Overdiagnosis in breast cancer screening

Lynge, Elsebeth; Beau, Anna-Belle; Christiansen, Peer; von Euler-Chelpin, My; Kroman, Niels; Njor, Sisse; Vejborg, Ilse

Published in:
European Journal of Cancer

DOI:
10.1016/j.ejca.2017.04.018

Publication date:
2017

Document version
Publisher's PDF, also known as Version of record

European Journal of Cancer 80 (2017) 26–29

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Current Perspective

Overdiagnosis in breast cancer screening: The impact of study design and calculations

Elsebeth Lynge [a,*], Anna-Belle Beau [a], Peer Christiansen [b], My von Euler-Chelpin [a], Niels Kroman [c], Sisse Njor [d], Ilse Vejborg [c]

[a] Department of Public Health, University of Copenhagen, Øster Farimagsgade 5, DK 1014 København K, Denmark
[b] Plastic and Breast Surgery, Breast Surgery Unit, Aarhus University Hospital/Randers Regional Hospital, Norrebrogade 44, DK 8000 Aarhus C, Denmark
[c] Copenhagen University Hospital, Rigshospitalet, Copenhagen, Blegdamsvej 9, DK 2100 København K, Denmark
[d] Department of Clinical Epidemiology, Aarhus University, Olof Palmes Allé 43-45, DK 8200 Aarhus N, Denmark

Received 15 March 2017; received in revised form 5 April 2017; accepted 17 April 2017
Available online 20 May 2017

KEYWORDS
Breast cancer;
Screening;
Overdiagnosis;
Study design;
Assumptions

Abstract Overdiagnosis in breast cancer screening is an important issue. A recent study from Denmark concluded that one in three breast cancers diagnosed in screening areas in women aged 50–69 years were overdiagnosed. The purpose of this short communication was to disentangle the study's methodology in order to evaluate the soundness of this conclusion. We found that both the use of absolute differences as opposed to ratios; the sole focus on non-advanced tumours and the crude allocation of tumours and person-years by screening history for women aged 70–84 years, all contributed to the very high estimate of overdiagnosis. Screening affects cohorts of screened women. Danish registers allow very accurate mapping of the fate of every woman. We should be past the phase where studies of overdiagnosis are based on the fixed age groups from routine statistics.

1. Introduction

When it comes to evaluation of breast cancer screening, Denmark is in a particular position. A population-based screening program was offered to around 20% of Danish women up to 17 years before it was offered to the rest of Danish women. This allows for comparison of women exposed and women not exposed to screening. The

* Corresponding author.
E-mail address: elsebeth@sund.ku.dk (E. Lynge).

http://dx.doi.org/10.1016/j.ejca.2017.04.018

screening program targeted women aged 50–69 years, and opportunistic screening was rare. On this basis, several studies have been undertaken on the impact of screening on breast cancer mortality [1–3], and overdiagnosis [4–6].

Recently, Jørgensen *et al.* [7] published a new Danish study on overdiagnosis. This study was based on the number of breast cancers divided into advanced (>20 mm in tumour diameter) and non-advanced cases. Person-years were estimated from the official Danish statistics. Incidence rates were calculated for women aged 35–49, 50–69 and 70–84 years and for two geographical areas (screening and non-screening area) and two periods (before and after screening started). The study concluded that 'one in every three women aged 50–69 years diagnosed with breast cancer was overdiagnosed in the screening area'.

If true, breast cancer screening in Denmark causes considerable harm and would be unjustified as a public health policy. It is therefore important to understand the analysis behind the conclusion by Jørgensen *et al.* With this purpose in mind, we looked into the details of the methodology. Jørgensen *et al.* estimated overdiagnosis in breast cancer screening in Denmark using two approaches, and we discussed them one by one.

2. Jørgensen approach 1

First, overdiagnosis was calculated based on absolute difference in changes of breast cancer incidence rates: (absolute difference between after and before in screening area)−(absolute difference between after and before in non-screening area) as numerator, and the incidence rate in the after period in the non-screening area as denominator. For women aged 50–84 years, this gave ([351.3−226.1]−[280−182.4])/280 = 0.099 or 9.9% (Jørgensen Table 3). All incidence rates were per 100,000, but for simplicity, we omitted the 'per 100,000' from both numerator and denominator.

However, the absolute difference in an outcome between an exposed and a non-exposed group cannot be used as a measure of the strength of the association between the exposure and the outcome. The size of the absolute difference depends not only on the changes over time but also on the levels before. Let us illustrate this with an example. In the data by Jørgensen *et al.*, (Jørgensen Table 3) for women aged 50–84 years, the incidence rate before screening was 226.1 in the screening area and 182.4 in the non-screening area. If both incidence rates increased by 10% they would become 248.7 and 200.6, respectively. The two areas had then undergone exactly the same changes over time, and the ratio of the rate ratios would be 1.00, but the absolute difference would be 4.4.

The actual impact of screening on breast cancer incidence would therefore be better measured with the ratio of rate ratios than with the absolute difference (Table 1). In the data by Jørgensen *et al.*, the ratio of the rate ratios was 1.01 (i.e. [351.3/226.1]/[280/182.4]) (Jørgensen Table 3). This would indicate an overdiagnosis of 1%, as compared with the overdiagnosis of 9.9%, Jørgensen *et al.* calculated from the absolute difference. Both calculations are, however, problematic due to limitations in the study design used by Jørgensen *et al.*, see in the following section.

3. Jørgensen approach 2

Second, overdiagnosis was calculated from the increase in non-advanced tumours in women aged 50–69 years. Advanced tumours were disregarded in the calculation for two reasons. The first reason was because the rate 'did not decrease in the screening area when incidence trends among women aged 35–49 years were accounted for'. For advanced tumours, the ratios of the rate ratios were 0.48, 0.66 and 0.69, respectively, for women aged 35–49, 50–69 and 70–84 years (Jørgensen Table 2). So, there was a decrease in all age groups, but the authors' argument seems to be that given a decrease in the rate in women below screening age, screening could not explain the decrease in the rate in women above screening age. The change in women below screening age is of course interesting but hardly tells about impact of screening.

The second reason for disregarding advanced tumours was because 'there was no compensatory decrease in the incidence of advanced tumours in older women'. Older women are here women aged 70–84 years. However, in older women, the ratio of the rate ratios for advanced tumours did in fact decrease more in screening than in non-screening areas; (154.8/124.0)/(162.2/89.7) = 0.69 (Jørgensen Table 2; 95% confidence interval 0.63–0.75, our calculation). It therefore seems strange that the authors found that overdiagnosis could be estimated based solely on data for non-advanced tumours.

In the second approach, the numerator was the absolute difference in changes of incidence rates in non-advanced breast cancer and the denominator was the incidence rate of advanced and non-advanced breast cancer in the non-screening area. For women aged

Table 1
Incidence rates of breast cancer per 100,000 person-years among women aged 50–84 years and changes over time calculated as a difference and as a ratio (data from Jørgensen, Table 3).

Area	Before	After	Difference	Ratio
Screening	226.1	351.3	125.2	1.55
Non-screening	182.4	280.0	97.6	1.54
Screening versus non-screening			27.6	1.01
Estimated overdiagnosis			9.9%[a] (i.e. [27.6/280] × 100)	1%

[a] The estimate of overdiagnosis by Jørgensen *et al.*

50–69 years, this gave ([258.9–111.4]–[142.0–95.6])/262.1 = 0.386, or 38.6% (Jørgensen Table 2). The authors added an absolute difference in incidence of ductal carcinoma *in situ* (DCIS) between screening and non-screening areas in the after period. This was (38.2–12.4) = 25.8. Together (101.1 + 25.8)/262.1 = 0.483, or 48.3%; resulting in the conclusion that '1 in every 3 women aged 50–69 years diagnosed with breast cancer was overdiagnosed'. Estimates of overdiagnosis are highly depending on not only the numerator but also on the denominator [8]. For women aged 50–84 years, Jørgensen estimated an overdiagnosis of 26% (Jørgensen Table 4), not including DCIS.

In the studied screening areas, screen-detected tumours constituted 54% of all tumours (screen-detected, interval cancers and tumours in non-participants) when DCIS was included and 51% when only invasive tumours were included [9]. Overdiagnosis can affect only the screen-detected part of the tumours. With the calculation by Jørgensen *et al.*, about half of the tumours detected at screening should be overdiagnosed (i.e. [1.486–0.46]/0.54 = x; x = 1.89).

4. Jørgensen study design

A key problem with the study by Jørgensen *et al.* was the design. Screening changes the age-specific pattern of breast cancer incidence in the targeted birth cohorts [10], including a prevalence peak during the first screen, an artificial ageing during the subsequent screens and a compensatory decrease after the end of screening age [11].

In Denmark, screening targeting women aged 50–69 years at the beginning of a biennial invitation round started in Copenhagen in April 1991; screening targeting women aged 50–69 years at the time of invitation started in Funen in November 1993 and screening targeting women aged 50–69 years started in Frederiksberg in June 1994 (program later merged with that of Copenhagen). During the first biennial rounds also women above the age of 70 years were invited.

For women 'no longer offered screening', Jørgensen *et al.* used data from the age group of 70–84 years for the period of 1991–2010 from Copenhagen and Frederiksberg and from 1994 to 2010 for Funen. For Copenhagen, these observations are illustrated in Fig. 1. However, some women contributing observations were never invited to screening (in red), and a small part of the women were still invited despite having passed the age of 70 years (in blue). Furthermore, women over screening age still develop new incident cancers unaffected by screening (in grey and green). In Denmark, the compensatory decrease occurred mainly within the first 8 years after women were no longer invited (in grey) [6].

Jørgensen *et al.* had 3530 invasive breast cancers in the age group of 70–84 years in the screening area in the after period (Jørgensen Supplementary Table). Data were not provided by screening program and detailed age group. However, Copenhagen constituted about 40% of the targeted population; Frederiksberg 10% and Funen 50%. Jørgensen *et al.* used breast cancer data for 20 years from Copenhagen and Frederiksberg and from 17 years from Funen. For the illustration here, it was therefore reasonable to assume that 1526 of the cancers came from Copenhagen; 382 from Frederiksberg and

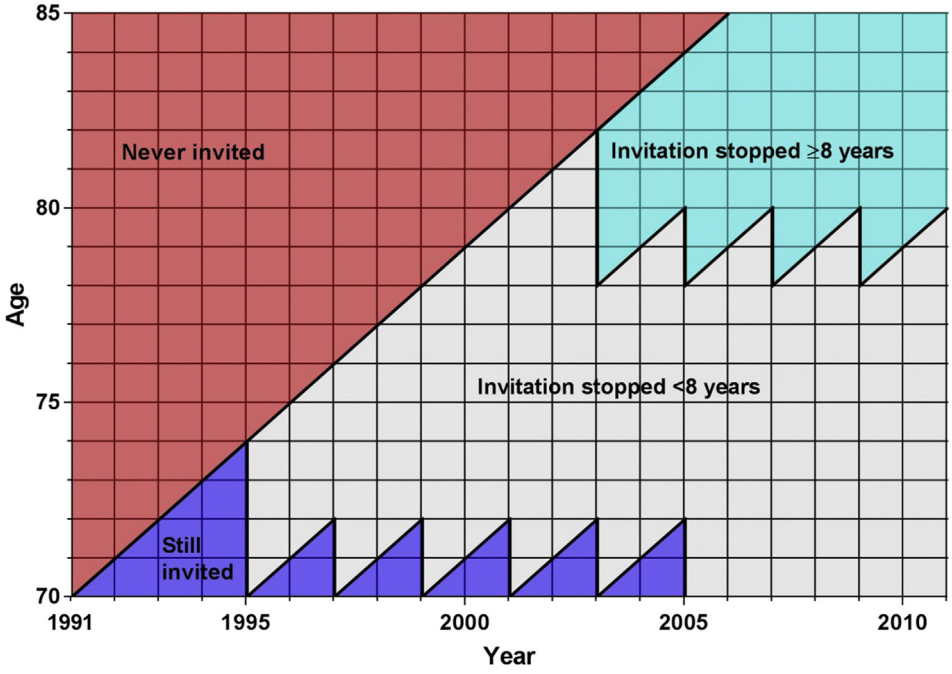

Fig. 1. Lexis' diagram. Women aged 70–84 years in the municipality of Copenhagen, Denmark, 1991–2010 by screening experience.

Table 2
Estimated distribution of 3530 invasive breast cancers by screening status for women aged 70–84 years in screening areas 1991/1994–2010 in the Jørgensen data.

Screening status	Copenhagen		Frederiksberg		Funen		Total breast cancer	
	% observations[a]	N breast cancer	% observations[a]	N breast cancer	% observations[a]	N breast cancer	N	%
Never invited	37.5%	572	52.5%	201	44.1%	715	1488	42.1%
Still invited	6.0%	92	2.7%	10	6.3%	102	204	5.8%
Post invitation <8 years	42.0%	641	34.0%	130	40.8%	662	1433	40.6%
Post invitation ≥8 years	14.5%	221	10.8%	41	8.8%	143	405	11.5%
Total	100%	1526	100%	382	100%	1622	3530	100%

[a] Percent of area in Lexis' diagram.

1622 from Funen. A rough estimation indicated that 42% of the breast cancers in the age group of 70–84 years occurred in women never offered screening, 6% in women still invited to screening, 41% in women who stopped being invited for screening within the last 8 years and 11% in women who stopped being invited for screening more than 8 years ago (Table 2). It is not possible to capture correctly the size of overdiagnosis with such crude data.

5. Conclusion

Our analysis illustrated that several factors contributed to the very high estimate of overdiagnosis presented by Jørgensen *et al.* [7]. First, the use of absolute differences as opposed to ratios; second, focus on only non-advanced cancers and third, an inadequate study design. Screening affects the age-specific incidence of breast cancer in cohorts of screened women, but this classic wisdom has not yet entered mainstream epidemiology. Studies of overdiagnosis are still published without proper consideration of the cohort perspective, and the study by Jørgensen *et al.* is an example. For the sake of screened women, researchers and journal editors should join efforts to find the best possible method for estimation of overdiagnosis.

Funding

None.

Conflict of interest statement

None declared.

References

[1] Olsen AH, Njor SH, Vejborg I, Schwartz W, Dalgaard P, Jensen M-B, et al. Breast cancer mortality in Copenhagen after introduction of mammography screening: cohort study. BMJ 2005;330:220. http://dx.doi.org/10.1136/bmj.38313.639236.82.

[2] Jørgensen KJ, Zahl P-H, Gøtzsche PC. Breast cancer mortality in organised mammography screening in Denmark: comparative study. BMJ 2010;340:c1241.

[3] Njor SH, Schwartz W, Blichert-Toft M, Lynge E. Decline in breast cancer mortality: how much is attributable to screening? J Med Screen 2015;22:20–7. http://dx.doi.org/10.1177/0969141314563632.

[4] Olsen AH, Agbaje OF, Myles JP, Lynge E, Duffy SW. Overdiagnosis, sojourn time, and sensitivity in the Copenhagen mammography screening program. Breast J 2006;12:338–42. http://dx.doi.org/10.1111/j.1075-122X.2006.00272.x.

[5] Jørgensen KJ, Zahl P-H, Gøtzsche PC. Overdiagnosis in organised mammography screening in Denmark. A comparative study. BMC Womens Health 2009;9:36. http://dx.doi.org/10.1186/1472-6874-9-36.

[6] Njor SH, Olsen AH, Blichert-Toft M, Schwartz W, Vejborg I, Lynge E. Overdiagnosis in screening mammography in Denmark: population based cohort study. BMJ 2013;346:f1064.

[7] Jørgensen KJ, Gøtzsche PC, Kalager M, Zahl P-H. Breast cancer screening in Denmark: a cohort study of tumor size and overdiagnosis. Ann Intern Med 2017;166:313. http://dx.doi.org/10.7326/M16-0270.9.

[8] de Gelder R, Heijnsdijk EAM, van Ravensteyn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. Epidemiol Rev 2011;33:111–21.

[9] Domingo L, Jacobsen KK, von Euler-Chelpin M, Vejborg I, Schwartz W, Sala M, et al. Seventeen-years overview of breast cancer inside and outside screening in Denmark. Acta Oncol 2013;52:48–56. http://dx.doi.org/10.3109/0284186X.2012.698750.

[10] Boer R, Warmerdam P, de Koning H, van Oortmarssen G. Extra incidence caused by mammographic screening. Lancet 1994;343:979. http://dx.doi.org/10.1016/S0140-6736(94)90105-8.

[11] Møller B, Weedon-Fekjaer H, Hakulinen T, Tryggvadóttir L, Storm HH, Talbäck M, et al. The influence of mammographic screening on national trends in breast cancer incidence. Eur J Cancer Prev Off J Eur Cancer Prev Organ ECP 2005;14:117–28.