FULL PAPER

BJC

British Journal of Cancer (2016) 114, 1038–1044 | doi: 10.1038/bjc.2016.68

Keywords: screening; mortality; case–control study; cohort study; long-term effects; selection bias

Impact of organised mammography screening on breast cancer mortality in a case–control and cohort study

Sirpa Heinävaara*[,1], Tytti Sarkeala[1] and Ahti Anttila[1]

[1]Finnish Cancer Registry/Mass Screening Registry, Unioninkatu 22, 00130 Helsinki, Finland

Background: The usefulness of case–control studies has been questioned. Our aim was to evaluate the long-term effect of screening on breast cancer mortality within the population-based mammography programme in Finland using a case–control design, and to compare the analyses with the earlier cohort study.

Methods: The cases were women invited to screening, diagnosed and died from breast cancer in 1992–2011 while being 50–84 years at death. We chose 10 controls for each case with non-restrictive eligibility criteria. Our data included 1907 cases and 18 978 matched controls. We analysed associations between the screening participation and the risk of breast cancer death using the conditional Cox proportional hazards model. The effect estimates were corrected for self-selection bias.

Results: An overall effect of screening was 0.67 (95% confidence interval (CI): 0.49–0.90), and that remained unchanged over time. Analyses with matching criteria comparable to the cohort study yielded an effect (0.70, 95% CI: 0.49–1.00) in 1992–2003 similar to that of the previous cohort analysis (0.72, 95% CI: 0.56–0.88).

Conclusions: Organised mammography screening decreases mortality from breast cancer by 33% among the participants. If made comparable, a case–cohort study can yield effect estimates similar to a cohort study.

The purpose of mammography screening is to decrease mortality from breast cancer. Screening has been shown to be effective in the majority of studies (e.g. Olsen *et al*, 2005; Swedish Organised Service Screening Evaluation Group, 2006; Ascunce *et al*, 2007; Allgood *et al*, 2008; Puliti *et al*, 2008; Sarkeala *et al*, 2008a; Paap *et al*, 2010; van Schoor *et al*, 2011; Nickson *et al*, 2012; Otto *et al*, 2012; Hofvind *et al*, 2013), but not in all (Paci *et al*, 2002; Fielder *et al*, 2004; Gabe *et al*, 2007; Kalager *et al*, 2010). However, decreasing breast cancer mortality trends, particularly below the target age of screening, indicate improved medical services and treatment (Autier *et al*, 2010). Thus, the magnitude of the screening effect can be challenged.

The effectiveness of screening is shown within an organised screening programme. However, if all women in the target age group have been invited to screening for years, the assessment of breast cancer mortality among the non-invited will be based on historical data, probably decades before. Subsequently, the estimation of the effectiveness will be challenging even if such cohort studies would otherwise be feasible with individual-level follow-up data on invitations, participation, breast cancer diagnoses and deaths. In reality, they often are not, and case–control studies are used as the next best alternative in assessing the effect of screening. However, as case–control studies have generally resulted in stronger effects of screening than cohort studies (Broeders *et al*, 2012), the validity of these studies have been questioned (Lauby-Secretan *et al*, 2015). The assessment of possible factors affecting the effect of screening in a case–control study can bring further understanding on the issue.

A long-standing, organised screening programme enables the estimation of a long-term screening effect. As screening may affect breast cancer incidence after the last screening round (Seppänen *et al*, 2006; Heinävaara *et al*, 2014), an analysis including postinvitation ages reflects the true effectiveness of screening. In a case–control study, inclusion of a long study period and a wide

*Correspondence: Dr S Heinävaara; E-mail: sirpa.heinavaara@cancer.fi

Received 28 October 2015; revised 19 February 2016; accepted 25 February 2016; published online 24 March 2016

age range raises concern on the eligibility criteria of controls. It has often been assumed that controls should be free of breast cancer at the cases' date of diagnosis and alive at the cases' date of death (Dubin *et al*, 1987; Weiss, 1994). Moreover, if these assumptions are made for a long time period, controls are unlikely to represent a truly random sample of the general female population. Therefore, special attention should be paid to the selection and handling of controls.

The aim of this study is to evaluate the long-term effect of organised mammography screening on incidence-based breast cancer mortality in Finland in 1992–2011 among 50–84-year-old women using a matched case–control design with non-restrictive eligibility criteria of controls. The current study covers the period, age group and area of the earlier cohort study (Sarkeala *et al*, 2008a). Subsequently, the aim is also to study whether matching criteria or any other design issues cause differences in effects between the cohort and case–control studies by comparing our case–control effect estimates with that of the cohort study for 50–69-year-old women in 1992–2003.

MATERIALS AND METHODS

The Finnish breast cancer screening programme. The population-based screening program for breast cancer was initiated in Finland in 1987 with a group-randomised design (Hakama *et al*, 1997). Since 1992, the programme has covered the entire country, and Finnish municipalities have been entitled to offer a free mammography to 50–59-year-old women every 2 years. In practise, however, the majority of municipalities (86%) have followed varying screening policies since the beginning by also inviting, irregularly or regularly, 60–69-year-old women (Sarkeala *et al*, 2008b). In addition, one municipality has regularly invited 40–74-year-old women. According to the bylaw established in 2007, the invitational age range will be gradually enlarged to 50–69 years by 2017.

All women who belong to an invitational age group and resident in a Finnish municipality with a mailing address are invited to mammography screening with no exclusion criteria. Invitations are sent by personal invitation letters, of which the majority includes prebooked times that can be changed by phone or on the web. Non-participants are sent one reminder letter.

The Mass Screening Registry, a section of the Finnish Cancer Registry, maintains the registration of all invitations to and participation in cancer screening programmes in Finland. Individual level data on screening invitations to and participation in mammography screening has been available from 1992 onwards from screening centres of the Cancer Society of Finland. These screening centres covered some 260 municipalities, ~50% of all activities of the population-based mammography screening in 1992–2004. The coverage of the Mass Screening Registry has improved with time, and it reached complete coverage of all service providers in 2005.

Study population. Potential cases, that is, women who were diagnosed with and who died from breast cancer in Finland between ages 50 and 84 years in 1992–2011 were identified from the Finnish Cancer Registry ($n = 9786$). They were linked with the Mass Screening Registry data with their personal identification number. The screening data have been complete from 1992 onwards in the municipalities screened by the Cancer Society of Finland, the area in the earlier cohort study (Sarkeala *et al*, 2008a), and data were therefore restricted to women living these municipalities. Residential municipality and participation in the screening was assessed from the most recent invitation to screening before the diagnosis of breast cancer, that is, an index invitation. The accumulation of 1911 cases (20% of all breast cancer deaths) with at least one invitation to screening is illustrated in Table 1.

A group of potential controls were drawn for each case from the Mass Screening Registry. They were matched to cases by the year of birth and the year of and the residential municipality at the index invitation ($n = 603\,413$). Matching by residential municipality was used as a surrogate of women's screening history with the number of and ages at invitations. Potential controls were linked with the Finnish Cancer Registry for the diagnoses of cancer, cause of death and dates of emigration and death. Those diagnosed with breast cancer ($n = 16\,691$, 3%), living abroad ($n = 399$) or died ($n = 1750$) before the cases' index invitation date were excluded, leaving us with 584 573 eligible controls. Random numbers were generated for each group of the case's potential controls, and women with the 10 smallest random numbers were chosen as matched controls for each case. Four cases were excluded as they did not have eligible controls. Overall, our data included 1907 cases aged 50–84 years at the death of their breast cancer and 18 978 matched controls. Thirty-eight cases (2%) had <10 controls.

As our study period covers two decades and a wide age range, the time from the index invitation to death can be almost two decades. Usually we would assume that controls must be alive and without breast cancer at the case's diagnosis of breast cancer and alive at case's death date. Now 10 such controls for each case are likely to represent a subsample of long-living individuals rather than a random sample of the general female population. We therefore allowed eligible controls to behave like members of the general female population as closely as possible. It was possible for controls to be diagnosed with breast cancer, emigrate or die before the case's diagnosis of breast cancer. In such situations, controls

Table 1. Accumulation of cases

	Excluded	Included
Diagnosed and died from breast cancer in 1992–2011 between ages 50 and 84 years		9786
Not invited to screening before the diagnosis of breast cancer	4879	
Invited to screening before the diagnosis		4907
Not resident in a Cancer Society of Finland municipality at the index invitation	2613	
Resident in a Cancer Society of Finland municipality at the index invitation		2294
Incomplete data on invitations and participations in screening in 1992 or later	372	
Complete data on invitations and participations in screening 1992 or later		1922
Aged < 50 years at the index invitation	11	
Eligible cases aged at least 50 years at the index invitation		1911
No eligible controls available	4	
Cases with eligible controls		1907

were followed from the date of their index invitation until their earliest date of breast cancer diagnosis, emigration or death. Accordingly, living controls without breast cancer at the date of the case's diagnosis could emigrate or die before the date of case's death. In all these situations, controls exited the follow-up before the case and are considered to be censored. Controls were censored from the follow-up by a diagnosis of breast cancer before the case ($n=201$, 1% of controls), a death from breast cancer after the case's diagnosis of breast cancer ($n=13$), emigration ($n=2$) or death from other causes ($n=1057$, 6% of controls) at any time during the follow-up. A case can be a control to another case, but all breast cancer deaths were included only once as cases in the data.

To compare effects between case–control and cohort designs, the emphasis was on women who had been diagnosed with and died from breast cancer in 1992–2003 between the ages 50 and 69 years. In the earlier cohort study, possible differences between residential municipalities were not taken into account in detail; municipalities were categorised into three recall rate groups that were used in modelling. In the 'Low recall rate' group, the range in recall rates was 0.9–1.9%, in the 'Intermediate recall rate' group 2.3–2.7% and in the 'High recall rate' group 2.8–3.5%. To increase comparability to the earlier study, we formed secondary data by matching controls to cases with respect to the recall rate category ('Low', 'Intermediate', 'High') at index invitation instead of residential municipality, whereas other matching factors and eligibility criteria were kept the same. It is notable that our data on cases for 1992–2003 include 92% of those in the earlier cohort study (Sarkeala *et al*, 2008a). This difference in the number of cases is because of the municipal amalgamations that have led to differences in municipality codes between the data sources: The Finnish Cancer Registry includes only the most current municipality code, whereas the Mass Screening Registry maintains that at the time of an event. Subsequently, as the municipality codes were used from different sources between the two studies, the final number of cases was a little smaller in the case–control data than in the cohort data.

For the assessment of screening history, participation in the screening was defined also at the first invitation. As data on invitations and participation have been reliably available at the Mass Screening Registry from 1992 onwards, the first invitation refers to that in 1992 or later. Thus, it is not truly the first one for those older women who received invitations before 1992.

Statistical analysis. Both data were analysed with the conditional Cox proportional hazards model with death from breast cancer (case–control status) as an outcome and the participation in screening (no/yes) at the index invitation as an explanatory variable. Follow-up time was taken into account from the index invitation date to the earliest of the cases' death date and the date of censoring (exit). The exact index invitation date is not available in the Mass Screening Registry. For the participants, it was replaced by the date of screening, and for the non-participants, it was replaced by an annual median date of those screened in a given municipality. The association between the participation in screening and the risk of breast cancer death was reported with hazard ratios (HRs) and 95% confidence intervals (CIs). To evaluate possible changes with time, a 5-year calendar period of death and index invitation (1992–1996, 1997–2001, 2002–2006, 2007–2011), and the quintiles of follow-up time were used. The heterogeneity of risk in the calendar period and follow-up time was evaluated with interactions and likelihood ratio tests.

For overall estimates of screening, crude HRs were corrected for a self-selection bias by allowing the risk of breast cancer death to differ between screening participants and non-participants (Duffy *et al*, 2002). For this correction, the participation rate in mammography screening and the relative risk of breast cancer death among the non-participants compared with the uninvited women were used. The participation rate for 1992–2011 was calculated to be 0.86 and that for 1992–2003 was 0.87 (The Mass Screening Registry, 2014). For a correction factor, we used the most valid Finnish estimate, 1.56 (95% CI: 1.25–1.91) for 50–69-year-old women in 1992–2003 (Sarkeala *et al*, 2008a).

For descriptive purposes, frequencies were compared between the cases and controls using the χ^2 test. Data were analysed with Stata, version 12 (StataCorp LP, College Station, TX, USA).

RESULTS

The mean (median) age at the index invitation was 58.0 (58.0) years, and <2% ($N=374$) of women were aged 70 years or more. The mean (median) age at the cases' diagnosis of breast cancer was 61.1 (60.0) years and those of death 65.3 (65.0) years. The mean (median) follow-up time from the index invitation date to death was 7.39 (6.50) years, and the mean (median) time from the index invitation date to the cases' diagnosis of breast cancer was 3.16 (1.42) years.

Screening history was assessed with the number of invitations before the index invitation, the year of the first invitation and the change of residential municipality from the first invitation to the index invitation. These indicators did not differ between the cases and controls (Table 2). Further, the pattern of participating in screening was consistent within the study population, as 92% of the cases and 89% of the controls participated similarly (no/yes) at their first and the index invitation.

Overall, the controls participated in screening (86%) more often than the cases (76%), and the participation decreased with age at the index invitation (Table 3).

An overall crude HR of screening was 0.39 (95% CI: 0.34–0.44), and the HR corrected for self-selection bias was 0.67 (95% CI: 0.49–0.90) (Table 4). The effect of screening was slightly larger when the data were restricted to women aged 50–69 years at death (HR = 0.61, 95% CI: 0.45–0.84). The effect of screening was not consistent with the 5-year calendar period of the index invitation ($P=0.0450$), whereas no heterogeneity was observed between the 5-year calendar periods of death ($P=0.6344$). The HR of screening decreased with the period of the index invitation, and the largest

Table 2. Descriptive statistics (*n* and %) for the indicators of screening history

	Cases, *n* (%)	Controls, *n* (%)
Number of invitations before index invitation		
0	611 (32)	6135 (32)
1	431 (23)	4259 (22)
2	327 (17)	3206 (17)
3	276 (14)	2733 (14)
4	155 (8)	1562 (8)
5–9	105 (6)	1055 (6)
10–12	2 (0)	28 (0)
Year of first invitation		
1992	683 (36)	6720 (35)
1993	621 (33)	6237 (33)
1994–1996	259 (14)	2530 (13)
1997–1999	167 (9)	1770 (9)
2000–2009	17 (9)	1721 (9)
Residential municipality at index invitation		
The same as in the first invitation	1812 (95)	18 073 (95)
Changed from the first invitation	95 (5)	905 (5)

Note: The differences in frequencies between the cases and controls: for the number of invitations before index invitations $P=0.999$, for the year of invitation $P=1.000$ and for the change in residential municipality $P=0.678$.

Table 3. Descriptive statistics of study population by the participation in screening at index invitation

	Cases		Controls	
	All, *n*	Participation in screening, *n* (%)	All, *n*	Participation in screening, *n* (%)
All	1907	1451 (76)	18 978	16 347 (86)
Age at death (years)				
50–54	123	95 (77)	1246	1140 (91)
55–59	371	286 (77)	3698	3352 (91)
60–64	433	334 (77)	4310	3878 (90)
65–69	416	325 (78)	4115	3573 (87)
70–74	302	230 (76)	3006	2473 (82)
75–84	182	127 (70)	1810	1366 (75)
Period of death				
1992–1996	117	90 (77)	1160	1021 (88)
1997–2001	361	268 (74)	3581	3111 (87)
2002–2006	585	452 (77)	5833	5018 (86)
2007–2011	844	641 (76)	8404	7197 (86)
Age at index invitation				
50–54	517	426 (82)	5177	4702 (91)
55–59	774	590 (76)	7677	6855 (89)
60–64	431	305 (71)	4289	3345 (78)
65–74	185	130 (70)	1835	1445 (79)
Period of index invitation				
1992–1996	917	704 (77)	9123	7612 (83)
1997–2001	606	461 (76)	6024	5317 (88)
2002–2006	302	236 (78)	3011	2702 (90)
2007–2011	93	59 (63)	930	816 (88)
Follow-up time (years)				
0–2.9	394	276 (70)	3913	3506 (89)
3–4.9	337	265 (79)	3376	3014 (89)
5–7.9	386	297 (77)	3824	3351 (87)
8–11.9	421	330 (78)	4188	3712 (88)
12–19.9	369	283 (77)	3677	2928 (79)

Note: Controls were categorised according to case's follow-up time category, that is, time from index invitation to death.

Table 4. Crude hazard ratios of screening, and hazard ratios corrected for selection bias

	Crude hazard ratio (95% CI)	Corrected hazard ratio (95% CI)
Age at death (years)		
50–84	0.39 (0.34–0.44)	0.67 (0.49–0.90)
50–69	0.36 (0.31–0.41)	0.61 (0.45–0.84)
Period of death		
1992–1996	0.37 (0.22–0.62)	0.63 (0.35–1.14)
1997–2001	0.33 (0.25–0.44)	0.57 (0.38–0.85)
2002–2006	0.42 (0.33–0.53)	0.72 (0.50–1.04)
2007–2011	0.40 (0.33–0.48)	0.68 (0.49–0.94)
Period of index invitation		
1992–1996	0.45 (0.37–0.56)	0.78 (0.55–1.10)
1997–2001	0.35 (0.28–0.44)	0.60 (0.42–0.86)
2002–2006	0.39 (0.29–0.52)	0.66 (0.44–1.00)
2007–2011	0.22 (0.13–0.36)	0.38 (0.21–0.67)
Follow-up time (years)		
0–2.9	0.24 (0.19–0.31)	0.41 (0.28–0.60)
3–4.9	0.42 (0.32–0.57)	0.73 (0.49–1.09)
5–7.9	0.43 (0.32–0.57)	0.73 (0.49–1.09)
8–11.9	0.41 (0.31–0.54)	0.70 (0.47–1.03)
12–19.9	0.63 (0.44–0.90)	1.08 (0.69–1.70)

Abbreviation: CI = confidence interval. *Note*: Controls are categorised according to case's follow-up time category, that is, time from index invitation to death. Heterogeneity, that is, the consistency of HRs with the variable category: for the period of death $P = 0.6344$, for the period of index invitation $P = 0.0450$ and for follow-up time $P = 0.0002$.

mortality effect was observed among those with the most recent 5-year calendar period (2007–2011, corrected HR = 0.38, 95% CI: 0.21–0.67). The HR of screening was strongly dependent on the follow-up time category ($P = 0.0002$), the effect being the largest among those with the shortest follow-up time from the index invitation to death.

To compare effects between case–control and cohort designs, controls were matched to cases with respect to recall rate category while other matching and eligibility criteria were the same as in the primary data. The analysis yielded a corrected HR of screening for 50–69-year-old women in 1992–2003 to be 0.70 (95% CI: 0.49–1.00), which is close to the RR in the earlier cohort study, 0.72 (95% CI: 0.56–0.88) (Table 5). Matching by residential municipality yielded stronger effects of screening than that by recall rate category (Table 5).

For further comparisons, the primary results of corresponding case–control data without censoring are presented Supplementary Appendix Table 1.

DISCUSSION

Organised screening decreases mortality from breast cancer by 33% in women attending screening, and this effect has not changed in Finland in 1992–2011. If a case–control study is made comparable to a cohort study, mortality effects can be analogous.

Our data included almost 2000 breast cancer diagnoses and deaths over two decades, a wide age group (50–84 years) and a large area covering 50% of the screening target population. As the oldest women were invited to screening at the age of 74 years, our study covers a minimum of 10 years of follow-up after the last invitation. In many municipalities, however, women were invited to screening until the age of 69 years, and thus potentially accumulating 15 years of follow-up after the last invitation.

Table 5. The corrected hazard ratios of case–control data, and corrected relative risk of the earlier cohort study with recall rate categories[a]

Study design	Case–control		Cohort
Matching by	Recall rate category	Municipality	
Age at death (years)	Hazard ratio (95% CI)	Hazard ratio (95% CI)	Relative risk (95% CI)
Period of death: 1992–2003			
50–69	0.70 (0.49–1.00)	0.64 (0.45–0.92)	0.72 (0.56–0.88)
Period of death: 1992–2011			
50–69	0.65 (0.48–0.89)	0.61 (0.45–0.84)	NA
50–84	0.73 (0.54–0.99)	0.67 (0.49–0.90)	NA

Abbreviations: CI = confidence interval; NA = not available.
[a]Where controls were matched to cases by recall rate category or by municipality while other matching criteria were kept the same.

Residential municipality at the index invitation was a sufficient surrogate for the screening history. It also represents an overall effect of confounding including underlying breast cancer risk as well as access to health services and opportunistic screening (Aro *et al*, 2001; Pukkala and Patama, 2010). The matching of cases and controls by the year of birth, and by the year of and the residential municipality at the index invitation was thus essential.

The selection of controls towards the general female population was maximised by allowing them to exit from the follow-up after the index invitation because of breast cancer diagnosis, emigration or death. The chosen approach can be useful, especially when the time period between the index invitation and death is long. It also increases resemblance to cohort studies.

The censoring of controls was mainly because of death from other causes than breast cancer. Participation in screening among the censored controls (68%) was lower compared with the overall participation rate of controls (86%). Our study thus confirms that non-participation is associated with an increased risk of dying due to any cause (Jousilahti *et al*, 2005; Dugué *et al*, 2014). Thus, if all controls must have been alive at a case's death, they would have been long-living participants more likely than non-participants. Therefore, *a priori*, censoring will decrease the effect of screening. However, as the censoring of controls was uncommon, our mortality impact of screening was only slightly smaller than that of the case–control data without censoring (see Supplementary Appendix for details). Breast cancer treatment and access to diagnostic services outside the screening programme have improved with time, potentially diminishing differences between various subgroups of women (Autier *et al*, 2010). Therefore, the corrected effects of screening imply a conservative assessment rather than an overestimation of the impact of screening.

The increasing impact of screening towards more recent periods of index invitation is because of the decreasing length of follow-up time (from index invitation to death). In the shortest follow-up time category (0–2.9 years), diagnoses of breast cancer were made on average 6 months after the index invitation, whereas in the longest follow-up time category (12–19 years), breast cancers were diagnosed on average 8 years after the index invitation. The decreasing effect of screening with follow-up time is thus understandable and in line with findings by van der Waal *et al* (2015). Interval-specific effects should therefore be interpreted cautiously.

The effect of screening on breast cancer mortality for 50–69-year-old women in 1992–2003 was compared with the earlier cohort study. There were some minor discrepancies between the data. First, the number of breast cancer deaths in the case–control data was slightly smaller than in the cohort study. Second, in the cohort study screening, the indicator was defined at the first invitation, whereas in the case–control study, it was at the index invitation. However, as women are consistent in their participation in screening, this difference is also of minor importance. Interestingly, the case–control data yielded larger effects of screening when the matching of controls was carried out by the residential municipality rather than by the recall rate category. This seems to indicate a varying effect of screening by screening history and/or municipality. Screening histories also reflect various screening policies conducted in municipalities, and the effect of screening is reported to vary by screening policy (Sarkeala *et al*, 2008b). In addition, women in the most populated, urban municipalities have the lowest attendance in screening (Hemminki *et al*, 2006) as they are likely to attend to a mammogram outside the organised screening programme (Aro *et al*, 2001). Thus, their effect of screening may differ from that in the less populated areas. Overall, the effect of matching by residential area and other criteria should be studied further in future studies.

Our crude effects for screening participation are of the same magnitude than those in other case–control studies (Allgood *et al*, 2008; Paap *et al*, 2010; van Schoor *et al*, 2011; Otto *et al*, 2012), but also a bit smaller than in some others (Fielder *et al*, 2004; Gabe *et al*, 2007; Nickson *et al*, 2012). A study methodologically closest to our study in the Netherlands with 282 cases and 1410 referents, a long study period (1975–2008) of invitations and deaths, and the matching of controls by residential area and age at invitations reported a crude odds ratio 0.35 (95% CI: 0.49–0.87) (van Schoor *et al*, 2011). The study also reported an increasing effect of screening with calendar years of index invitation, which is in line with our findings. As the correction factor for self-selection in the Dutch study is much smaller than in our study, the corrected effect is remarkably larger than our corresponding effect.

Distribution of follow-up times from the index invitation to death has not been reported in case–control studies, which reduces possibilities to compare our results with the previous studies. If participation in screening was assessed only for a short time period before the diagnosis of breast cancer, individual follow-up times tended to be short, and this might explain at least partially the strong effects. Also, the potential influence of matching by residential area in one study is not easily comparable to another study in another country. The reported studies also differ by invitational age group, age group at death, participation rate, the length of the study period and time since the beginning of mammography screening, and thus a conclusive summary is difficult to achieve.

The women who did not participate in screening may have a higher risk of dying from breast cancer than the average population, causing a so-called self-selection bias in participants (Duffy *et al*, 2002; Swedish Organised Service Screening Evaluation Group, 2006; Sarkeala *et al*, 2008a,b). When participants are compared with non-participants, the correct adjustment for self-selection bias is crucial. The participation rate in the Finnish mammography screening has been higher compared with that in the other studied programmes (Lerda *et al*, 2014), and subsequently our correction factor is the highest reported (Fielder *et al*, 2004; Gabe *et al*, 2007; Allgood *et al*, 2008; Puliti *et al*, 2008; Paap

et al, 2010; van Schoor *et al*, 2011; Otto *et al*, 2012). The Finnish correction factor was assessed from the incidence-based mortality data with individual level mortality and follow-up data, and with the emphasis on the comparability of the periods before and after the introduction of the screening programme (Sarkeala *et al*, 2008a). Such detailed data on deaths from incident cases and person-years may not necessarily have been available from the same population in all countries (Fielder *et al*, 2004; Allgood *et al*, 2008; Puliti *et al*, 2008; Paap *et al*, 2011; van Schoor *et al*, 2011; Otto *et al*, 2012). On the other hand, it has been shown for a nested case–control study within a randomised trial that correction factors are likely to vary depending on the definition of screening (first *vs* non-attender, never *vs* ever) and data used for its estimation (case–control or person-years data) (van der Waal *et al*, 2015). However, variation in correction factors was not clearly associated with either of these factors, possibly because of the small sample size, limiting conclusions. It may be that our correction factor from cohort data with attendance at screening at the first invitation is not fully precise for the case–control data with attendance at screening at the index invitation. In any case, extrapolation from one time period to another (i.e., from 1992–2003 to 1992–2011) has introduced uncertainty to corrected HRs as correction factors are likely vary with time as well (van der Waal *et al*, 2015).

In many countries, case–control studies are the best alternative in assessing the effect of screening on breast cancer mortality. They show some evidence (crude odds ratios), which is not directly dependent on breast cancer mortality among the non-invited. The breast cancer mortality rate among the non-invited is still needed for the correction of self-selection, but a possible magnitude of this effect can be elaborated with a sensitivity analysis. Therefore, we cannot ignore case–control studies altogether, but we need to gain understanding in factors resulting in differences in effects between case–control and cohort studies.

Our study demonstrates that organised mammography screening is effective in reducing breast cancer mortality. The effectiveness of screening is, however, strongly affected by the length of follow-up time, that is, the time from the index invitation to death. Case–control studies with long follow-up times, and applying an appropriate matching and eligibility criteria of controls can be valid in assessing effectiveness of population-based screening programmes.

DISCLAIMER

Researchers of the study were independent from Cancer Society of Finland (CFS) as the employee and sole funder.

AVAILABILITY OF MATERIALS AND DATA

Data are available for research projects from the legal administrator of the data, National Institute for Health and Welfare. For permissions of use, send your request to kirjaamo@cancer.fi.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Allgood PC, Warwick J, Warren RM, Day NE, Duffy SW. A case–control study of the impact of the East Anglian breast screening programme on breast cancer mortality (2008) *Br J Cancer* **98**: 206–209.

Aro AR, de Koning HJ, Absetz P, Schreck M (2001) Two distinct groups of non-attenders in an organized mammography screening program. *Breast Cancer Res Treat* **70**(2): 145–153.

Ascunce EN, Moreno-Iribas C, Barcos Urtiaga A, Ardanaz E, Ederra Sanz M, Castilla J, Eqüés N (2007) Changes in breast cancer mortality in Navarre (Spain) after introduction of a screening programme. *J Med Screen* **14**: 14–20.

Autier P, Boniol M, LaVecchia C, Vatten L, Gavin A, Héry C, Heanue M (2010) Disparities in breast cancer mortality trends between European countries: retrospective trend analysis of WHO mortality database. *BMJ* **341**: c3620.

Broeders M, Moss S, Nyström L, Njor S, Jonsson H, Paap E, Massat N, Duffy S, Lynge E. Paci EUROSCREEN Working Group (2012) The impact of mammography screening on breast cancer mortality in Europe: a review of observational studies. *J Med Screen* **19**(Suppl 1): 14–25.

Dubin N, Friedman DR, Toniolo PG, Pasternack BS (1987) Breast cancer detection centers and case–control studies of the efficacy of screening. *J Chronic Dis* **40**: 1041–1050.

Duffy SW, Cuzick J, Tabar L, Vitak B, Hsiu-His Chen T, Yen MF, Smith RA (2002) Correcting for non-compliance bias in case-control studies to evaluate cancer screening programmes. *Appl Statist* **51**(Part 2): 235–243.

Dugué PA, Lynge E, Rebolj M (2014) Mortality of non-participants in cervical screening: Registry-based cohort study. *Int J Cancer* **134**(11): 2674–2682.

Fielder HM, Warwick J, Brook D, Gower-Thomas K, Cuzick J, Monypenny I, Duffy S (2004) A case–control study to estimate the impact on breast cancer death of the breast screening programme in Wales. *J Med Screen* **11**: 194–198.

Gabe R, Tryggvadóttir L, Sigfússon BF, Olafsdóttir GH, Sigurdsson K, Duffy SW (2007) A case–control study to estimate the impact of the Icelandic population-based mammography screening program on breast cancer death. *Acta Radiol* **48**: 948–955.

Hakama M, Pukkala E, Heikkilä M, Kallio M (1997) Effectiveness of the public health policy for breast cancer screening in Finland: Population based cohort study. *BMJ* **314**: 864–867.

Heinävaara S, Sarkeala T, Anttila A (2014) Overdiagnosis due to breast cancer screening: updated estimates of the Helsinki service study in Finland. *Br J Cancer* **111**(7): 1463–1468.

Hemminki E, Sevon T, Tanninen K, Pukkala E, Anttila A (2006) Use of private gynaecologist does not relate better prevention outcomes – an ecological analysis from Finland. *BMC Health Serv Res* **6**: 27.

Hofvind S, Ursin G, Tretli S, Sebudegård S, Møller B (2013) Breast cancer mortality in participants of the Norwegian breast cancer screening program. *Cancer* **119**: 3106–3112.

Jousilahti P, Salomaa V, Kuulasmaa K, Niemelä M, Vartiainen E (2005) Total and cause specific mortality among participants and non-participants of population based health surveys: a comprehensive follow up of 54 372 Finnish men and women. *J Epidemiol Community Health* **59**: 310–315.

Kalager M, Zelen M, Langmark F, Adami HO (2010) Effect of screening mammography on breast-cancer mortality in Norway. *N Engl J Med* **363**: 1203–1210.

Lauby-Secretan B, Scoccianti C, Loomis D, Benbrahim-Tallaa L, Bouvard V, Bianchini F, Straif K. for the International Agency for Research on Cancer Handbook Working Group (2015) Breast-Cancer Screening – viewpoint of the IARC Working Group. *N Engl J Med* **372**: 2353–2358.

Lerda D, Deandrea S, Freeman C, López-Alcalde J, Neamtiu L, Nicholl C, Nicholson N, Ulütuck A, Villanueva S (2014) *Report of a European Survey on the Organisation of Breast Cancer Care Services*. Publications Office of the European Union: Luxembourg; doi:10.2788/51070.

Nickson C, Mason KE, English DR, Kavanagh AM (2012) Mammographic screening and breast cancer mortality: a case–control study and meta-analysis. *Cancer Epidemiol Biomarkers Prev* **21**: 1479–1488.

Olsen A, Njor S, Vejborg I, Schwartz W, Dalgaard P, Jensen MB, Tange UB, Blichert-Toft M, Rank F, Mouridsen H, Lynge E (2005) Breast cancer mortality in Copenhagen after introduction of mammography screening: cohort study. *BMJ* **330**: 220.

Otto SJ, Fracheboud J, Verbeek AL, Boer R, Reijerink-Verheij JC, Otten JD, Broeders MJ, de Koning HJ. National Evaluation Team for Breast Cancer Screening (2012) Mammography screening and risk of breast cancer death: a population-based case–control study. *Cancer Epidemiol Biomarkers Prev* **21**: 66–73.

Paap E, Holland R, den Heeten GJ, van Schoor G, Botterweck AA, Verbeek AL, Broeders M (2010) A remarkable reduction of breast cancer

deaths in screened versus unscreened women: a case-referent study. *Cancer Causes Control* **21**: 1569–1573.

Paap E, Verbeek A, Puliti D, Broeders M, Paci E (2011) Minor influence of self-selection bias on the effectiveness of breast cancer screening in case–control studies in the Netherlands. *J Med Screen* **18**: 142–146.

Paci E, Giorgi D, Bianchi S, Vezzosi V, Zappa M, Crocetti E, Rosselli del Turco M (2002) Assessment of the early impact of the population-based breast cancer screening programme in Florence (Italy) using mortality and surrogate measures. *Eur J Cancer* **38**: 568–573.

Pukkala E, Patama T (2010) Small-area based map animations of cancer incidence in Finland, 1953–2008. Finnish Cancer Registry. Available at http://astra.cancer.fi/cancermaps/suomi5308/index.html (last accessed 26 January 2016).

Puliti D, Miccinesi G, Collina N, De Lisi V, Federico M, Ferretti S, Finarelli AC, Foca F, Mangone L, Naldoni C, Petrella M, Ponti A, Segnan N, Sigona A, Zarcone M, Zorzi M, Zappa M, Paci E. IMPACT Working Group (2008) Effectiveness of service screening: a case–control study to assess breast cancer mortality reduction. *Br J Cancer* **99**: 423–427.

Sarkeala T, Heinävaara S, Anttila A (2008a) Organised mammography screening reduces breast cancer mortality: a cohort study from Finland. *Int J Cancer* **122**(3): 614–619.

Sarkeala T, Heinävaara S, Anttila A (2008b) Breast cancer mortality with varying invitational policies in organized mammography. *Br J Cancer* **98**(3): 641–645.

Seppänen J, Heinävaara S, Anttila A, Sarkeala T, Virkkunen H, Hakulinen T (2006) Effects of different phases of an invitational screening program on breast cancer incidence. *Int J Cancer* **119**: 920–924.

Swedish Organised Service Screening Evaluation Group (2006) Reduction in breast cancer mortality from organized service screening with mammography: 1. Further confirmation with extended data. *Cancer Epidemiol Biomarkers Prev* **15**(1): 45–51.

The Mass Screening Registry (2014) Available at: http://www.cancer.fi/syoparekisteri/en/mass-screening-registry/statistics/koko-maa/ (last accessed 18 August 2014).

van der Waal D, Broeders MJM, Verbeek ALM, Duffy SW, Moss SM (2015) Case–control studies on the effectiveness of breast cancer screening: Insights from the UK Age Trial. *Epidemiology* **26**(4): 590–596.

van Schoor G, Moss SM, Otten JDM, Donders R, Paap E, den Heeten GJ, Holland R, Broeders M, Verbeek AL (2011) Increasingly strong reduction in breast cancer mortality due to screening. *Br J Cancer* **104**: 910–914.

Weiss NS (1994) Application of the case–control method in the evaluation of screening. *Epidemiol Rev* **16**(1): 102–108.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)