

Review

Overdiagnosis and overtreatment of breast cancer

Overdiagnosis and overtreatment in service screening

Eugenio Paci¹ and Stephen Duffy²

¹Unit of Clinical and Descriptive Epidemiology, CSPO, Research Institute of the Tuscany, Via di SanSalvi, Florence, Italy

²Centre for Epidemiology, Mathematics and Statistics, Wolfson Institute of Preventive Medicine, Charterhouse Square, London EC1M 6BQ, UK

Corresponding author: Eugenio Paci, e.paci@cspo.it

Published: 10 November 2005

This article is online at <http://breast-cancer-research.com/content/7/6/266>

© 2005 BioMed Central Ltd

Breast Cancer Research 2005, **7**:266-270 (DOI 10.1186/bcr1339)

Abstract

Screening mammography has been shown to be effective for reducing breast cancer mortality. According to screening theory, the first expected consequence of mammography screening is the detection of the disease at earlier stages and this diagnostic anticipation changes the population incidence curve, with an observed increase in incidence rates at earlier ages. It is unreasonable to expect that the age-specific incidence will ever return to pre-screening levels or to anticipate a significant reduction of incidence at older ages immediately after the first screening round. The interpretation of incidence trends, especially in the short term, is difficult. Methodology for quantification of overdiagnosis and statistical modelling based on service screening data is not well developed and few population-based studies are available. The overtreatment issue is discussed in terms of appropriateness of effective treatment considering the question of chemotherapy in very early stages and the use of breast conserving surgery.

Introduction

The results of eight randomised clinical trials have shown screening mammography to be effective in reducing breast cancer mortality [1,2]. Evidence of efficacy was shown for women aged 50 years and over and service screening was implemented at national or regional levels in many countries [2,3]. The challenge today is to evaluate service screening in Europe to assess the outcome of the programmes [4] in terms of mortality and disease stage at diagnosis.

The aim of breast cancer screening has been shown to be achieved by the detection of cancer (*in situ* or invasive) at an earlier stage of the natural history of the disease and by the subsequent use of effective treatment in the early phase of natural history. According to screening theory, the first expected consequence of mammography screening, spontaneous or organised, is the detection of the disease at earlier stages and this diagnostic anticipation changes the population incidence curve, with an increase in incidence

rates at earlier ages. The shift of the curve at younger ages is expected to be more evident at the time of prevalence screening, but it will also continue over the subsequent rounds of the screening programme.

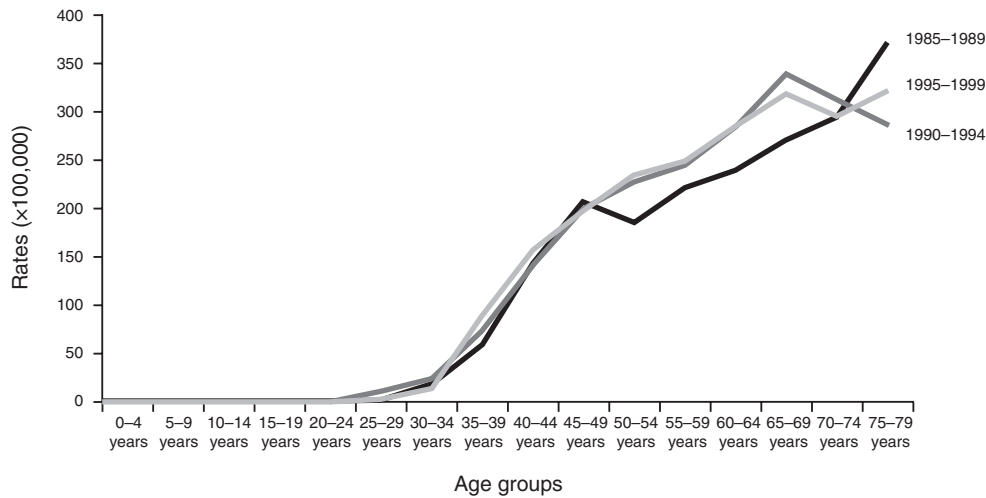
This excess in incidence should not be confused with overdiagnosis. Overdiagnosis has been defined as the detection of *in situ* or invasive breast cancers at screening that would have never clinically surfaced in the absence of screening. It is the combination of two causes: the natural history of the disease, that is, the low potential of progression of certain lesions; and the presence of competing causes of death, such that the detected disease would not have been diagnosed in the lifespan of the subject in the absence of screening.

In these terms, overdiagnosis is largely an epidemiological concept, because there is no marker today to classify a cancer as a pseudodisease. The possibility of overdiagnosis is inherent to the process of screening, however, and the quantification of such overdiagnosis is a current challenge to the epidemiological community.

Excess incidence

The excess in breast cancer incidence related to service screening with mammography has been described in several areas. For example, in Fig. 1, the population-based incidence in the city of Florence during the first and subsequent rounds of screening is compared with that of the pre-screening era, showing the excess for different age groups invited to screening. The screening epoch from 1990 onwards shows a higher incidence in the 50 to 69 years age group invited to screening and a lower incidence in the 75 to 79 age group.

In a recent paper, Zahl *et al.* [5] presented data comparing areas with or without service screening in Norway and in Sweden. They estimated that the incidence of invasive breast

Figure 1

Breast cancer incidence rates in the city of Florence by calendar period.

cancer in women aged 50 to 69 years increased by 54% in Norway and 45% in Sweden over the study period. They argued that as there was no corresponding, statistically significant decline in incidence at 70 to 74 years of age, there must be substantial overdiagnosis.

In the Norwegian counties in which screening started in 1996, there was no significant increase in incidence before screening and a subsequent increase in the 50 to 69 years age group after 1995. In the year 2000 (4 years after the start), the 70 to 74 years age group showed a non-significant 11% reduction in incidence. It is likely, however, that four years is insufficient time to see the full reduction in incidence in the post-screening age groups which are not targets for service screening. In Sweden, an increase in incidence was evident during the period 1971 to 1985; nationwide screening was gradually implemented after 1985. The incidence was not declining by the year 2000 for women aged 70 to 74 years, but a statistically significant reduction in incidence of 12% was shown for ages 75 to 79 years. This reduction was considered small by the authors, but a reduction of 12% in this age group represents a substantial number of the incidence at lower ages, in absolute terms. In Sweden, service screening has also been offered to 70 to 74 year old women in several areas. Thus, the conclusion of substantial overdiagnosis from these data may be unwarranted.

Olsen *et al.* [6] compared incidence in three Danish municipalities providing organised screening programmes with the rest of Denmark. They found a temporary increase in incidence corresponding to the first screen, followed by a return to levels close to those in the pre-screening period in two of the three municipalities. In the third municipality, a

small area of Copenhagen, the increase corresponded more to the second round of screening, possibly due to poor sensitivity at the first. The authors concluded that there was no serious overdiagnosis.

In the UK, McCann *et al.* [7] projected pre-screening trends in incidence into the screening epoch and found an excess incidence in the early 1990s in the screening 50 to 64 years age group and a deficit in incidence in the late 1990s in the 65 to 69 years age group. They found that accounting for the later deficit using the earlier excess was more complete if ductal carcinoma *in situ* (DCIS) cases were included.

Both Anttila *et al.* [8] in Finland and Fracheboud *et al.* [9] in the Netherlands observed increased incidences of breast cancer with the introduction of screening. Both groups noted, however, that these included underlying increases in incidence that were taking place in any case. In Finland and the Netherlands, screening programmes were introduced gradually and the excess incidence will, therefore, be spread over several years.

The research reviewed above points to a lead time effect as being responsible, at least in part, for the excess incidence observed with screening. This does not rule out overdiagnosis, which may also be partly responsible for the excess. The challenge is, therefore, to consider the possible multiple causes of excess incidence in screened cohorts, and to estimate the extent of overdiagnosis, taking the other causes into account.

Quantification of overdiagnosis

One should note first that a fixed, discrete cohort, the kind of population studied in a randomised clinical trial, is extremely

different from a dynamic population, where several ageing cohorts and newcomers are monitored for varying periods of time. In the HIP study, the cumulative breast cancer incidence in the control group was observed to catch up with the study group when screening stopped, and this was confirmed by statistical modelling [10]. This sort of analysis is not available in the service screening setting.

The possible reasons for an observed excess of incidence in the service screening context are:

1. In almost all countries, incidence of breast cancer was increasing before screening programmes were introduced.
2. There is inevitably a surge in incidence at the time of introduction of screening, due to the prevalence screening of a large population. The size of the surge will depend on how long it takes to complete coverage. Most of this is composed of anticipated tumours that would have occurred in any case in the following five years.
3. A continued surge at the lower end of the age range for screening as the women reaching the lower age limit have a prevalence screen.
4. There will be a shift in the age-incidence curve due to lead time. If the screening programme is achieving an average lead time of three years, say, then we will observe age 53 incidence at age 50, age 54 incidence at age 51, and so on.
5. Depending on the temporal pattern of screening activity, there may also be periodic excesses due to anticipated tumours from incidence screening, balanced by periodic deficits in clinical cancer incidence between such screens.
6. There may also be overdiagnosis.

Reason 3 and 4 will remain active as long as the screening programme is in place. Thus it is unreasonable to expect that the age-specific incidence will ever return to pre-screening levels. Also, it should be noted that a deficit in incidence above the age limits for screening can only occur in cohorts that have actually been through the screening programme. One cannot expect, therefore, to observe immediately after the first round of screening a significant reduction in incidence at older ages. So the interpretation of incidence trends, especially in the short term, is difficult.

Methodology for quantification of overdiagnosis is not yet well developed but there are some examples in the literature. In the city of Florence, service screening started at the beginning of 1990, offering high-quality mammography every 2 years to women aged 50 to 69 years. An evaluation of overdiagnosis due to service screening was performed after 10 years [11]. This compared incidence in the period 1990 to 1999 with that expected in the absence of screening, but with adjustment for that part of the excess that was due to lead time alone, and not to overdiagnosis.

All breast cancer cases were partitioned by diagnostic method (screen detected versus clinically detected).

Considering the mean sojourn time estimate of 3.7 years for breast cancer cases and an exponential distribution of the sojourn time, the probability that a screen detected case would have remained asymptomatic up to the end of the study period was calculated. The sum of the probabilities of clinical incidence of screen detected cases within the study period, added to the observed clinically detected breast cancer cases was compared with the expected incidence in the absence of screening. We estimated the overdiagnosis first for invasive tumours only, then for all cancers including DCIS. Overdiagnosis of invasive breast cancer cases was estimated at 2% (non-significant). The inclusion of *in situ* cases in the model increased the risk of overdiagnosis to 5% (statistically significant), supporting the view that DCIS could be majorly responsible for the excess. For the Florence data, we estimated an excess of incidence for women aged 50 to 84 years at about 15% at short term follow up and 11% at long term. The excess corrected for lead time was 12% at short term and 2% at long term. The implications are that an evaluation of overdiagnosis in the short term of service screening – when the prevalence screening is mainly ongoing – can be misleading, and that when long term data are available, correction for lead time yields a much more modest estimate of overdiagnosis. These are probably overestimates of overdiagnosis, using the incidence rates of 1985 to 1989 without consideration for the increasing trend in breast cancer incidence.

More detailed modelling of overdiagnosis

The possibility of overdiagnosis has been a cause for concern, in particular with regard to the occurrence of carcinoma *in situ* [12]. Detection of *in situ* lesions is a feature of mammography screening and natural history and the probability of progression of this kind of lesion is not fully understood.

Yen *et al.* [13] reviewed the rates of DCIS and invasive cancers from the Swedish Two County Study and from various service screening programmes to: derive tentative estimates of DCIS detection rates that should be typically observed; describe the typical range of absolute detection rates of DCIS; and estimate the proportion of DCIS detected at screening that truly represents overdiagnosis.

They used a six-state Markov Model that fitted the data reasonably. In their conclusions, 37% of DCIS cases at the prevalence screening were estimated to be non-progressive; the corresponding figure at incidence screens was 4%. On the basis of the estimates, a woman attending for prevalent screening has a 1 in 3,300 chance of being diagnosed with a non-progressive DCIS. The probability of being diagnosed with a progressive DCIS or invasive carcinoma was 1 in 175. They concluded that there was an element of overdiagnosis of DCIS in mammographic screening; however, this element is modest in comparison with the likely benefit of mammography. The increasing number of DCIS cases poses the challenge to

therapy to develop treatment protocols taking into account the potential aggressiveness of the detected lesion.

The increased incidence of DCIS has been found in randomized controlled trials and service screening to be at least partly balanced by a later reduction in invasive cancer incidence [7,14].

This evidence is in contrast with the conclusion of the International Agency for Research on Cancer (IARC) expert group, which states "studies of populations...provide no evidence of a decrease of incidence of invasive cancers" [15]. The expert group conclusion was based on descriptive trends from SEER, the US cancer registry network, without any estimate of the impact of lead time in the excess of the observed carcinoma *in situ* incidence. The above results suggest that where overdiagnosis is explicitly estimated, taking into account other causes of increased incidence, the estimate is usually small. There is, however, a need for further quantification of overdiagnosis from other screening programmes and more detailed models.

Overtreatment

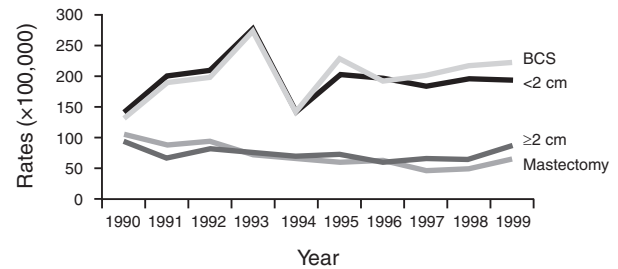
Overtreatment may be considered to occur in two ways (although these are related). Firstly, if there are overdiagnosed cases, any treatment of these is unnecessary. As one cannot tell when a breast cancer is diagnosed whether it would or would not progress in the absence of treatment, some treatment, particularly excision, is inevitable. As noted above, empirical estimates of overdiagnosis, which take into account effects of lead time, suggest that the proportion of overdiagnosed tumours is small, but there is no room for complacency. To minimise the burden of overtreatment of this kind, research must continue on tumour biology to further quantify the aggressive potential of screen-detected cancers, notably DCIS. In the meantime, treatment should be individually decided on the basis of the aggressive potential, as detected, for example, by stage or grade of the lesion diagnosed.

The second major manifestation of overtreatment is the administration of more aggressive therapies than is necessary to 'true' but very early stage cancers [16,17]. The one-size-fits-all philosophy of cytotoxic chemotherapy for all invasive lesions is inappropriate when one considers that node negative tumours smaller than 10 mm have survival rates in excess of 90% without chemotherapy. In such cases the

This article is part of a review series on *Overdiagnosis and overtreatment of breast cancer*, edited by Nick E Day, Stephen Duffy and Eugenio Paci.

Other articles in the series can be found online at http://breast-cancer-research.com/articles/review-series.asp?series=BCR_Overdiagnosis

Figure 2



Breast conserving surgery and tumour size (<pt2) rates in Florence service screening for the period 1990 to 1999.

benefits and risks to life of cytotoxic agents may actually have a negative balance. The first response to this problem should be to tailor the treatment to the tumour.

There is evidence from the Florence programme that this is occurring in terms of surgery [18]. With the introduction of the screening programme, absolute numbers of breast conserving surgery episodes increased and absolute numbers of mastectomies fell (Fig. 2). The rates of the two types of operations paralleled very closely the rates of early and late stage tumours. There is an onus on the oncological community to ensure that surgical treatment and adjuvant therapies are administered on the basis of tumour characteristics.

Conclusion

Overdiagnosis in breast cancer screening is probably a minor phenomenon, but further quantification is needed from multiple service screening programmes. Estimates of overdiagnosis should take into account other causes of observed excess incidence, such as lead time. The large numbers of early stage tumours being diagnosed in screening programmes suggest that care should be taken to minimise harm from over-aggressive therapy for such lesions.

Competing interests

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

References

- Smith RA, Duffy SW, Gabe R, Tabar L, Yen AMF, Chen THH: **The randomized trials of breast cancer screening: what have we learned?** *Radiol Clin N Am* 2004, **42**:793-806.
- Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskiold B, Rutqvist LE: **Long-term effects of mammography screening: updated overview of the Swedish randomised trials.** *Lancet* 2002, **359**:909-919. Erratum: *Lancet* 2002, **360**:724.
- Ballard-Barbash R, Klabunde C, Paci E, Broeders M, Coleman EA, Fracheboud J, Bouchard F, Rennert G, Shapiro S: **Breast cancer screening in 21 countries: delivery of services, notification of results and outcomes ascertainment.** *Eur J Cancer Prev* 1999, **8**:417-426.
- Cuzick J: **Breast cancer screening - time to move forward.** *Breast* 2002, **11**:209-210.
- Zahl PH, Strand BH, Maehlen J: **Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study.** *Br Med J* 2004, **328**:921-924.

6. Olsen AH, Jensen A, Njor SH, Villadsen E, Schwartz W, Vejborg I, Lyng E: **Breast cancer incidence after the start of mammography screening in Denmark.** *Br J Cancer* 2003, **88**:362-365.
7. McCann J, Treasure P, Duffy S: **Modelling the impact of detecting and treating ductal carcinoma in situ in a breast screening programme.** *J Med Screen* 2004, **11**:117-125.
8. Anttila A, Koskela J, Hakama M: **Programme sensitivity and effectiveness of mammography service screening in Helsinki, Finland.** *J Med Screen* 2002, **9**:153-158.
9. Fracheboud J, Otto SJ, van Dijk JAAM, Broeders MJM, Verbeek ALM, de Koning HJ and National Evaluation Team for Breast Cancer Screening: **Decreased rates of advanced breast cancer due to mammography screening in The Netherlands.** *Br J Cancer* 2004, **91**:861-867.
10. Boer R, Warmerdam P, de Koning H, van Oortmarssen G: **Extra incidence caused by mammographic screening.** *Lancet* 1994, **343**:979.
11. Paci E, Warwick J, Falini P, Duffy SW: **Overdiagnosis in screening: is the increase in breast cancer incidence rates a cause for concern?** *J Med Screen* 2004, **11**:23-27.
12. Ernster VL, Ballard-Barbash R, Barlow WE, Zheng Y, Weaver DL, Cutter G, Yankaskas BC, Rosenberg R, Carney PA, Kerlikowske K, Taplin SH, Urban N, Geller BMR: **Detection of ductal carcinoma in situ in women undergoing screening mammography.** *J Natl Cancer Inst* 2002, **94**:1546-1554.
13. 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.
14. Duffy SW, Tabar L, Vitak B, Day NE, Smith RA, Chen HH, Yen MF: **The relative contributions of screen-detected in situ and invasive breast carcinomas in reducing mortality from the disease.** *Eur J Cancer* 2003, **39**:1755-1760.
15. International Agency for Research on Cancer Expert Group: **Breast Cancer Screening.** In *IARC Handbooks of Cancer Prevention*. Vol 7. Lyon, France: IARC; 2002:248.
16. Barth RJ, Gibson GR, Carney PA, Mott LA, Becher RD, Poplack SP: **Detection of breast cancer on screening mammography allows patients to be treated with less-toxic therapy.** *Am J Roentgenol* 2005, **184**:324-329.
17. Duffy SW, Tabar L: **Screening for breast cancer.** *Lancet* 1995, **346**:852.
18. Paci E, Duffy SW, Giorgi D, Zappa M, Crocetti E, Vezzosi V, Bianchi S, Cataliotti L, del Turco MR: **Are breast cancer screening programmes increasing rates of mastectomy? Observational study.** *Br Med J* 2002, **325**:418.