–8.

[2] Nyström L, Larsson L-G, Wall S, et al. An overview of the Swedish randomised mammography trials: total mortality pattern and the representivity of the study cohorts. J Med Screen 1996;3:85–7.

[3] StatBank Denmark. Available at: URL:http://www.statistikbanken.dk (accessed 7 September 2004).

[4] Otto S J, Fracheboud J, Looman C W, 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.

[5] Blanks R G, Moss S M, McGahan C E, et al. Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990–8: comparison of observed with predicted mortality. Br Med J 2000;321:665–9.

[6] Duffy S W, Tabar L, Chen H H, et al. The impact of organized mammography service screening on breast carcinoma mortality in seven Swedish counties. Cancer 2002;95:458–69.

[7] Jonsson H, Nyström L, Törnberg S, et al. Service screening with mammography of women aged 50–69 years in Sweden: effects on mortality from breast cancer. J Med Screen 2001;8:152–60.