IJC
International Journal of Cancer

Benefit-to-harm ratio of the Danish breast cancer screening programme

Anna-Belle Beau[1], Elsebeth Lynge[1], Sisse Helle Njor[2], Ilse Vejborg[3] and Søren Nymand Lophaven[1]

[1] Department of Public Health, University of Copenhagen, DK-1014 Copenhagen, Denmark
[2] Department of Clinical Epidemiology, University of Aarhus, DK-8200 Aarhus, Denmark
[3] Department of Radiology, Copenhagen University Hospital (Rigshospitalet), DK-2100 Copenhagen, Denmark

Cancer Epidemiology

The primary aim of breast cancer screening is to reduce breast cancer mortality, but screening also has negative side-effects as overdiagnosis. To evaluate a screening programme, both benefits and harms should be considered. Published estimates of the benefit-to-harm ratio, the number of breast cancer deaths prevented divided by the number of overdiagnosed breast cancer cases, varied considerably. The objective of the study was to estimate the benefit-to-harm ratio of breast cancer screening in Denmark. The numbers of breast cancer deaths prevented and overdiagnosed cases [invasive and ductal carcinoma *in situ* (DCIS)] were estimated per 1,000 women aged 50–79, using national published estimates for breast cancer mortality and overdiagnosis, and national incidence and mortality rates. Estimations were made for both invited and screened women. Among 1,000 women invited to screening from age 50 to age 69 and followed until age 79, we estimated that 5.4 breast cancer deaths would be prevented and 2.1 cases overdiagnosed, under the observed scenario in Denmark of a breast cancer mortality reduction of 23.4% and 2.3% of the breast cancer cases being overdiagnosed. The estimated benefit-to-harm ratio was 2.6 for invited women and 2.5 for screened women. Hence, 2–3 women would be prevented from dying from breast cancer for every woman overdiagnosed with invasive breast cancer or DCIS. The difference between the previous published ratios and 2.6 for Denmark is probably more a reflection of the accuracy of the underlying estimates than of the actual screening programmes. Therefore, benefit-to-harm ratios should be used cautiously.

The primary aim of breast cancer screening is to reduce breast cancer mortality. Other benefits include more conservative treatment and the positive psychological effect of knowing that a negative test is associated with a reduced breast cancer risk. But screening also has negative side-effects as overdiagnosis, that is, a diagnosis of breast cancer that would in the absence of screening not have led to clinically manifest disease in the woman's lifetime. Further harms to be considered are longer time as a cancer patient due to the lead time, false positive and false negative tests, radiation exposure, negative psychological effects in terms of anxiety about waiting for the mammogram report and the false reassurance to women later diagnosed with interval cancers, Figure 1.

To evaluate a screening programme, both the benefits and harms should be considered. However, it is difficult to weight all benefits and harms at the same time because they are measured in very diverse ways. It has, therefore, become an established procedure to report a benefit-to-harm ratio, which is the number of breast cancer deaths prevented divided by the number of overdiagnosed breast cancer cases. It should be noted that both indicators can be measured only at the population level, as it is not possible to identify the individual woman with a prevented breast cancer death or with an overdiagnosed breast cancer.

In 2012, the EUROSCREEN group estimated that for every 1,000 women screened, 7 to 9 breast cancer death would be prevented, and 4 cases would be overdiagnosed, giving a ratio between 1.8 and 2.3.[1] Using the same methodology, Hofvind *et al.* found a benefit-to-harm ratio of 0.7 for women screened in the Norwegian Breast Cancer Screening Program.[2] The published estimates of the benefit-to-harm ratio have thus varied considerably.

Key words: screening, breast cancer, breast cancer death prevented, overdiagnosed cases, benefit-to-harm ratio
Additional Supporting Information may be found in the online version of this article.
Contributors: E.L., S.L. and A.B.B. designed the study. A.B.B. conducted the analyses and wrote the first draft. All authors commented on this draft and contributed to the final version. All authors had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. A.B.B. is study guarantor.
Competing interests: E.L. collaborates with the Danish company Biomediq, receives no fee from the company, neither has shares in the company. Other authors declare no interest.
Grant sponsor: Olga and Esper Boel's Fund
DOI: 10.1002/ijc.30758
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
History: Received 25 Jan 2017; Accepted 13 Apr 2017; Online 3 May 2017
Correspondence to: Anna-Belle Beau, Department of Public Health, University of Copenhagen, Øster Farimagsgade 5, DK 1014 København K, Denmark, Tel.: +4535331951, E-mail: abeau@sund.ku.dk,

What's new?

Breast cancer screening reduces breast cancer mortality, but one negative side-effect is overdiagnosis. Published estimates of the benefit-to-harm ratio–the number of prevented breast cancer deaths divided by the number of overdiagnosed breast cancer cases–vary considerably. This study reports a benefit-to-harm ratio of 2.6 for women invited to breast cancer screening in Denmark. Among 1,000 invited women from age 50 and followed up until 79, 2–3 women would be prevented from dying from breast cancer for every overdiagnosed woman. International variations in benefit-to-harm ratios probably reflect differences in the accuracy of underlying estimates more than differences between screening programmes.

In Denmark, population based breast cancer screening was implemented in the Copenhagen municipality on April 1, 1991 and in the Funen county on November 1, 1993. Together, these programmes covered 20% of Danish women. Nationwide implementation of screening took place from the late 2007 to 2010. We found implementation of screening to be followed by a 25% decrease in breast cancer mortality in Copenhagen[3] and by a 22% decrease in Funen.[4] We found overdiagnosis, including invasive and ductal carcinoma *in situ* (DCIS), to account for 2.3% of the breast cancer cases in Copenhagen and Funen.[5] This article aims to estimate the benefit-to-harm ratio of breast cancer screening in Denmark using the long-term cohort data on the screening outcomes.

Methods

In Denmark, breast cancer screening targets women aged 50–69 years with biennial invitations, initially using one-or two view(s) mammography and later two views only. All mammograms are read independently by two radiologists.

We assumed screening to affect the incidence and mortality from breast cancer in the age group 50–79 years.[6] The estimates of breast cancer mortality reduction and overdiagnosis were extracted from our previous studies, which used the difference-in-differences method based on four cohorts with individual follow-up data.[3–5] The methodology presented by the EUROSCREEN group was used to estimate the number of breast cancer deaths prevented and the number of overdiagnosed cases.[1] As the EUROSCREEN group estimation is based on cumulative rates for age 50 to 79, the results of our present study will differ from our previously published estimates on number of breast cancer deaths prevented[4] and on number of overdiagnosed cases[5] using other methods. The results were presented for every 1,000 women aged 50 years invited to screening until age 69 years and followed until age 79 years.

Estimates on breast cancer mortality reduction and overdiagnosis

Estimates were derived from observational studies of birth cohorts of women targeted by breast cancer screening. Three control groups were used: a regional control group including the same birth cohorts from the areas of Denmark where screening had not yet been implemented; a historical control group including birth cohorts from screening areas prior to screening; and finally a historical, regional control including birth cohorts from non-screening areas prior to screening. Individual-level data and incidence-based mortality analysis were used. The methodology was a difference-in-differences analysis. The result of this analysis is a ratio of rate ratios. This means that it estimated the change from before to during screening in the screening area, controlled for the change over this time period in the non-screening areas. In this way, it estimated the effect of screening controlled for both regional and time related differences in breast cancer risk.

In Copenhagen, a decrease in breast cancer mortality of 25% (95% confidence interval (CI): 11 to 37%) was found for invited women and of 37% for screened women after controlling for selection bias.[3] In Funen, the decrease in breast cancer mortality for women invited to screening was 22% (95% CI: 11 to 32%) and 28% (95% CI: 13 to 41%) for screened women.[4] We used person-years accumulated during the first ten years of the respective screening programmes for calculation of weighted estimates. Copenhagen contributed 430,823 person-years and Funen 514,556 person-years.[3,4] This gave a weighted estimated effect for invited women of 23.4% (= [(430,823 × 25%) + (514,556 × 22%)]/(430,823 + 514,556)). For screened women, the figure became 32.1% (= [(430,823 × 37%) + (514,556 × 28%)]/(430,823 + 514,556)).

Overdiagnosis was estimated, in the Njor *et al.* article,[5] as the cumulative incidence of breast cancer in a screening region compared with the expected cumulative incidence in the absence of screening. The expected incidence in the absence of screening was estimated from the incidence in the historical control group (including women living in a screening area in the period before screening) controlled for the change in incidence from before to after screening time in the national control group (including women living outside any screening area). It was assumed that the observed number of incident breast cancer cases followed a Poisson distribution. The overdiagnosis estimate was calculated with a Poisson regression analysis which is a standard method for comparison of incidence rates. Njor *et al.*[5] included period (before or after screening), area (in or outside any screening area), exposure to screening (yes or no), and 2-year birth cohort in the analysis. This resulted in an effect of screening, measured as a ratio of rate ratios, on the cumulative breast cancer incidence of 1.023 (95% CI: 0.97 to 1.08). The lower limit of the CI thus corresponded to a 3% reduction in the cumulative incidence of breast cancer in invited as opposed to non-invited women. While this makes sense as long as we

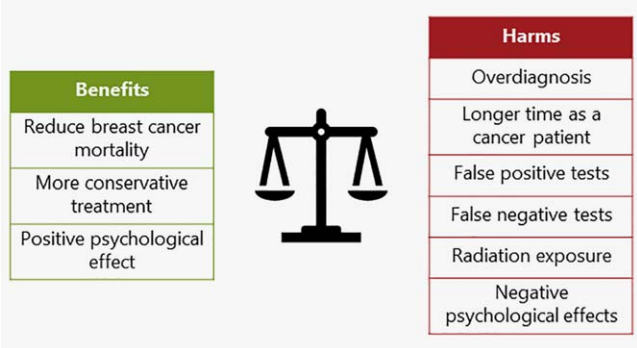

Figure 1. Benefits and harms of breast cancer screening.

talk about the cumulative incidence, it may seem strange to have a negative estimate of overdiagnosis. Therefore, we have reported this lower limit as 1, meaning no overdiagnosis.

For overdiagnosis for invited women, we used the increase in the pooled cumulative incidence of breast cancer (invasive and DCIS) of 2.3% (95% CI: −3 to 8%) reported by Njor *et al.*[5] For screened women, overdiagnosis amounted to 1% in Funen and to 5% in Copenhagen according to Njor *et al.*[5] Person-years of women not followed beyond screening age are not relevant for estimation of overdiagnosis. Therefore, to pool the estimates from Copenhagen and Funen, we weighted them with the person-years reported for all women in the Njor study. Copenhagen contributed 456,499 person-years and Funen 323,363 person-years.[5] This gave a weighted estimate for screened women of 3.3% (= [(456,499 × 5%) + (323,363 × 1%)]/(456,499 + 323,363)).

Measures of cumulative risk of breast cancer and breast cancer death in absence of screening

Cumulative risk of breast cancer (invasive + DCIS) and cumulative risk of breast cancer death from 50 to 79 years in the absence of screening were computed. These risks were expressed as the probability that a woman was diagnosed with breast cancer (invasive or DCIS) or died from the disease between ages 50 and 79 in the absence of screening. From NORDCAN,[7] we extracted the age-specific breast cancer (invasive only) incidence and mortality rates for Danish women aged 50–79 years in the 1996–2000 (*i.e.*, before the programme started in regions outside Copenhagen and Funen). The cumulative risk of breast cancer was 8.8%, and assuming an increase due to DCIS of 2.8%[5], it became 9.1%. The cumulative risk of breast cancer death was 3.4%. Thus, among 1,000 women, we estimated 91 breast cancer (invasive + DCIS) cases and 34 breast cancer deaths (18 among women aged 50–69 years and 16 among those aged 70–79; Table 1).

Based on data used by Olsen *et al.*[3] 21% of breast cancer deaths occurring in women aged 50–69 derived from breast cancer diagnosed before age 50, thus before the women were invited to screening. 51% of deaths from breast cancer in women aged 70–79 derived from breast cancer diagnosed under age 50 or over age 69. In these cases, breast cancer screening could not have influenced the probability of dying from breast cancer. Thus, the number of breast cancer death occurring at aged 50–79 in which screening may have had a protective effect was 22 of 34 (= 18 × (1–0.21) + 16 × (1–0.51)) (Table 1).

In 1996–2000, breast cancer screening programmes were already implemented for the 20% of Danish women living in Copenhagen and Funen. Thus, we estimated the expected cumulative risk of breast cancer in the absence of screening (A) by subtracting from the estimated 91 breast cancer cases an extra 2.3% for the 20% of women coming from Copenhagen and Funen ($A = 91 - (0.2 \times A \times 0.023)$) (Table 1). We estimated the expected cumulative risk of breast cancer mortality in the absence of screening (B) by adding to the estimated 22 breast cancer deaths an extra 23.4% for the 20% of the population coming from Copenhagen and Funen ($B = 22 + (0.2 \times B \times 0.234)$) (Table 1).

Hence, in the absence of screening, the expected numbers of breast cancers and breast cancer deaths potentially affected by screening were 90.6 and 23.1 per 1,000 women aged 50–79 (Table 1). For screened women, the figures were 90.4 and 23.5 per 1,000 women aged 50–79 (Table 1).

Effect of screening

We estimated the effect of screening as the difference between the estimated risks of breast cancer and breast cancer death assuming nationwide screening coverage, and the expected risks in the absence of screening. The numbers of breast cancer deaths prevented and overdiagnosis cases (invasive and DCIS) were estimated per 1,000 women aged 50–79, along with estimates using lower and upper points of the CI of mortality reduction and overdiagnosis estimates. Estimates were made for both invited and screened women.

Several sensitivity analyses were conducted. First, we assumed a theoretical 10% increase in the cumulative risk of breast cancer due to DCIS (instead of the observed 2.8%). Second, we used the NORDCAN rates for 1990 before any screening started in Denmark, showing cumulative risks of 6.9% for breast cancer (invasive + DCIS) and 3.4% for breast cancer death. Hence, the expected numbers of breast cancers and breast cancer deaths in which screening may had an effect were 69 and 22 per 1,000 women aged 50–79. Third, rates of breast cancer incidence and mortality might have progressed over time independently of screening. Simulations, using the observed trend in Denmark from 1970 to 1990,[7] of a 30% increase in cumulative incidence and a 15% increase in cumulative mortality were applied to the 1990 rates. In this estimation, the cumulative risk of breast cancer (invasive + DCIS) was 8.8%, and the cumulative risk of death from breast cancer was 3.9%. Hence, the expected numbers of breast cancers and breast cancer deaths potentially affected by screening were 88 and 26 per 1,000 women aged 50–79.

Table 1. Estimation of cumulative risk of breast cancer (BC) and BC death for every 1,000 women in the Danish breast cancer screening programme aged 50, and followed until age 79

Estimation	For 1,000 women invited biennially from 50 to 69 years	For 1,000 women screened biennially from 50 to 69 years
Breast cancer		
BC (invasive and DCIS) from 50 to 79 years based on age-specific 1996-2000 incidence rates	91	91
Expected BC, corrected for early start programmes (*A*)	90.6 ($A = 91 - (0.2 \times A \times 0.023)$)	90.4 ($A = 91 - (0.2 \times A \times 0.033)$)
Breast cancer death		
BC death (invasive and DCIS) from 50 to 79 years based on age-specific 1996–2000 mortality rates	34	34
BC death, potentially affected by screening	22	22
Expected BC death, corrected for early start programmes (*B*)	23.1 ($B = 22+(0.2 \times B \times 0.234)$)	23.5 ($B = 22+(0.2 \times B \times 0.321)$)

Results

For every 1,000 women aged 50 invited to breast cancer screening biennially to age 69 and followed until age 79, we estimated that 5.4 (=23.1 × 0.234) breast cancer deaths would be prevented and 2.1 (=90.6 × 0.023) cases would be overdiagnosed, under the observed scenarios in Denmark of a breast cancer mortality reduction of 23.4 and 2.3% of the breast cancer cases being overdiagnosed (Fig. 2). The estimated benefit-to-harm ratio for invited women was 2.6 (=5.4/2.1) (Table 2).

For screened women, these numbers were 7.5 (= 23.5 × 0.321) breast cancer deaths prevented, 3.0 (=90.4 × 0.033) cases overdiagnosed, and a benefit-to-harm ratio of 2.5 (Fig. 2, Table 2).

The sensitivity analyses, assuming a theoretical increase of breast cancer incidence of 10% due to DCIS, resulted in an expected cumulative risk of breast cancer in absence of screening of 95.6 per 1,000 women aged 50–79. Among invited women, this led to 2.2 overdiagnosed cases, giving a benefit-to-harm ratio of 2.5 (=5.4/2.2). Among screened women, it gave 3.1 overdiagnosed cases, resulting in a ratio of 2.4 (=7.5/3.1).

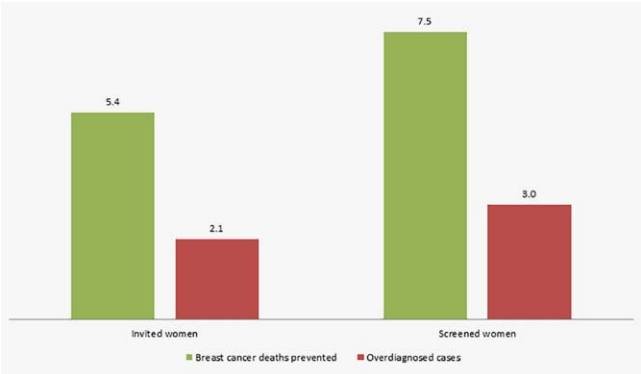

Figure 2. Estimated number of breast cancer deaths prevented and overdiagnosed cases for every 1,000 women aged 50, invited to screening or screened biennially in the Danish breast cancer screening programme to age 69, and followed until age 79. Estimates were based on 23.4% reduction in mortality from breast cancer (invasive + DCIS) and 2.3% overdiagnosis for invited women;[3–5] figures were 32.1 and 3.3% for screened women.[3–5] Estimates were applied to observed incidence of and mortality from breast cancer in Denmark in 1996–2000 (women aged 50–79 years).[7]

Using rates from 1990 resulted in 5.1 breast cancer deaths prevented, 1.6 overdiagnosed cases, and a benefit-to-harm ratio of 3.2 for invited women. For screened women, 7.1 breast cancer deaths would be prevented and 2.3 women would be overdiagnosed, leading to a ratio of 3.1. Lastly, simulating an increase in the 1990 rates of 30% for cumulative breast cancer incidence and of 15% for cumulative breast cancer mortality led to 6.1 breast cancer deaths prevented, 2.0 overdiagnosed cases, and a benefit-to-harm ratio of 3.1 for invited women. For screened women, 8.3 breast cancer deaths would be prevented and 2.9 women overdiagnosed, leading to a ratio of 2.9.

Discussion

Main findings

Among women invited to breast cancer screening in Denmark, the benefit-to-harm ratio was 2.6; meaning that 2–3 women would be prevented from breast cancer death for every woman overdiagnosed. Hence, in term of number of breast cancer deaths prevented, the benefit was two to three times the harm in terms of overdiagnosed cases. An alternative way of expressing the results is that the number of women needed to be invited for screening in order to prevent one breast cancer death was 185. The number of women needed to be screened to prevent one breast cancer death was 133. The sensitivity analyses showed a benefit-to-harm ratio ranging from 2.4 to 3.2 for invited and screened women. Our estimated benefit-to-harm ratios were similar for invited and screened women, as the effects of screening on breast cancer incidence and mortality changed proportionally for invited and screened women.

Strength and weakness

With the comprehensive register data and the long-time interval between implementation of screening in

Table 2. Estimates of the benefit-to-harm ratio in the Danish breast cancer screening programme

	For every 1,000 women aged 50, invited to screening biennially to age 69, and followed until age 79	For every 1,000 women aged 50, screened biennially to age 69, and followed until age 79
Estimate of mortality reduction using[3,4]	23.4% (95% CI: 11 to 34.3%)	32.1%*
Estimate of overdiagnosis using[5]	2.3% (95% CI: −3 to 8%)	3.3%*
Number of breast cancer deaths prevented	5.4 (2.5 to 8.1)	7.5
Number of overdiagnosed cases	2.1 (1 to 7.2)	3.0
Benefit-to-harm ratio	2.6	2.5

Note:
*Confidence intervals not reported in the original studies.

Copenhagen/Funen and in the rest of the country, Denmark offers an ideal setting of estimation of the benefit-to-harm ratio of breast cancer screening. The mortality reduction[3,4] and overdiagnosis[5] used in our estimations came from observational studies following cohorts of women offered and not offered screening; using incidence-based mortality analysis; controlling for pre-screening differences between regions; and including a long post-screening follow-up time. The mortality reduction[3,4] had been estimated using the follow-up analysis found in the Swedish randomized controlled trials to give the most conservative estimate.[8] The overdiagnosis[5] was estimated without the youngest cohorts due to their short post-screening follow-up time, which might have resulted in a slight underestimation.

In the analyses of breast cancer mortality and overdiagnosis, both area and time had been controlled for. But, it was assumed that in the absence of screening, the breast cancer mortality and the breast cancer incidence would over time have changed to the same extent in screening and non-screening area. In statistical terms this means that there was no interaction. In analyses of breast cancer mortality, this assumption was considered reasonable based on the time trends in breast cancer mortality in the pre-screening period.[9] In the analysis of overdiagnosis, Njor *et al.*[5] tested this assumption and observed for the oldest birth cohorts an interaction between area and time period in the Copenhagen data. Breast cancer incidence data from these birth cohorts prior to screening age was used to correct for this interaction. No interaction was found in the Funen data.

Despite the use of the national Danish data, the point estimates of reduction in breast cancer mortality of 23.4% and of overdiagnosis of 2.3% came with some uncertainty. For reduction in breast cancer mortality the 95% CI was from 11 to 34.3% and for overdiagnosis it was from 0 to 8%. If we combined these into the worst scenario for the benefit-to-harm ratio, it became 0.3 (=2.5/7.2), thus indicating almost three extra breast cancer cases diagnosed for each breast cancer death prevented. The best scenario for the benefit-to-harm ratio became 8.1 if we used only the upper limit of the CI for reduction in breast cancer mortality. This means eight breast cancer deaths prevented for each overdiagnosed case. However, none of these extremes seemed realistic, so we considered the point estimate to be the best representation of the benefit-to-harm ratio in Denmark.

The outcomes of breast cancer screening in Denmark were reported also in other studies. Jørgensen *et al.* "were unable to detect any effect of the Danish screening programmes on breast cancer mortality."[10] This study was based on routine breast cancer mortality data including deaths in women diagnosed before screening started. Furthermore, the authors focused on the similarity of the average annual changes (equal to the slopes) in the trends of death rates in screening and non-screening regions from pre-and post-screening periods. They ignored a statistically significant change in the relative level of the trends occurring between their pre-and post-screening periods. Jørgensen *et al.* reported that "one in every three women aged 50 to 69 years diagnosed with breast cancer was overdiagnosed."[11] Estimation was only based on rates of non-advanced tumors because "there was no clear compensatory decrease in the incidence of advanced cancer in older women" (*i.e.*, women aged 70 to 84 years). However, almost half of the observations in the age-group 70–84 years were from women never invited to screening or still invited despite having passed the age of 70 years. These women could not contribute to the compensatory dip. Therefore, we did not include these studies in our estimations.

We present a benefit-to-harm ratio derived from simple age-specific breast cancer incidence and mortality rates from Denmark. A more complex analysis using individual level data might have led to different results, but the simplicity of the interpretation the possibility to compare with other studies might have been lost.

Moreover, our results focused only on breast cancer mortality reduction and overdiagnosis. But other benefits and harms of mammography screening should be considered in evaluating a population-based screening programme. Other benefits include more conservative treatment and the positive psychological effect of knowing that a negative test is associated with a reduced breast cancer risk. Other harms of mammography screening include longer time as a cancer patient due to the lead time, false positive and false negative tests. Further harms to be considered are radiation exposure and negative psychological effects in terms of anxiety about

waiting for the mammogram report and the false reassurance to women later diagnosed with interval cancers (Fig. 1).

Comparison with other studies

Our benefit-to-harm ratio was more in favor of screening than the results from the Marmot report[12] and from the Norwegian Breast Cancer Screening Program,[2] but similar to results reported by Duffy *et al.*[13] and the EUROSCREEN group.[1]

The Marmot report[12] used results from the randomized controlled trials (RCTs) to estimate the benefit and risk of breast cancer screening. For every 1,000 women invited to screening, 4.3 breast cancer deaths would be prevented and 12.9 cases would be overdiagnosed. Combining these estimates would lead to a benefit-to-harm ratio of 0.33. Similarly, for every 1,000 women screened, 5.6 breast cancer deaths would be prevented and 16 cases would be overdiagnosed, leading to a ratio of 0.35. While the impact of screening on breast cancer mortality was estimated based on all nine RCTs, overdiagnosis was estimated only from the three RCTs "that did not systematically screen the control group at the end of the screening period."[12] However, even these RCTs were not well suited for estimation of overdiagnosis.[14] In the Malmö RCT, the excess cumulative breast cancer incidence derived exclusively from women screened on average to age 76.5, which left limited time for post-screening follow-up. The two Canadian RCTs targeted women aged 40–49[15] and 50–59 years,[16] but shortly after the trials stopped, service screening was implemented for women aged 50–69 in the majority of Canadian provinces from which the trial populations were recruited.[17]

Hofvind *et al.*[2] presented estimates from the Norwegian Breast Cancer Screening Program. For every 1,000 invited women, 4.0 breast cancer deaths would be prevented and 8.9–10.6 cases overdiagnosed, giving a benefit-to-harm ratio of 0.38–0.45. For screened women, the numbers were 5.7, 8.2, and a ratio of 0.7. The impact of screening on breast cancer mortality was estimated from a Norwegian study showing a 28% reduction for invited women.[18] The same study gave an estimate of 36.8% for screened women,[18] while another study gave an estimate of 43% based on comparison of screened with non-screened women.[19] For screened women, an overdiagnosis percentage was combined from two studies. One from the Norwegian screening programme where excess cases in regular screened women during screening + deficit cases in ever screened in post-screening were compared with numbers in never screened women giving an estimate of 18.5%.[20] The second estimate of 7% came from comparison of programme-screened with outside screened + never screened women in the Norwegian Women and Cancer Study.[21]

The Danish and Norwegian screening programmes are very similar. Therefore, it is noteworthy that while the estimated impact on breast cancer mortality was almost the same in the two countries, the Norwegian estimates of overdiagnosis based on approximations were considerably higher than the Danish estimate based on the observed excess cumulative incidence. While both the Danish and the Norwegian estimates divide the number of overdiagnosed breast cancer cases with all breast cancer cases diagnosed in the follow-up period, there are, nevertheless, methodological differences between the two estimates of overdiagnosis. The Danish estimate is the only one based on observational data with enough time for sufficient post-screening follow-up, while the estimate from Norway was based predominantly on observations from the time during or shortly after end of invitation to screening.

Duffy *et al.*[13] presented data from the Swedish Two-County RCT and the UK Breast Screening Programme in England. For every 1,000 screened women, the Two-County RCT found 8.8 breast cancer deaths would be prevented. The control group in this RCT was offered screening after the end of the trial. To estimate overdiagnosis, Duffy *et al.* excluded breast cancer cases diagnosed during the prevalence screen of the control group and estimated 4.3 extra breast cancer cases diagnosed per 1,000 women in the intervention group as compared with the control group. This gave a benefit-to-harm ratio of 2.0. For England, breast cancer incidence and mortality rates from 1974–1988 were used to predict the expected number in the absence of screening after the start of screening in 1989. In 1995–2004, 28% fewer breast cancer deaths were observed in women aged 50–69 years, as compared with the expected based on the rates for 1974–1988, and adjusted for changes over time in the younger and older age-groups. This translated into 5.7 prevented breast cancer deaths per 1,000 screened women. Overdiagnosis was estimated as the observed number of cases in 1989–2003 in women aged 45+ compared with the expected number based on pre-1988 rates for the same age-groups and standardized to the age-group below 45 years. This resulted in 2.3 overdiagnosed cases per 1,000 screened women. This number might, however, be too high as screening was implemented only gradually for women aged 50–64 years in 1989–1993, and women aged 65–69 years were offered screening in 2002–2003. Limited time had therefore been available for accumulation of post-screening person years. The English data resulted in a benefit-to-harm ratio of 2.5.

The EUROSCREEN estimates were based on joint data from the European service screening programmes available in 2012.[1] For screened women, the reduction in breast cancer mortality was 38% in incidence-based mortality studies and 48% in the case-control, giving 7–9 prevented breast cancer deaths for 1,000 screened women. Overdiagnosis had been assessed with various methodologies, providing a joint estimate of 6.5%, giving four extra cases for 1,000 screened women. The benefit-to-harms ratio thus became 1.8–2.3.

The data and methodology used in the benefit-to-harm studies are summarized in the Supporting Information Table 1. Estimates of overdiagnosis are sensitive to methodology.[22] We used only estimates, where there was correspondence

between the percentage of overdiagnosis reported and the age-group to which this percentage was applied. In the Marmot report, the numbers were given for women aged 50 years and followed for 20 years.[12] In the Two County Study for women aged 50–69 years at randomization,[13] in the English data for women aged 45 and above,[13] while both the Norwegian[2] and the EUROSCREEN[1] studies used data for women aged 50–79 years, as we did.

The reported benefit-to-harm ratios varied due to differences in number of prevented breast cancer deaths and/or number of overdiagnosed cases. The reductions in breast cancer mortality varied from 20% in the Marmot report[12] to 32% from the Two County Study.[13] This reflected that the Marmot report used all RCTs, including the two Canadian outliers,[15,16] and that the Two County Study was the single RCT with the largest reduction in breast cancer mortality. The proportion of overdiagnosed cases varied from 2.3% in the Danish data to 15.6% in the Norwegian data. To generate valid data on overdiagnosis, high quality RCTs with no screening of women in the control arm at any point and with long-follow-up would be required. Unfortunately, such RCT data are not available. Previous studies of overdiagnosis have, therefore, used various modeling (Two County Study) and/or approximation (Norway) or short-term observational data (England).

In order to implement truly shared decision making in breast cancer screening, information is requested on benefits and harms of screening. Our comparison of the reported ratios illustrated considerable variation. The difference between the ratios for invited women of 0.4 for Norway and 2.6 for Denmark is probably more a reflection of the accuracy of the underlying estimates than of the actual screening programmes. Observational data rely on assumption about what does and what does not vary over time and geography, and/or modeling to deal with the uncertainties. Therefore, all benefit-to-harm ratios estimated from available observational data should be used cautiously.

Conclusions

Our study estimated that for invited women the benefit-to-harm ratio of the Danish breast cancer screening programme was 2.6. Among 1,000 women invited to screening from age 50 and followed until 79, 2–3 women would be prevented from dying from breast cancer for every woman overdiagnosed with invasive breast cancer or DCIS.

Acknowledgments

The funders had no role in the study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. Researchers were independent of the funders.

References

1. Paci E. 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;19:5–13.
2. Hofvind S, Román M, Sebuødegård S, et al. Balancing the benefits and detriments among women targeted by the Norwegian Breast Cancer Screening Program. *J Med Screen* 2016;23:203–9.
3. Olsen AH, Njor SH, Vejborg I, et al. Breast cancer mortality in Copenhagen after introduction of mammography screening: cohort study. *BMJ* 2005;330:220.
4. Njor SH, Schwartz W, Blichert-Toft M, et al. Decline in breast cancer mortality: How much is attributable to screening? *J Med Screen* 2015;22: 20–7.
5. Njor SH, Olsen AH, Blichert-Toft M, et al. Overdiagnosis in screening mammography in Denmark: population based cohort study. *BMJ* 2013; 346:f1064.
6. Boer R, Warmerdam P, de Koning H, et al. Extra incidence caused by mammographic screening. *Lancet* 1994;343:979.
7. NORDCAN [Internet]. Available at: http://www-dep.iarc.fr/NORDCAN.htm. (Accessed on Sept. 19, 2016)
8. Nyström L, Andersson I, Bjurstam N, et al. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet Lond Engl* 2002;359:909–19.
9. Andreasen AH, Andersen KW, Madsen M, et al. Regional trends in breast cancer incidence and mortality in Denmark prior to mammographic screening. *Br J Cancer* 1994;70:133–7.
10. Jørgensen KJ, Zahl P-H, Gøtzsche PC. Breast cancer mortality in organised mammography screening in Denmark: comparative study. *BMJ* 2010;340:c1241.
11. Jørgensen KJ, Gøtzsche PC, Kalager M, et al. Breast Cancer Screening in Denmark: a cohort study of tumor size and overdiagnosis. *Ann Intern Med* 2017;166:313–23.
12. Marmot MG, Altman DG, Cameron DA, et al. The benefits and harms of breast cancer screening: an independent review. *Br J Cancer* 2013;108:2205–40.
13. Duffy SW, Tabar L, Olsen AH, et al. Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. *J Med Screen* 2010;17:25–30.
14. Njor SH, Garne JP, Lynge E. Over-diagnosis estimate from The Independent UK Panel on Breast Cancer Screening is based on unsuitable data. *J Med Screen* 2013;20:104–5.
15. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. *Ann Intern Med* 2002;137:305–12.
16. Miller AB, To T, Baines CJ, et al. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50–59 years. *J Natl Cancer Inst* 2000;92:1490–9.
17. Organized Breast Cancer Screening Programs in Canada—Report on Program Performance in 2005 and 2006. Available at: http://www.phac-aspc.gc.ca/cd-mc/publications/cancer/obcsp-podcs05/index-eng.php. (Accessed on Sept 19, 2016).
18. Weedon-Fekjær H, Romundstad PR, Vatten LJ. Modern mammography screening and breast cancer mortality: population study. *BMJ* 2014;348: g3701.
19. Hofvind S, Ursin G, Tretli S, et al. Breast cancer mortality in participants of the Norwegian Breast Cancer Screening Program. *Cancer* 2013;119: 3106–12.
20. Falk RS, Hofvind S, Skaane P, et al. Overdiagnosis among women attending a population-based mammography screening program. *Int J Cancer* 2013;133:705–12.
21. Lund E, Mode N, Waaseth M, et al. Overdiagnosis of breast cancer in the Norwegian Breast Cancer Screening Program estimated by the Norwegian Women and Cancer cohort study. *BMC Cancer* 2013;13:614.
22. de Gelder R, Heijnsdijk EAM, van Ravesteyn NT, et al. Interpreting overdiagnosis estimates in population-based mammography screening. *Epidemiol Rev* 2011;33:111–21.