Review
Overdiagnosis and overtreatment of breast cancer
Overdiagnosis in randomised controlled trials of breast cancer screening
Sue Moss

Cancer Screening Evaluation Unit, Institute of Cancer Research, Sutton, Surrey, UK

Corresponding author: Sue Moss, sue.moss@icr.ac.uk

Abstract

Data from randomised controlled trials of mammographic screening can be used to determine the extent of any overdiagnosis, as soon as either a time equivalent to the lead-time has elapsed after the final screen, or the control arm has been offered screening. This paper reviews those randomised trials for which breast cancer incidence data are available. In recent trials in which the control group has not been offered screening, an excess incidence of breast cancer remains after many years of follow-up. In those trials in which the control arm has been offered screening, although there is a possible shift from invasive to *in situ* disease, there is no evidence of overdiagnosis as a result of incident screens.

Introduction

Overdiagnosis in mammographic screening is taken here to mean the diagnosis of invasive or *in situ* breast cancer that, in the absence of screening, would not have presented clinically during the woman's lifetime.

In studying overdiagnosis, randomised controlled trials have the advantage that data on the incidence of breast cancer in the intervention and control arms are usually available in detail at an individual level. Overdiagnosis of both ductal carcinoma *in situ* (DCIS) and invasive cancer may occur; however, it is not easy to determine to what extent an excess of DCIS is due to stage-shifting from invasive disease, although estimates can be made where sufficiently detailed information is available [1]. Most trials have provided relatively little information on the treatment of breast cancer cases, so that the extent of overtreatment is difficult to quantify.

Overdiagnosis can be studied in randomised controlled trials by comparing the cumulative incidence of breast cancer in the intervention and control arms at different times from date of entry or randomisation. While screening is continuing in the intervention arm of a trial, incidence in that arm will be increased because of the advancement of diagnosis by the lead-time in screen-detected cancers, as well as by any overdiagnosis. This 'prevalence peak' will be followed by a corresponding decrease once screening ceases. Over-diagnosis can therefore be estimated only after a time equivalent to the lead-time has elapsed following the final screen. In several trials, women in the control arm have subsequently been offered screening. Once this has occurred, only overdiagnosis due to incident, not prevalent, screens would be observable, because women in both arms of the trial would be subject to any overdiagnosis occurring at prevalent screens.

The extent of any overdiagnosis in trials of breast screening may be affected by the 'intensity' of screening (one or two views, modalities employed, screening frequency and recall policy), and by the uptake of screening in the intervention arm. It may also depend on the age range of women included in the trial, both because of variation in the natural history of the disease with age and because of increased mortality from other causes in older women during the 'lead-time' before a screen-detected cancer would have presented clinically. The extent to which overdiagnosis is observed will also depend on the extent of 'contamination' in the control arm by opportunistic screening.

Method

This review considers those randomised trials that include screening by mammography (with or without clinical examination). There are eight randomised controlled trials of mammography that have so far completed and reported mortality results, and for which data on breast cancer

---

BSE = breast self-examination; DCIS = ductal carcinoma *in situ*; HIP = Health Insurance Plan; MP = mammography plus physical examination; NBSS = National Breast Screening Study; PO = physical examination alone.

Table 1

Randomised controlled trials of mammographic screening

Study	Age range at entry (years)	Screening interval (months)	Number of views	Uptake rate (%) (prevalent screen)	Relative risk of breast cancer mortality	Follow-up (years)
HIP	40–64	12	2	67	0.78	18
Two County	40–74	24–33	1	89	0.70	11
Malmö I	45–70	18–24	2	74	0.81	19
Malmö II	43–49	18–24	2	75–80	0.65	9
Stockholm	40–64	28	1	81	0.90	15
Goteborg	39–59	18	2	84	0.78	13
NBSS I[a]	40–49	12	2	100	1.06	13
NBSS II[b]	50–59	12	2	100	1.02	13
Edinburgh	45–64	12[c]	2	61	0.78	13

[a]Comparison of annual mammography, clinical breast examination (CBE) and breast self-examination (BSE) instruction with a single CBE and BSE instruction. [b]Comparison of annual mammography with annual CBE. [c]Mammography and CBE every 24 months; CBE only in intervening years. HIP, Health Insurance Plan; NBSS, National Breast Screening Study.

incidence are available [2]. The main characteristics of these trials are described in Table 1.

Data have been abstracted from published reports on the cumulative incidence of breast cancer in intervention and control arms in each trial; where available, data have been abstracted for a period of follow-up extending sufficiently beyond the final screen to allow for lead-time, or after women in the control arm have been invited for screening. The absolute excess per 1,000 women years in the intervention arm compared with the control arms is presented in Table 2 for invasive breast cancers, for DCIS, and for invasive cancers and DCIS combined, together with 95% confidence intervals for the absolute excess. The ratio of the incidence rate of all breast cancers in the intervention arm to that in the control arm is shown in Fig. 1.

For studies in which rates have been published only per 1,000 women, rates per 1,000 women years have been calculated on the basis of estimated mean follow-up.

The randomised trials
The earliest of the randomised trials was the Health Insurance Plan (HIP) study performed in the United States in the 1960s, in which women in the intervention arm were offered annual screening by both mammography and clinical breast examination for four rounds. By the end of 5 years from date of entry (about 1.5 years after the last screen) the incidence of breast cancer in the two arms was similar (2.03 per 1,000 women years in the intervention arm, and 1.94 in the control arm [3]). An earlier report indicates that the percentages of *in situ* cancer were 13% and 8%, respectively, in the two arms [4]. There have been considerable improvements in the sensitivity of mammography since this trial was conducted,

and it therefore provides little indication of the potential for overdiagnosis with current techniques.

The four randomised trials conducted in Sweden all used screening by mammography alone, with screening intervals ranging from 18 to 33 months. In all except one of these the control group has subsequently been offered screening.

The Swedish Two County Study, which began in 1977/8, included 143,867 women aged 40 to 74 years at the date of randomisation [5]. Women in the intervention arm were invited to screening at intervals of 24 to 33 months.

In 1985, after two to four rounds of screening in the intervention arm, women in the control arm were offered screening. After this screen had taken place, rates of invasive cancer have been reported as 16.90 and 17.79 per 1,000 women in the intervention and control arms, respectively; assuming an average of 7 years of follow-up, the estimated rates per 1,000 women years are 2.41 and 2.54, respectively, with rates of DCIS of 0.23 and 0.12 per 1,000 women years, respectively [6].

The Gothenberg trial invited women aged between 35 and 59 years, between 1982 and 1991. Again, women in the control arm were offered a single screen at approximately the same time as the final screen in the intervention arm.

The cumulative incidence of invasive breast cancer was greater in the intervention arm than in the control arm until year 6, at about the time of the first screen in the control arm. At the end of the 'screening phase' of the trial (that is, the period up to and including the first screen in the control arm), rates of DCIS were slightly higher in the intervention arm (about 0.22 versus 0.17

231

Table 2

Randomised controlled trials of mammography screening: differences in breast cancer incidence between intervention and control arms in the follow-up period

Study	Cumulative incidence of invasive breast cancer (per 1,000 women years)			Cumulative incidence of DCIS (per 1,000 women years)			Cumulative incidence of all breast cancer (invasive + DCIS) (per 1,000 women years)			Follow-up (years)	Control group screened
	Control group	Absolute excess in intervention arm	95% CI	Control group	Absolute excess in intervention arm	95% CI	Control group	Absolute excess in intervention arm	95% CI		
Two County	2.54	−0.13[a]	−0.29, 0.04	0.12	0.11	0.07, 0.15	2.66	−0.02	−0.19, 0.15	11	Yes
Stockholm	0.85	−0.04	−0.18, 0.10	0.06	0.03	−0.00, 0.07	0.91	−0.03	0.15, 0.14	5	Yes
Goteberg	1.73	−0.17	−0.42, 0.08	0.17	0.05	−0.03, 0.14	1.90	−0.11	−0.38, 0.15	8	Yes
HIP	1.94	0.09[a]	−0.24, 0.39	No data	No data	No data	1.94	0.07	−0.24, 0.39	5	No
Malmö I	2.12	0.50	0.18, 0.81	0.27	0.23	0.11, 0.36	2.39	0.73	0.39, 1.07	8.8	No
NBSS I	1.68	0.12	−0.08, 0.32	0.09	0.13	0.07, 0.18	1.77	0.25	0.04, 0.46	13	No
NBSS II	2.38	0.04	−0.22, 0.31	0.06	0.21	0.14, 0.28	2.44	0.26	−0.02, 0.54	13	No
Edinburgh	1.93	0.11	−0.15, +0.38	0.05	0.13	0.07, 0.20	1.98	0.25	−0.00, 0.53	10	No

[a]All breast cancers. CI, confidence interval; DCIS, ductal carcinoma *in situ*; HIP, Health Insurance Plan; NBSS, National Breast Screening Study.

per 1,000 women years), with a slightly lower rate of invasive cancers (1.56 versus 1.73 per 1,000 women years) [7]. The relative risk of overall breast cancer incidence in the intervention arm relative to the control arm at follow-up of up to 14 years is reported as 0.98 (95% confidence interval 0.88 to 1.09).

The first Malmö trial invited women between 45 and 69 years old to five rounds of screening between 1976 and 1978. Women in the control arm were not invited for screening (although those in the youngest cohort were eventually invited in 1992/3). At a mean follow-up of 8.8 years, when screening was still continuing, the rates of invasive breast cancers in the intervention and control arms were 2.62 and 2.12 per 1,000 women years, respectively. The rates of DCIS were 0.50 and 0.27 per 1,000 women years, respectively [8]. It was estimated that 24% of the control arm had been screened, although most only once. No incidence data have been published on the subsequent (Malmö II) trial.

The Stockholm trial randomised about 60,000 women in 1981; there were two screening rounds using single-view mammography 28 months apart; during 1986 the control group was invited to a single screen. Uptake at first screen was 81% in the intervention arm and 77% in the control arm [9]. At the end of 1986 there was no difference in overall cancer incidence between the two arms (0.90 versus 0.91 per 1,000 women years). There was a slightly higher rate of DCIS (0.09 versus 0.06) and lower rate of invasive cancers (0.81 versus 0.85) in the intervention arm, but the differences were not significant.

Two trials conducted in Canada both used volunteer populations, resulting in high uptake. The NBSS II trial was designed to compare mammography plus physical examination (MP), with physical examination alone (PO) in women aged 50 to 64 years [10]. An initial excess of invasive cancers in the MP arm mainly disappeared with continued follow-up; at 13 years of follow-up, rates per 1,000 women years were 2.43 and 2.38 in the MP and PO arms, respectively [11]. Rates of DCIS were 0.28 and 0.06 per 1,000 women years, respectively. The NBSS I trial was designed to compare breast cancer mortality in women aged 40 to 49 years randomised to either screening by annual mammography, physical examination and instruction on breast self-examination (BSE) or a single physical examination and BSE instruction [12]. After 13 years of follow-up, the cumulative rates of DCIS were 0.22 and 0.09 per 1,000 women years in the screening and 'usual care' groups, respectively [13]. The rates of invasive breast cancers were 1.81 and 1.68, respectively, per 1,000 women years.

The Edinburgh trial recruited women aged 45 to 64 years into the initial cohort during 1978 to 1981, with randomisation by general practice. Women in the intervention arm were offered annual screening for 7 years, by mammography and physical examination every 2 years, and physical examination only in the intervening years. At 10 years of follow-up the incidence rates of invasive breast cancer were 2.04 and 1.93 per 1,000 women years in the intervention and control arms, respectively; rates of DCIS were 0.19 and 0.05, respectively, per 1,000 women years [14].

Figure 1
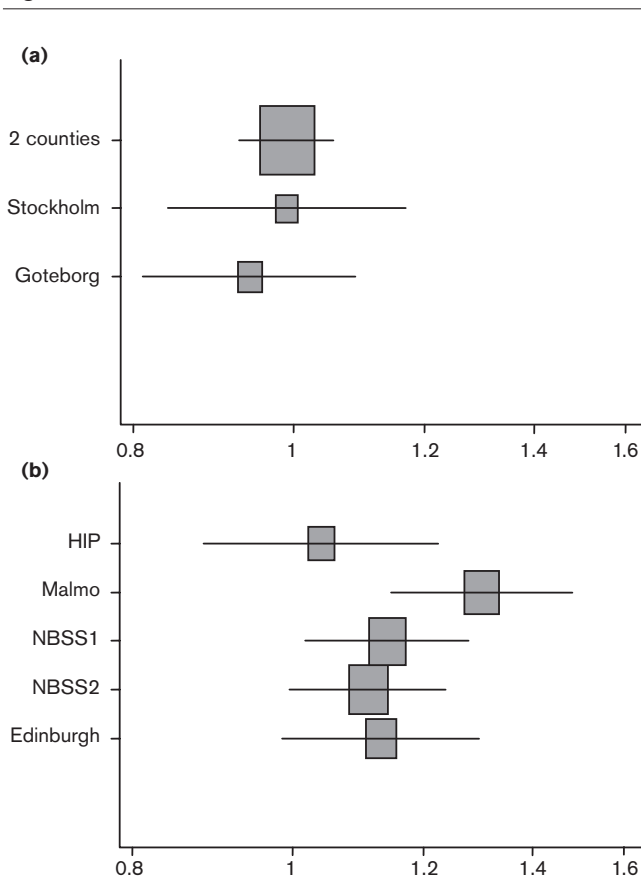
Relative incidence of all breast cancers; ratios of intervention arm to control arm. **(a)** Randomised controlled trials with screening in the control arm. **(b)** Randomised controlled trials without screening in the control arm. HIP, Health Insurance Plan; NBSS, National Breast Screening Study.

However, the cluster randomisation in this trial led to an imbalance in socio-economic status, reflected in all-cause mortality, which is likely to have resulted in an increased risk of breast cancer in the intervention arm. Adjustment for this altered the rate ratio of breast cancer mortality at 14 years of follow-up from 0.87 to 0.79 [15].

Summary of the trials
In Table 2 the trials are grouped according to whether the control group had been offered screening. For the three trials in which this occurred the absolute excess of all breast cancers in the intervention arm ranged from −0.02 to −0.11 per 1,000 women years; the ratio of the incidence in the intervention arm to that in the control arm was 0.94 to 0.99. For DCIS the absolute excess ranged from 0.05 to 0.11, and for invasive cancer from −0.84 to −0.17.

By contrast, in those trials in which the control group had not been offered screening, there was an excess of both invasive cancers and DCIS in the intervention arm, although in the

Malmö trial screening was still in progress at the time that rates were reported. The two Canadian trials are the most informative because they were conducted most recently and have 13 years of follow-up; these show an absolute excess of all breast cancers of 0.25 to 0.26 per 1,000 women years; the ratio of the incidence in the intervention arm to that in the control arm was 1.11 to 1.14.

UK trials of age and frequency
Two further trials in the UK have not yet reported mortality results. The 'age' trial is offering annual mammography from the age of 40 or 41 years to an intervention arm [16]; women in both arms will be invited as part of the national programme at ages 50 to 52 years. At the time of an interim analysis, when screening was still in progress, there was an 8% excess diagnosis of invasive breast cancers and a 17% excess of all breast cancers in the intervention arm in comparison with the control arm [17]. Once all women in both arms have been invited for screening in the national programme, any excess diagnosis should be as a result of incident screens.

The 'frequency' trial has compared annual versus three-yearly screening in women aged 50 to 64 years within the UK National Health Service breast screening programme [18]. Because all women had received a prevalent screen before randomisation, any excess diagnosis should be the result of more frequent screening (or difference in uptake). At 3 years of follow-up (that is, when both arms had been reinvited) a non-significant increase in breast cancers of 19% (13% invasive) was observed in the annual screening arm. Although the authors did not consider the difference to be real, they acknowledged a possible effect of increased diagnostic activity. The (31%) increase in DCIS was also non-significant.

Conclusion
This paper summarises the evidence for overdiagnosis in randomised trials of mammography, on the basis of comparisons of cumulative incidence, in the intervention and control arms of such trials. It is noted that differences in breast cancer incidence can also arise from bias in randomisation, and indeed equality of incidence has been used as evidence for a lack of such bias [7]. No mathematical modelling of the extent of overdiagnosis has been attempted in the present paper, because this will be the subject of a later paper in this series.

In trials in which the control group has not been offered screening, only once sufficient time has elapsed since the end of screening can overdiagnosis be estimated. If screening in the intervention arm is still continuing at the time incidence is reported, an increase in the intervention arm would be expected owing to advanced diagnosis, as observed in the Malmö trials. In the Canadian trials there is an 11 to 14% excess of all cancers in the intervention arm at 13 years of follow-up, largely of DCIS, suggesting the existence of overdiagnosis. However, in those recent trials in which the control group have been invited for screening, although there is a possible shift from invasive disease to DCIS, there is no evidence of overdiagnosis of all breast cancers as a result of incident screens.

Competing interests

The author(s) declare that they have no competing interests.

References

1. Yen MF, Tabar L, Vitak B, Smith RA, Chen HH, Duffy SW: **Quantifying the potential problem of overdiagnosis of ductal carcinoma in situ in breast cancer screening.** *Eur J Cancer* 2003, **39:**1746-1754.
2. Vainio H, Bianchini F: *Breast Cancer Screening: IARC Handbooks of Cancer Prevention.* Lyon: IARC Press; 2002.
3. Shapiro S, Venet W, Strax P, Venet L, Roeser R: **Ten- to fourteen-year effect of screening on breast cancer mortality.** *J Natl Cancer Inst* 1982, **69:**349-355.
4. Shapiro S: **Evidence on screening for breast cancer from a randomized trial.** *Cancer* 1977, **39:**2772-2782.
5. Tabar L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Grontoft O, Ljunquist U, Lundstrom B, Manson JC, Eklund G, *et al.*: **Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health & Welfare.** *Lancet* 1985, **i:**829-832.
6. Duffy SW, Tabar L, Vitak B, Day NE, Smith RA, Chen HHT, Yen MFA: **The relative contributions of screen-detected insitu and invasive breast carcinomas in reducing mortality from the disease.** *Eur J Cancer* 2003, **39:**1755-1760.
7. Bjurstam N, Bjorneld L, Warwick J, Sala E, Duffy SW, Nystrom L, Walker N, Cahlin E, Eriksson O, Hafstrom LO, *et al.*: **The Gothenburg Breast Screening Trial.** *Cancer* 2003, **97:**2387-2396.
8. Andersson I, Aspegren K, Janzon L, Landberg T, Lindholm K, Linell F, Ljungberg O, Ranstam J, Sigfusson B: **Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial.** *BMJ* 1988, **297:**943-948.
9. Frisell J, Lidbrink E, Hellstrom L, Rutqvist L-E: **Follow up after 11 years – update of mortality results in the Stockholm mammographic screening trial.** *Breast Cancer Res Treat* 1997, **45:**263-270.
10. Miller AB, Baines CJ, To T, Wall C: **Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years.** *Can Med Assoc J* 1992, **147:**1477-1488.
11. Miller AB, To T, Baines CJ, Wall C: **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:**1490-1499.
12. Miller AB, Baines CJ, To T, Wall C: **Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years.** *Can Med Assoc J* 1992, **147:**1459-1476.
13. 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.** *Ann Intern Med* 2002, **137:**E-305-E-315.
14. Alexander FE, Anderson TJ, Brown HK, Forrest AP, Hepburn W, Kirkpatrick AE, McDonald C, Muir BB, Prescott RJ, Shepherd SM: **The Edinburgh randomised trial of breast cancer screening: results after 10 years of follow-up.** *Br J Cancer* 1994, **70:**542-548.
15. Alexander FE, Anderson TJ, Brown HK, Forrest APM, Hepburn W, Kirkpatrick AE, Muir BB, Prescott RJ, Smith A: **14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening.** *Lancet* 1999, **353:**1903-1908.
16. Moss S: **A trial to study the effect on breast cancer mortality of annual mammographic screening in women starting at age 40. Trial Steering Group.** *J Med Screen* 1999, **6:**144-148.
17. Moss S, Waller M, Anderson TJ, Cuckle H (writing committee for the Trial Management Group): **Randomised controlled trial of mammographic screening in women from age 40: predicted mortality based on surrogate outcome measures.** *Br J Cancer* 2005, **92:**955-960.
18. The Breast Screening Frequency Trial Group: **The frequency of breast cancer screening: results from the UKCCCR Randomised Trial.** *Eur J Cancer* 2002, **38:**1458-1464.