Review

Overdiagnosis and overtreatment of breast cancer

Microsimulation modelling estimates based on observed screen and clinical data

Harry J de Koning, Gerrit Draisma, Jacques Fracheboud and Arry de Bruijn

Erasmus MC, Department of Public Health, PO Box 1738, 3000 DR Rotterdam, The Netherlands

Corresponding author: HJ de Koning, h.dekoning@erasmusmc.nl

Published: 21 December 2005
This article is online at http://breast-cancer-research.com/content/8/1/202

Breast Cancer Research 2006, **8:**202 (doi:10.1186/bcr1369)

Abstract

There is a delicate balance between the favourable and unfavourable side-effects of screening in general. Overdiagnosis, the detection of breast cancers by screening that would otherwise never have been clinically diagnosed but are now consequently treated, is such an unfavourable side effect. To correctly model the natural history of breast cancer, one has to estimate mean durations of the different pre-clinical phases, transition probabilities to clinical cancer stages, and sensitivity of the applied test based on observed screen and clinical data. The Dutch data clearly show an increase in screen-detected cases in the 50 to 74 year old age group since the introduction of screening, and a decline in incidence around age 80 years. We had estimated that 3% of total incidence would otherwise not have been diagnosed clinically. This magnitude is no reason not to offer screening for women aged 50 to 74 years. The increases in ductal carcinoma *in situ* (DCIS) are primarily due to mammography screening, but DCIS still remains a relatively small proportion of the total breast cancer problem.

Introduction

Breast cancer screening has been effective in reducing breast cancer mortality. Both randomised controlled trials and nation-wide screening programmes have shown a roughly 25% reduction in disease-specific mortality for women aged 50 years and over invited to screening [1-4]. This benefit applies to the group as a whole, but at the individual level it is impossible to determine who will actually benefit or who will receive more harm than benefit from such a programme: there is a delicate balance between the favourable and unfavourable side-effects of screening in general [5]. For example, detecting breast cancers by screening that would otherwise never have been clinically diagnosed, but are now treated, is such an unfavourable side effect. Because of lead time and length-biased sampling, the screening test will generally detect more early lesions with possibly different biological behaviour and also more slowly growing tumours, especially ductal carcinoma *in situ* (DCIS). Screening at older ages will, due to existing co-morbidity, lead to the detection of clinically relevant diseases; however, women may not necessarily benefit because they more often die due to other diseases.

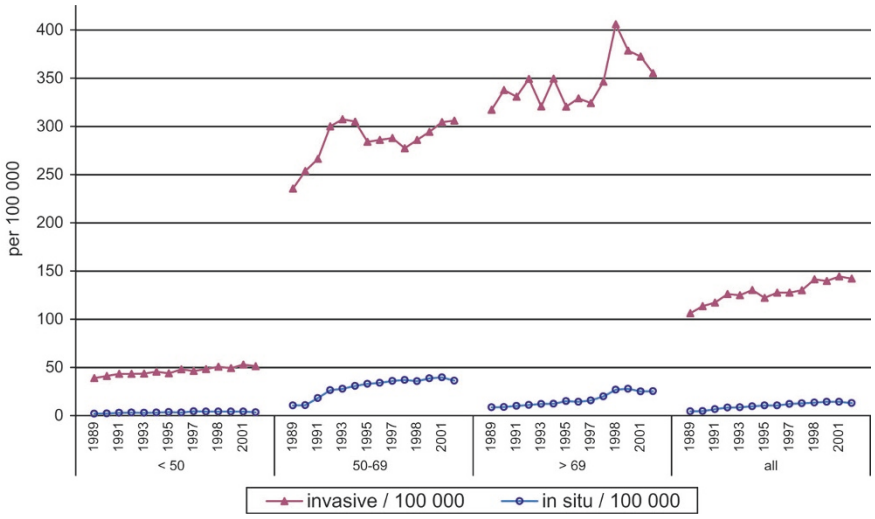
This paper presents quantitative estimates of overdiagnosis in breast cancer screening based on microsimulation modelling, with special emphasis on DCIS. In this study, overdiagnosis is defined as diagnosing cancers that would not have been diagnosed clinically if there were no screening programme.

Observations

Starting to screen a population systematically for breast cancer will lead to the detection of cancers about three to four years earlier than without such an approach [6]; therefore, the number of detected cancers at the population level is expected to increase. Because screening is continued every consecutive year, this number is higher, and remains so, than if there were no systematic screening. Figure 1 shows the Dutch national data since 1989, when screening was gradually being implemented [7,8]. For women aged 50 to 69 years, implementation took place in the period 1990 to 1997. After an initial increase of around 30%, incidence in the 50 to 69 year old age group stabilized at 16% higher than without screening. Furthermore, the last years of screening have resulted in an additional 10% increase, probably due to more referrals and better screening performance. From 1999 onwards in the Netherlands women aged 70 to 74 were also invited for screening. Compared to the year 1989, the number of breast cancers diagnosed each year has increased by 40%. Proportionally, this increase is largest for DCIS. Figure 2a,b shows the increase in Dutch hospital admissions for non-invasive breast cancer in the years 1990 to 1992 (at the start of nation-wide screening) in municipalities that had started screening compared to municipalities that had not; in the age group invited for screening (at that time 50 to 69 year olds), the increase was

DCIS = ductal carcinoma *in situ*.

Figure 1



Age-specific incidence of breast cancer (invasive/non-invasive) from 1989 to 2002 in the Netherlands [7,8].

3 to 5-fold (Fig. 2a). Strikingly, there was also an increase outside the screening municipalities (Fig. 2b). Non-invasive breast cancer, however, still accounted for only 4% of the total incidence [8].

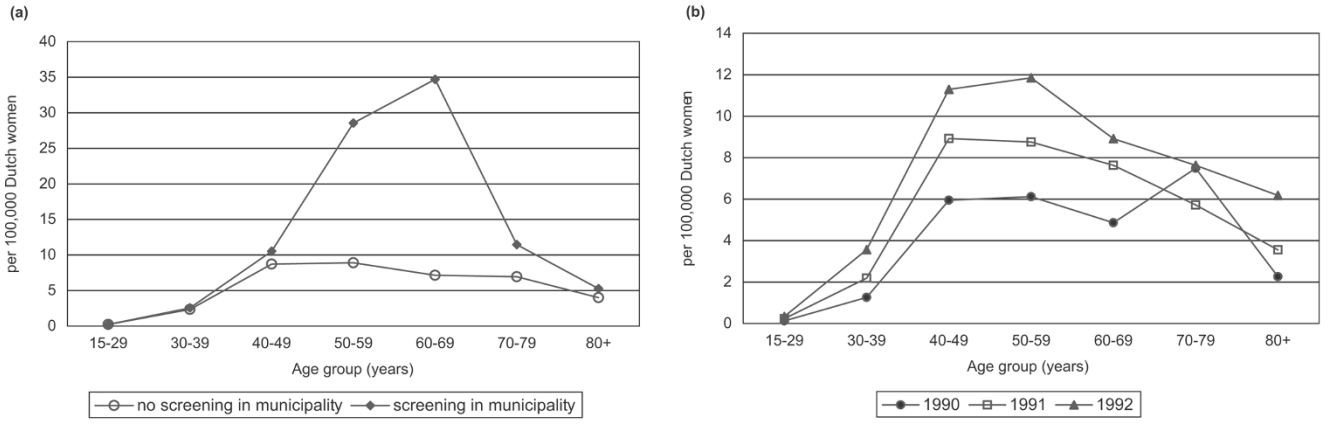
Modelling

These increases in incidence represent real overdiagnosis to only a limited extent. From the observed rates, one can not easily determine to what extent overdiagnosis is involved because screening is still being continued. In these circumstances, modelling of the natural history of breast cancer and its early lesions, and what screening is estimated to depict, is crucial and provides a 'best guess'. Using the microsimulation model MISCAN [6,9], we first simulate individual life histories for women in the absence of screening and then assess how these histories would change as a consequence of a screening programme. The natural history is modelled as a progression from no breast cancer through pre-clinical disease (DCIS, T1a, 1b, 1c, T2+) to clinical disease (same stages). From a given pre-clinical state, a cancer may be detected by screening or become clinically apparent or, if undiagnosed, progress to the next pre-clinical state. To correctly model this natural history of breast cancer for women in a certain age group, one has to estimate mean durations of the different pre-clinical phases, transition probabilities, and sensitivity of the applied test [10]. Basically, one therefore needs data from two sources: observed screen and clinical data. These data include clinical incidence data by age and stage in the situation without screening, data on screen-detected cancers by stage, screening round (and interval) and age, and corresponding clinical incidence data when screening is being implemented [11]. Although the observed data can often be explained by a small range of parameters (e.g., a somewhat higher sensitivity and shorter mean duration of the stage may also result in a good fit), by having more detailed data from several screening rounds, by screening different age groups and/or by using different screening intervals, best parameters often fall into a smaller range [12]. In the Netherlands, such detailed data have been used: in the past using pilot data [6], and more recent data from the annual monitoring by the National Evaluation Team for Breast cancer screening [7].

The fit of the model to the breast cancer screening pilot data [6,12], as well as to the Dutch nation-wide data [9], has been reported as quite satisfactory.

We also used the MISCAN approach to analyse the results of the Health Insurance Plan trial study. These comparisons show the potential power of modelling: the parameter values for the invariant part of the natural history of pre-clinical breast cancer are indeed the same, whereas the increase in the sensitivity reflects the improvement in mammography. Taking the obvious differences between HIP and Nijmegen (one of the two Dutch pilot studies) into account, the model shows that there is a good correspondence between the screening data from these studies. The findings about the duration of pre-clinical disease and the sensitivity of screening can be compared with results from other modelling approaches. Day and colleagues [13] applied this model to data from Utrecht (the other Dutch pilot study). The study reports a good fit of the model (chi-square of 7.2 and 7 degrees of freedom) when assuming a sensitivity of 99% and a mean duration of 2.8 years. It is not indicated exactly what data from Utrecht were used, but it is clearly a less detailed subset of the data than we used for testing model assumptions. An adapted version of the Day and Walter model was applied to the Nijmegen data [14]. In general, the

Figure 2



Hospital admissions for non-invasive breast cancer in the early years (1990 to 1992) of screening. **(a)** Admissions in Dutch municipalities with screening compared to those with no screening. **(b)** Admissions per calendar year in municipalities without screening.

estimated parameters are comparable to the values found with the MISCAN approach presented here, especially regarding the age-dependency of the estimated duration of the preclinical stage. The reported average duration is somewhat shorter, however, for example, 2.5 years in the 50 to 64 year old age group.

Data on the natural history at older ages have been very limited, but are slowly emerging now that the Dutch programme includes women aged 70 to 74 years [15]. Data on the natural history of DCIS are scarce [16], but parameters concerning the screen-detectable pre-clinical period can be estimated, based on the aforementioned data.

In our first analyses, we have assumed that 10% of invasive breast cancers are preceded by a screen-detectable DCIS phase and that the chance of progressing to invasive cancer or clinical DCIS is almost 90% in the long term. Recent data from randomised treatment trials support a high progression rate in the long term [17]. The observed screen data are then consistent/compatible with a mammography sensitivity of 40% and a mean screen-detectable duration of 5 years.

Results on overdiagnosis

Figure 3a,b shows the model-estimated changes in breast cancer incidence (by age) in the Netherlands in a programme for women aged 50 to 74 years screened every 2 years (assuming an 80% attendance rate), compared to no screening. Incidence rises at the starting ages, because all young women have (in principle) never been screened before, which means that cancers are detected that have already progressed over time through the pre-clinical stages. Figure 3a,b also clearly depicts the true extent of over-diagnosis. Because of the earlier detection, cancers that would have surfaced at ages 75 to 85 years have now been detected earlier; clinical incidence at these ages must, therefore, be lower. In Fig. 3a,b, the difference between the left area (extra cancers detected by screening) and the right area (less cancers diagnosed clinically) represents overdiagnosis. We had estimated this to be 3% of the total incidence, or 8% of screen-detected cancers. The Dutch data clearly show the decline in incidence at around age 80 years. The higher incidence than expected by the model around ages 55 to 65 years in 2002 (as estimated before the Dutch programme started) confirms the better screening performance in more recent years. It also illustrates the difficulty of estimating overdiagnosis in a situation where a nation-wide screening programme is being implemented.

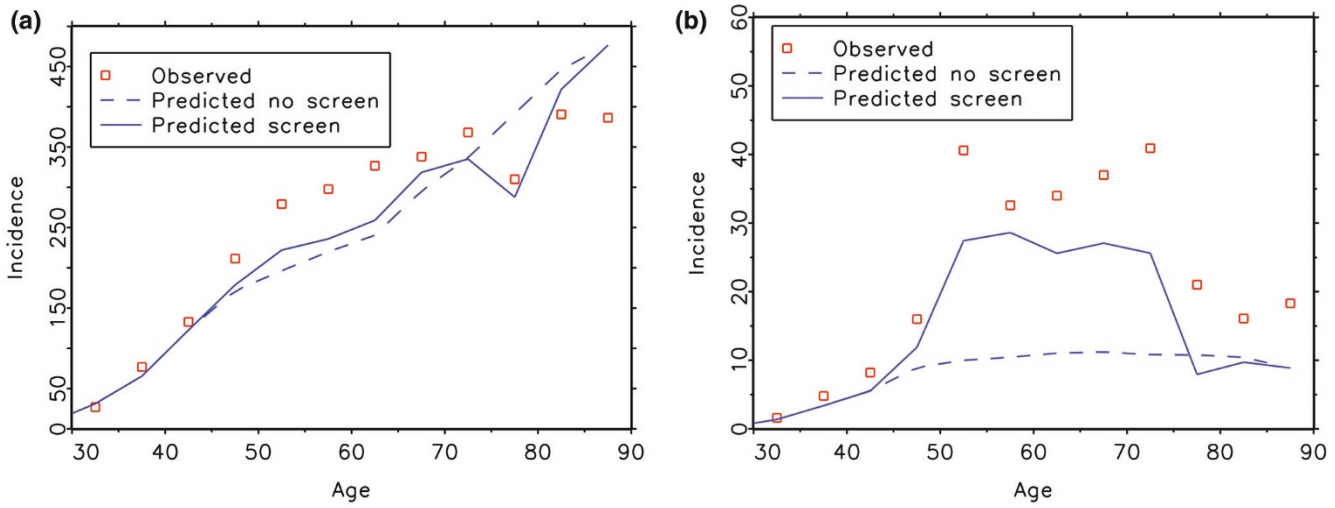
During the first years of screening, the increase in newly diagnosed cases in the age group invited to screening will not yet be reflected by a decrease in incidence at older ages as these are different cohorts of women. In the later years of screening, the increase in newly diagnosed cases and the decrease in incidence should be at a steady state, although this isn't always the case, because of other changes in the screening programme.

Figure 3b shows that the change in DCIS detection is especially striking, although some lesions would have progressed to an invasive disease (not shown in the figure). The amount of overdiagnosis, the increase in primary surgery/

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 3



Model estimations of breast cancer incidence. Predicted and observed rates are shown for **(a)** total incidence and **(b)** DCIS incidence in the 50 to 74 year old age group with screening every 2 years (assuming 80% attendance rate) or without screening in the Netherlands in 2002 [8].

radiotherapy, and the longer time frame since diagnosis has to be weighed against the favourable effects of screening: about 750 breast cancer deaths prevented per year (16%), reduction of treatments for advanced disease and its consequences for quality of life, and 15 life-years gained if dying from breast cancer has been prevented. We consider this to be a very acceptable balance at the population level [18].

Conclusion

Overdiagnosis is inherent to screening. The crucial issue is the extent to which it happens and what the consequences are for the population involved. This then has to balance with the favourable effects of screening in order to be able to decide on an appropriate screening policy. In breast cancer screening, overdiagnosis is not negligible but is relatively limited. The increases in DCIS are primarily due to mammography screening, but they remain a relatively small proportion of the breast cancer problem. The screen data observed in this study provide workable assumptions on the natural history of DCIS and do not lead to a major difference in conclusions regarding overdiagnosis. More and more women with DCIS are being treated by breast conservation, and in the Netherlands screen-detected DCIS is more often treated by conservation than clinically diagnosed DCIS. Categorisation of DCIS lesions into high risk versus low risk lesions (by screening) is still urgently needed.

Competing interests

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

References

1. Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE: **Long-term effects of mammography screening: updated overview of the Swedish randomised trials.** *Lancet* 2002, **359:**909-919. Erratum in *Lancet* 2002, **360:**724.
2. Otto SJ, Fracheboud J, Looman CWN, Broeders MJM, Boer R, Hendriks JHCL, Verbeek ALM, de Koning HJ and the National Evaluation Team for Breast Cancer Screening: **Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review.** *Lancet* 2003, **361:**1411-1417.
3. Olsen AH, Njor SH, Vejborg I, Schwartz W, Dalgaard P, Jensen MB, Tange UB, Blichert-Toft M, Rank F, Mouridsen H, Lynge E: **Breast cancer mortality in Copenhagen after introduction of mammography screening: cohort study.** *Br Med J* 2005, **330:** 220-224.
4. Rijnsburger AJ, van Oortmarssen GJ, Boer R, Draisma G, To T, Miller AB, de Koning HJ: **Mammography benefit in the Canadian National Breast Screening Study-2: a model evaluation.** *Int J Cancer* 2004, **110:**756-762.
5. de Koning HJ: **Assessment of nationwide cancer-screening programmes - Commentary.** *Lancet* 2000, **355:**80-81.
6. van Oortmarssen G: *Evaluation of mass screening for cancer: a model-based approach.* PhD Thesis. Department of Public Health, Erasmus Universiteit, Rotterdam; 1995.
7. Fracheboud J, Otto S, Dijck JAAM van, Broeders MJM, Verbeek ALM, de Koning HJ, and the 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.
8. Netherlands Cancer Registry: **Trends in breast cancer incidence 2002.** Utrecht; 2005.
9. Groenewoud JH, Pijnappel RM, van den Akker-van Marle ME, Birnie E, Buijs-van der Woude T, Mali WPTM, de Koning HJ, Buskens E: **Cost-effectiveness of stereotactic large-core needle biopsy for nonpalpable breast lesions compared to open-breast biopsy.** *Br J Cancer* 2004, **90:**383-392.
10. Boer R, Plevritis S, Clarke L: **Diversity of model approaches for breast cancer screening: a review of model assumptions by the Cancer Intervention and Surveillance Network (CISNET) Breast Cancer Groups.** *Stat Meth Med Res* 2004, **13:**1-14.
11. Day NE, Walter SD: **Simplified models of screening for chronic disease: estimation procedures from mass screening programmes.** *Biometrics* 1984, **40:**1-14.
12. van Oortmarssen GJ, Boer R, Habbema JD: **Modelling issues in cancer screening.** *Stat Methods Med Res* 1995, **4:**33-54.
13. Day NE, Walter SD, Tabar L, Fagerberg CJG, Collette HJA: **The sensitivity and lead time of breast cancer screening: a comparison of the results of different studies.** In *Screening for Breast Cancer.* Edited by Day NE, Miller AB. Toronto: H Huber Publishers; 1988:105-109.

14. Verbeek ALM, Straatman H, Hendriks JHCL: **Sensitivity of mammography in Nijmegen women under age 50: some trials with the Eddy model.** In *Screening for Breast Cancer*. Edited by Day NE, Miller AB. Toronto: H Huber Publishers; 1988:29-32.
15. Fracheboud J, Groenewoud JH, Boer R, Draisma G, de Bruijn AE, Verbeek ALM, de Koning HJ: **The upper age limit of 75 for breast cancer screening: results from the Dutch nation-wide screening programme.** *Int J Cancer*, in press.
16. 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.
17. Julien JP, Bijker N, Fentiman IS, Peterse JL, Delledonne V, Rouanet P, *et al.*: **Radiotherapy in breast-conserving treatment for ductal carcinoma in situ: first results of the EORTC randomised phase III trial 10853. EORTC Breast Cancer Cooperative Group and EORTC Radiotherapy Group.** *Lancet* 2000, **355:**528-533.
18. de Haes JC, de Koning HJ, van Oortmarssen GJ, van Agt HM, de Bruyn AE, van der Maas PJ: **The impact of a breast cancer screening programme on quality-adjusted life-years.** *Int J Cancer* 1991, **49:**538-544.