Predicting the Benefits and Harms of Breast Cancer Screening: Current debates and future directions

Rianne de Gelder

ISBN 978-94-6169-204-7

Predicting the Benefits and Harms of Breast Cancer Screening: Current debates and future directions
Thesis, Erasmus University

Design: Christine de Gelder
Lay-out and print: Optima Grafische Communicatie, Rotterdam, The Netherlands

The studies reported in this thesis were funded by the National Institute for Public Health and the Environment - Centre for Population Research (RIVM - CvB) (Chapters 2, 3, 6 and 7), the Netherlands Organisation of Health Research and Development (ZonMw) (Chapter 4), the Swiss Cancer League and Oncosuisse (Chapter 5).

This thesis financially supported by the Department of Public Health Erasmus MC and the Erasmus University Rotterdam.

Predicting the Benefits and Harms of Breast Cancer Screening: Current debates and future directions

Het voorspellen van de gunstige en ongunstige effecten van borstkankerscreening: Actuele debatten en toekomstige ontwikkelingen

Proefschrift

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van rector magnificus

Prof.dr. H.G. Schmidt

en volgens het besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 4 april 2012 om 15.30 uur

door
Rianne de Gelder
geboren te Ridderkerk

PROMOTIECOMMISSIE

Promotor:	Prof.dr. H.J. de Koning
Overige leden:	Prof.dr. M.G.M Hunink
	Prof.dr. ir. F.E. van Leeuwen
	Prof.dr. M. Rutten-van Mölken

CONTENTS

General introduction

1.1 BREAST CANCER SCREENING IN THE NETHERLANDS

Breast cancer in the Netherlands

Breast cancer is the most common cancer among women in Western countries.[1] Presently, women in the Netherlands have a 1 : 7 chance of developing breast cancer during their lifetime.[2] This means that in 2008, almost 15,000 women were newly diagnosed with the disease.[3] The incidence of breast cancer in the Netherlands is among the highest in Europe.[1, 4] In 2006, for instance, the age-standardized incidence rate was 128 per 100,000 woman-years. As a comparison, the average in Europe was 94.3 per 100,000 woman-years.[1] Although the probability of cure has improved over the last decennia, for a third of all women with breast cancer, the disease will be fatal.[2] This makes breast cancer the most common cause of cancer death in women in Europe.[1] The age-standardized mortality rate in the Netherlands in 2006 was 29.8 per 100,000 woman-years, *versus* 26.0 per 100,000 woman-years in Europe.[1]

Rationale for breast cancer screening

Breast cancer develops as a single malignant cell with uncontrolled cell growth to a tumour of several millimeters or centimeters in diameter. At some point in time, the tumour may reach a size at which it becomes symptomatic. The larger the size, the less likely it is that the tumour can be cured.[5–8] It is therefore thought that by diagnosing cancer at an earlier phase, for instance by screening, the probability of survival can be increased. Several methods for early detection of breast cancer exist: breast self examination, examination by a clinician or a nurse ('clinical breast examination'), MRI or ultrasonography. Mammography, which involves one or more X-ray images of the breasts, is considered the best tool for examining postmenopausal women with an average risk for the disease,[9] because it can reach a high sensitivity (>70%) and specificity (>95%) when it is applied on a large scale. At the same time, costs are moderate: 50 euro per screening examination in the Netherlands.[10] The effects of breast cancer screening using mammography were therefore studied in several randomized controlled trials (RCTs), with various screening ages and intervals. Screening women aged 50 and older resulted in statistically significant reductions in breast cancer mortality, of between 25%–30% in those women that were randomized in the screening arms of the trials.[11–12] Soon after the first positive trial outcomes, two pilot projects with large-scale mammography screening were started in the Netherlands, which also showed substantial reductions in breast cancer mortality in screened women.[13–14] A cost-effectiveness analysis showed a favourable balance between screening costs and potential life years gained.[15] Based on these findings, mammography screening was implemented in the Netherlands and other western countries.

Population-based breast cancer screening in the Netherlands

Nation-wide breast cancer screening in the Netherlands started around 1990. The programme originally targeted women in the age group 50–69 years, but from 1998 onwards, women aged 70–74 years are also invited to have a biennial mammogram. Since its start, the screening performance has been carefully monitored and evaluated by the National Evaluation Team for Breast cancer screening (NETB). Screening can only be effective if certain criteria are met, such as a sufficiently high attendance-, referral- and detection rate, a favourable tumour stage distribution of screen-detected compared to clinically diagnosed breast cancers, and a relatively low rate of interval cancers.

In the 20 years that the screening programme is now running, more than 11 million screening examinations have been performed. Annually, over 1 million women are invited to participate; 82% of them actually attend the programme. In 2007, 1.8% of all screened women were referred for further diagnostic assessment; in a third of these cases, a cancer was diagnosed.[10] This means that in 2007 approximately 5,000 tumours were detected by screening. Of these, 15.1% were ductal carcinoma *in situ* ('DCIS') and 8% were in a poor prognostic stage, being larger than 20 mm with positive lymph nodes ('T2+N+'). Since 1990, the fraction T2+N+ tumours of all diagnosed breast cancers had decreased. At the same time, interval cancer rates remained more or less stable over time: at first and subsequent screening rounds respectively 1.3 per 1,000 and 1.2 per 1,000 woman-years (2004).[10] These favourable results could forecast a reduction in breast cancer mortality.

1.2 BENEFITS OF BREAST CANCER SCREENING

The effects of mammography screening

Randomized controlled trials have clearly demonstrated the effects of breast cancer screening on breast cancer mortality under relatively controlled circumstances, but the benefits of mammography in a population setting may be different. First, the performance of screening likely improved since the trials, because quality assurance, training of radiographers and radiologists and mammography techniques may have improved.[16] Digital mammography, for instance, has the potential to further increase the accuracy of mammography because of its improved contrast resolution.[17–20] As a consequence, the number of prevented breast cancer deaths may increase. A trial population may also be healthier than the general screened population, which could raise the potential benefits of screening. On the other hand, the treatment of breast cancer has developed since the trials, in particular because adjuvant systemic therapy became more commonly used. This may interfere with the effects of early detection.[21]

The effects of population-based breast cancer screening on breast cancer mortality were therefore assessed in several studies. In 2003, Otto *et al.* published a study that calculated the reduction in breast cancer mortality based on the moment that screening was implemented in municipalities in the Netherlands.[22] Because the implementation of screening occurred gradually – in each municipality at a different time – breast cancer mortality could be assessed before and after the implementation of screening, without being affected by other time trends. The breast cancer mortality reduction that was observed could directly be related to the start of screening; the reduction since screening implementation in women aged 55–74 years was 1.7% per year. These findings were confirmed by a recent analysis of age-specific breast cancer incidence- and mortality trends in the Netherlands, which showed that despite increases in incidence, breast cancer mortality decreased in age groups that were targeted for screening, whereas no improvements were observed in women that were not yet invited. The moment at which the mortality trend changed was significantly correlated to the moment that screening started.[23] Two recent case-control studies in the Netherlands confirmed the beneficial effects of screening on breast cancer mortality. In the region of Nijmegen, a mortality reduction of 70% was observed in screened compared to non-screened women;[24] in the region of Rotterdam, the mortality reduction was 56%.[25] Despite large differences in the organization of screening, invitation policy, the performance of screening, and the reading- and referral policy of mammograms, the effectiveness of population-based screening was also demonstrated in several other countries.[26]

Screening effectiveness debated

Despite the cumulating evidence of the beneficial effects of breast cancer screening, critical objections to breast cancer screening persist. Some argue that the effects of screening may be smaller than generally assumed, while on the other hand, the risks may be larger. In 2001, the Cochrane Collaboration published a meta-analysis of the results of the randomized controlled trials, stating that mammography screening does not improve survival, and that the effects on breast cancer mortality are inconclusive.[27] According to them, the trials that showed significant results were methodologically inferior to those who did not show effects, for instance because incorrect methods of randomization were used. The only two trials which they judged to be of reasonable quality showed a non-significant 12% reduction in breast cancer mortality. Although recognizing the methodological differences between trials, the comments were later refuted.[28–31] In 2006, the same Cochrane Collaboration nuanced their conclusions, stating that breast cancer screening likely reduces breast cancer mortality, with a relative risk reduction of 15%.[32]

The debate about the effects of screening was stirred again recently, after two publications that showed no or only small reductions in breast cancer mortality in women

that could have benefitted from screening. The first compared incidence-based breast cancer mortality in Norwegian women living in regions with screening with breast cancer mortality in women that lived in regions without screening. After an average follow-up of 2.2 years, the relative risk of breast cancer death in the screening group was 10% lower than in the control group, but only a third of this difference could be attributed to screening alone.[33] The second study compared breast cancer mortality trends in Denmark in areas that had mammography screening with mortality in areas that had not. During a 10-years follow-up, an annual decline of 1% was observed in the screening regions, in the age group that could have benefitted from screening. At the same time, the mortality in the unscreened areas had dropped even stronger: an annual decline of 2% was reported, suggesting no mortality benefit.[34] Because these findings contradict those of the randomized trials and case-control studies, it remains relevant to critically examine breast cancer mortality trends.

1.3 RISKS AND COST-EFFECTIVENESS OF BREAST CANCER SCREENING

The harms of mammography screening

Inevitably, breast cancer screening also involves certain harms. Because mammography screening is offered to a healthy population, its risks should be kept to a minimum, while the benefits should be sufficiently high. The balance between harms and benefits, however, has often been debated.[35–36]

One of the most controversial screening harms is overdiagnosis. Overdiagnosis is the detection of a preclinical cancer that would not have been diagnosed during a woman's lifetime if she had not been screened. This could happen when a lesion grows slowly or not at all (or regress), and the woman eventually dies of something else than breast cancer. More rarely, overdiagnosis may also occur when a preclinical cancer has an average growth rate, but the woman dies an unnatural death, such as a traffic accident for instance.

On an individual level, it is impossible to distinguish patients with 'real' or 'pseudo' cancers: both are treated in a similar fashion. Most women will therefore perceive the early detection of cancer as a good, rather than a bad thing. On a population level, however, overdiagnosis means that more women will be diagnosed with cancer than in a situation without screening, which increases the costs of diagnostics and treatment. More importantly, more women will perceive distress and anxiety because they live with the diagnosis 'cancer'. And even if the tumour itself was never deemed to be fatal, its treatment involves an increased risk of complications and side-effects.[9]

It is difficult to assess the extent to which overdiagnosis occurs. Some argue that the risk is minimal, with estimates varying between 1% of all diagnosed cancers in a screened population[37] and 3% of all expected cancers in an unscreened population.[38] Others estimated the overdiagnosis risk to be substantially higher, with estimates up to 52%–54% of all expected cancers in women in the screening age.[36, 39] This thesis aims to explain why overdiagnosis estimates vary to such extent. The potential consequences of digital mammography screening, which was shown to increase the detection of non-invasive breast cancer (DCIS),[40–41] may be of particular interest in this context. Because the natural history of such lesions is largely unknown, an increased detection of DCIS at digital mammography screening could raise the number of overdiagnosed breast cancers. On the other hand, digital mammography may also prevent extra breast cancer deaths.

Another potential harm of breast cancer screening is the risk of 'false reassurance'. False reassurance is a delay in the cancer diagnosis due to having participated in screening. It occurs after a negative screen result, when a patient or doctor, perceiving the risk of developing cancer to be small, is consequently less alert to present symptoms or the need for further evaluation. Prolonged patient delays have been associated with increased tumour sizes,[42–43] more positive lymph nodes[42] and with decreased long-term survival.[42,44] False reassurance may therefore influence tumour stage and prognosis. So far, the phenomenon of delayed symptom presentation after mammography screening has not been examined yet.

Other downsides of screening include the inconvenience of the mammogram itself, possible distress and anxiety caused by false positive mammograms, and the longer time of living with a diagnosis for women that will not benefit from screening.[29] These harms will not be discussed in this thesis.

Cost-effectiveness of screening

For population screening to be justified, the costs of the intervention should be reasonably low. In the Netherlands and other Western countries, centrally organized mammography for women at average risk for breast cancer was shown to be highly cost-effective.[45–47] Mammography screening, however, is not necessarily cost-effective in every situation. Cost-effectiveness depends on country-specific demographic characteristics, breast cancer incidence, tumour stage distribution and breast cancer mortality before the initiation of screening, and the characteristics of screening, such as attendance, targeted screening ages and screening interval. The organisation of a health care system and the costs of screening, diagnostics and treatment also determine whether mammography screening is cost-effective. Under certain circumstances, 'opportunistic' mammography screening in asymptomatic women, the predominant form of screening in several European coun-

tries, may be a cost-effective alternative to programme-based mammography screening. In this thesis, the cost-effectiveness of both screening modalities is compared.

1.4 FUTURE DEVELOPMENTS IN BREAST CANCER SCREENING

Developments in or around mammography screening

Current and future developments in the early detection and treatment of breast cancer may affect the benefit-risk ratio of mammography screening. The growing use of adjuvant systemic therapies, the expansion of the lower age limit for screening and the implementation of digital mammography may be of particular interest in this context.

Since the mid-seventies, adjuvant systemic therapy became increasingly used for breast cancer patients.[48] In a meta-analysis of several trials on the effects of adjuvant chemo- and endocrine therapy, the Early Breast Cancer Trialists' Collaborative Group showed that such therapies could significantly reduce the risk of breast cancer death.[49–50] An effective adjuvant treatment could mean that the need for mammography screening becomes smaller. That is, if adjuvant treatment reduces the breast cancer mortality, the absolute benefits that can be obtained by early detection may decrease.[21] In the evaluation of the effects of breast cancer screening, the use and benefits of adjuvant treatment should thus be taken into account.

In 1998, the upper age limit of the Dutch screening programme was extended with 3 additional screening rounds between age 70 and 74. This extension was predicted to prevent extra breast cancer deaths, against reasonably increased screening risks.[10, 51] Currently, the lower age limit for screening is also under debate. The UK Age trial, offering annual screening to women between age 40 and 48, showed a reduction in breast cancer mortality.[52] This effect was not statistically significant, but longer follow-up or future improvements in screening may further increase the number of prevented deaths. On the other hand, the use of adjuvant screening therapy in younger women could reduce the screening potential. This thesis therefore assesses the long-term benefits of screening under the age of 50 in the presence of adjuvant systemic therapy.

In the discussion whether or not to screen before age 50 the radiation risks of screening younger women should be taken into account. That is, the risk of radiation-induced cancer was found to increase with younger age,[53] which could jeopardize the balance between benefits and harms of mammography screening. For women screened between age 50 and 69, the risk of radiation-induced breast cancer was estimated to be relatively small compared to the screening benefits.[54–55] However, a new model has recently been

developed with which the observed risks of high doses of radiation can be extrapolated to the doses that are observed at mammography. At the same time, the average absorbed radiation dose in women screened in the Netherlands was calculated for the first time. The balance between radiation risks and screening benefits should be assessed in the light of these developments.

1.5 RESEARCH QUESTIONS AND APPROACH

Research questions and outline of this thesis

The main objective of this thesis is to assess the benefits and risks of population-based breast cancer screening. Part 1 (Chapters 2–5) covers the present situation of mammography screening in the Netherlands and the current debate on overdiagnosis. In Part 2 of the thesis (Chapter 6 and 7), future directions of mammography screening will be discussed.

The following research questions will be addressed:

1. What are the benefits and harms of population-based mammography screening in the Netherlands? (Chapter 2)
   a. What are the benefits and harms of screen-film mammography screening?
   b. How has the implementation of digital mammography affected the breast cancer mortality and risk of overdiagnosis compared to screen-film mammography?

2. Why do estimates of overdiagnosis that are reported in literature vary to a large extent? (Chapter 3)

3. Do women with breast cancer symptoms, who participated in breast cancer screening, unnecessarily delay a visit to a doctor because of 'false reassurance'? (Chapter 4)

4. How does the organizational form of breast cancer screening affect its benefits and costs? (Chapter 5)
   a. What is the effectiveness of opportunistic breast cancer screening compared to organized screening, in terms of breast cancer mortality and life years gained?
   b. What is the cost-effectiveness of opportunistic *versus* organized breast cancer screening?

5. What are the effects of adjuvant systemic breast cancer treatment and mammography screening on breast cancer mortality in women older and younger than 50? (Chapter 6)
    a. What are the predicted effects of adjuvant systemic therapy on breast cancer mortality in women younger and older than 50?
    b. In the presence of adjuvant systemic therapy, what are the predicted effects of mammography screening of women aged 50–74 years on breast cancer mortality?
    c. What are the predicted effects of adjuvant treatment and screening starting between age 40 and 50 on breast cancer mortality?

6. How do the benefits of mammography screening in women younger and older than 50 relate to the risk of radiation-induced breast cancer and breast cancer death? (Chapter 7)
    a. What are the radiation-induced risks of breast cancer screening in women aged 50–74 years, as compared to the benefits of screening?
    b. How would an extension of the lower age limit for screening to ages below 50 affect radiation risks, as compared to the benefits?

Part 3 (Chapter 8) concludes this thesis with summary answers to and further discussion of the research questions and directions for further research.

Approach

The consequences of screening on population health can be assessed using microsimulation models, such as the MIcrosimulation SCreening ANalysis model 'MISCAN'.[56–58] With MISCAN, the results of randomized trials can be extrapolated to different screening ages, intervals and (improved) tests. The model consists of a part that simulates the demography of the population under study, a part that simulates the natural history of breast cancer, and a part that models the influence of screening on this natural history. By simulating a situation without and with screening, the effects and risks of screening can be assessed. The model is described in the Annex.

REFERENCES

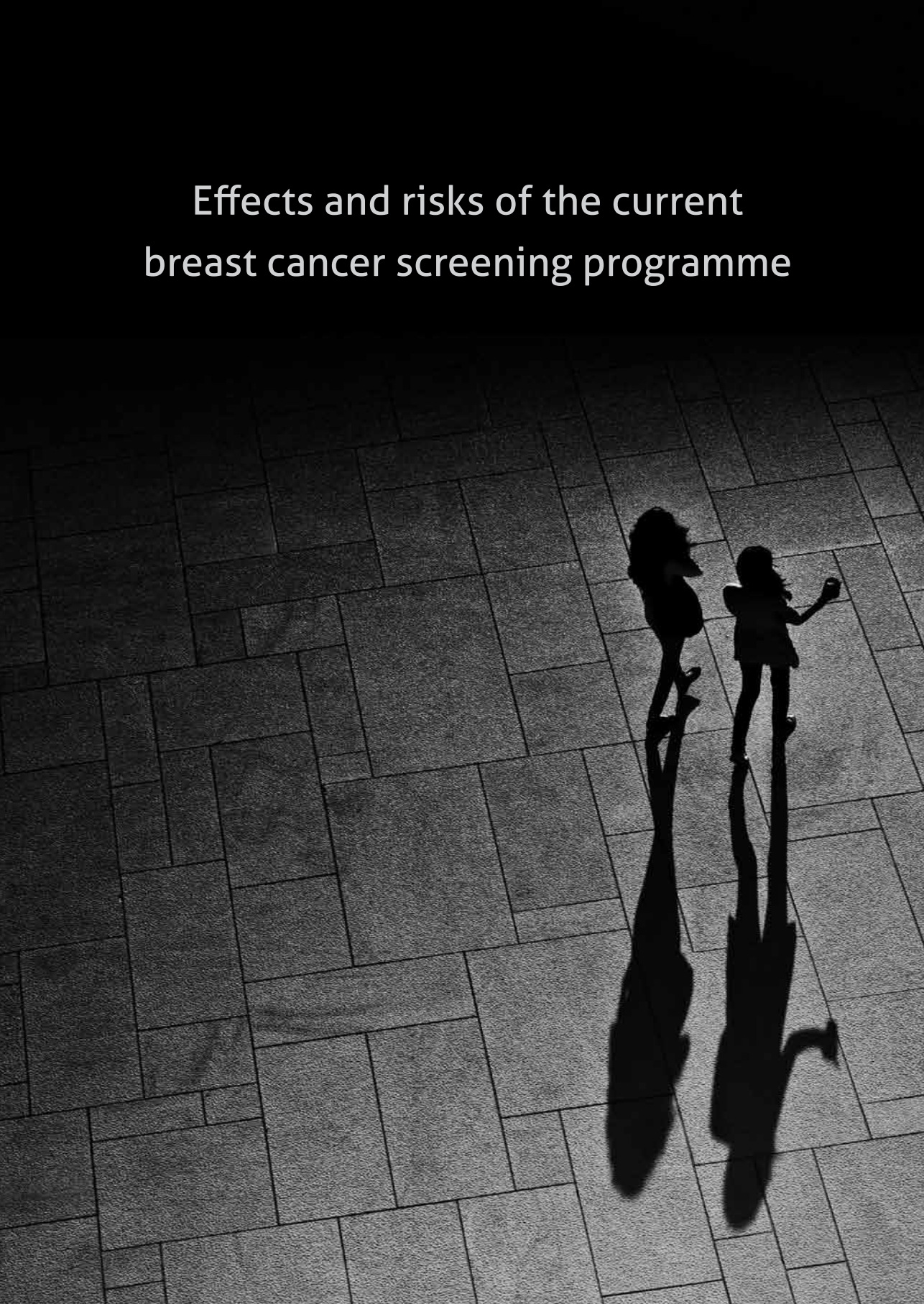
1. Ferlay J, Autier P, Boniol M, Heanue M, Colombet M, Boyle P. Estimates of the cancer incidence and mortality in Europe in 2006. Ann Oncol 2007;18:581–92.
2. Paap E, Broeders MJ, van Schoor G, Otten JD, Verbeek AL. Large increase in a Dutch woman's lifetime risk of developing breast cancer. Eur J Cancer 2008;44:1485–7.
3. IKCnet. http://www.ikcnet.nl/. http://www.ikcnet.nl/. April 6, 2010
4. Karim-Kos HE, de Vries E, Soerjomataram I, Lemmens V, Siesling S, Coebergh JW. Recent trends of cancer in Europe: a combined approach of incidence, survival and mortality for 17 cancer sites since the 1990s. Eur J Cancer 2008;44:1345–89.
5. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. Cancer 1989;63:181–7.
6. Michaelson JS, Silverstein M, Sgroi D, Cheongsiatmoy JA, Taghian A, Powell S, Hughes K, Comegno A, Tanabe KK, Smith B. The effect of tumor size and lymph node status on breast carcinoma lethality. Cancer 2003;98:2133–43.
7. Michaelson JS, Silverstein M, Wyatt J, Weber G, Moore R, Halpern E, Kopans DB, Hughes K. Predicting the survival of patients with breast carcinoma using tumor size. Cancer 2002;95:713–23.
8. Sant M, Allemani C, Capocaccia R, Hakulinen T, Aareleid T, Coebergh JW, Coleman MP, Grosclaude P, Martinez C, Bell J, Youngson J, Berrino F. Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. Int J Cancer 2003;106:416–22.
9. Wobbes T, Nortier JW, Koning CCE. Handboek Mammacarcinoom.ed. Utrecht: De Tijdstroom, 2007.
10. National Evaluation Team for Breast cancer screening (NETB), National evaluation of breast cancer screening in the Netherlands 1990–2007. Twelfth evaluation report. Erasmus MC University Medical Center Rotterdam, Radboud University Nijmegen Medical Centre, 2009.
11. Nystrom 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;359:909–19.
12. de Koning HJ, Boer R, Warmerdam PG, Beemsterboer PM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. J Natl Cancer Inst 1995;87:1217–23.
13. Collette HJ, Day NE, Rombach JJ, de Waard F. Evaluation of screening for breast cancer in a non-randomised study (the DOM project) by means of a case-control study. Lancet 1984;1:1224–6.
14. Verbeek AL, Hendriks JH, Holland R, Mravunac M, Sturmans F, Day NE. Reduction of breast cancer mortality through mass screening with modern mammography. First results of the Nijmegen project, 1975–1981. Lancet 1984;1:1222–4.
15. de Koning HJ, van Ineveld BM, van Oortmarssen GJ, de Haes JC, Collette HJ, Hendriks JH, van der Maas PJ. Breast cancer screening and cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors. Int J Cancer 1991;49:531–7.
16. van Schoor G, Moss SM, Otten JD, Donders R, Paap E, den Heeten GJ, Holland R, Broeders MJ, Verbeek AL. Increasingly strong reduction in breast cancer mortality due to screening. Br J Cancer 2011;104:910–4.

17. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, Conant EF, Fajardo LL, Bassett L, D'Orsi C, Jong R, Rebner M, *et al.* Diagnostic performance of digital *versus* film mammography for breast-cancer screening. N Engl J Med 2005;353:1773–83.
18. Lewin JM, D'Orsi CJ, Hendrick RE, Moss LJ, Isaacs PK, Karellas A, Cutter GR. Clinical comparison of full-field digital mammography and screen-film mammography for detection of breast cancer. AJR Am J Roentgenol 2002;179:671–7.
19. Lewin JM, Hendrick RE, D'Orsi CJ, Isaacs PK, Moss LJ, Karellas A, Sisney GA, Kuni CC, Cutter GR. Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4,945 paired examinations. Radiology 2001;218:873–80.
20. Skaane P, Young K, Skjennald A. Population-based mammography screening: comparison of screen-film and full-field digital mammography with soft-copy reading--Oslo I study. Radiology 2003;229:877–84.
21. Cronin KA, Feuer EJ, Clarke LD, Plevritis SK. Impact of adjuvant therapy and mammography on U.S. mortality from 1975 to 2000: comparison of mortality results from the cisnet breast cancer base case analysis. J Natl Cancer Inst Monogr 2006:112–21.
22. Otto SJ, Fracheboud J, Looman CW, Broeders MJ, Boer R, Hendriks JH, Verbeek AL, de Koning HJ. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. Lancet 2003;361:1411–7.
23. Otten JD, Broeders MJ, Fracheboud J, Otto SJ, de Koning HJ, Verbeek AL. Impressive time-related influence of the Dutch screening programme on breast cancer incidence and mortality, 1975–2006. Int J Cancer 2008;123:1929–34.
24. Paap E, Holland R, den Heeten GJ, van Schoor G, Botterweck AA, Verbeek AL, Broeders MJ. A remarkable reduction of breast cancer deaths in screened *versus* unscreened women: a case-referent study. Cancer Causes Control 2010;21:1569–73.
25. Otto S, Fracheboud J, Verbeek ALM, Boer R, Reijerink-Verheij JCIY, Otten JDM, Broeders MJM, de Koning HJ. Mammography Screening and Risk of Breast Cancer Death: A Population-Based Case–Control Study. Cancer Epidemiol Biomarkers Prev. 2011;Accepted.
26. Schopper D, de Wolf C. How effective are breast cancer screening programmes by mammography? Review of the current evidence. Eur J Cancer 2009;45:1916–23.
27. Olsen O, Gotzsche PC. Screening for breast cancer with mammography. Cochrane database of systematic reviews (Online) 2001:CD001877.
28. de Koning HJ. Mammographic screening: evidence from randomised controlled trials. Ann Oncol 2003;14:1185–9.
29. International Agency for Research on Cancer. Volume 7: Breast Cancer Screening. In: Vainio H, Bianchini F. IARC Handbooks of Cancer Preventioned. Lyon: IARCPress, 2002.
30. de Koning HJ. Assessment of nationwide cancer-screening programmes. Lancet 2000;355:80–1.
31. Gezondheidsraad. Het nut van bevolkingsonderzoek naar borstkanker. Publicatienr 2002/03: Den Haag: Gezondheidsraad, 2002.
32. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane database of systematic reviews (Online) 2006:CD001877.
33. Kalager M, Zelen M, Langmark F, Adami HO. Effect of screening mammography on breast-cancer mortality in Norway. N Engl J Med 2010;363:1203–10.
34. Jorgensen KJ, Zahl PH, Gotzsche PC. Breast cancer mortality in organised mammography screening in Denmark: comparative study. BMJ (Clinical research ed 2010;340:c1241.

35. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane database of systematic reviews (Online) 2011:CD001877.
36. Jorgensen KJ, Gotzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. BMJ (Clinical research ed 2009;339:b2587.
37. Duffy SW, Agbaje O, Tabar L, Vitak B, Bjurstam N, Bjorneld L, Myles JP, Warwick J. Overdiagnosis and overtreatment of breast cancer: estimates of overdiagnosis from two trials of mammographic screening for breast cancer. Breast Cancer Res 2005;7:258–65.
38. de Koning HJ, Draisma G, Fracheboud J, de Bruijn A. Overdiagnosis and overtreatment of breast cancer: microsimulation modelling estimates based on observed screen and clinical data. Breast Cancer Res 2006;8:202.
39. Morrell S, Barratt A, Irwig L, Howard K, Biesheuvel C, Armstrong B. Estimates of overdiagnosis of invasive breast cancer associated with screening mammography. Cancer Causes Control 2010;21:275–82.
40. Del Turco MR, Mantellini P, Ciatto S, Bonardi R, Martinelli F, Lazzari B, Houssami N. Full-field digital *versus* screen-film mammography: comparative accuracy in concurrent screening cohorts. AJR Am J Roentgenol 2007;189:860–6.
41. Vigeland E, Klaasen H, Klingen TA, Hofvind S, Skaane P. Full-field digital mammography compared to screen film mammography in the prevalent round of a population-based screening programme: the Vestfold County Study. Eur Radiol 2008;18:183–91.
42. Afzelius P, Zedeler K, Sommer H, Mouridsen HT, Blichert-Toft M. Patient's and doctor's delay in primary breast cancer. Prognostic implications. Acta oncologica (Stockholm, Sweden) 1994;33:345–51.
43. Burgess CC, Ramirez AJ, Richards MA, Love SB. Who and what influences delayed presentation in breast cancer? Br J Cancer 1998;77:1343–8.
44. Richards MA, Smith P, Ramirez AJ, Fentiman IS, Rubens RD. The influence on survival of delay in the presentation and treatment of symptomatic breast cancer. Br J Cancer 1999;79:858–64.
45. Groenewoud JH, Otten JD, Fracheboud J, Draisma G, van Ineveld BM, Holland R, Verbeek AL, de Koning HJ. Cost-effectiveness of different reading and referral strategies in mammography screening in the Netherlands. Breast cancer research and treatment 2007;102:211–8.
46. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. J Natl Cancer Inst 2006;98:774–82.
47. Carles M, Vilaprinyo E, Cots F, Gregori A, Pla R, Roman R, Sala M, Macia F, Castells X, Rue M. Cost-effectiveness of early detection of breast cancer in Catalonia (Spain). BMC cancer 2011;11:192.
48. Vervoort MM, Draisma G, Fracheboud J, van de Poll-Franse LV, de Koning HJ. Trends in the usage of adjuvant systemic therapy for breast cancer in the Netherlands and its effect on mortality. Br J Cancer 2004;91:242–7.
49. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687–717.
50. Davies C, Godwin J, Gray R, Clarke M, Cutter D, Darby S, McGale P, Pan HC, Taylor C, Wang YC, Dowsett M, Ingle J, *et al.* Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet 2011;378:771–84.

51. Fracheboud J, Groenewoud JH, Boer R, Draisma G, de Bruijn AE, Verbeek AL, de Koning HJ. Seventy-five years is an appropriate upper age limit for population-based mammography screening. Int J Cancer 2006;118:2020–5.
52. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet 2006;368:2053–60.
53. Preston DL, Mattsson A, Holmberg E, Shore R, Hildreth NG, Boice JD, Jr. Radiation effects on breast cancer risk: a pooled analysis of eight cohorts. Radiation research 2002;158:220–35.
54. Berrington de Gonzalez A, Reeves G. Mammographic screening before age 50 years in the UK: comparison of the radiation risks with the mortality benefits. Br J Cancer 2005;93:590–6.
55. Mattsson A, Leitz W, Rutqvist LE. Radiation risk and mammographic screening of women from 40 to 49 years of age: effect on breast cancer rates and years of life. Br J Cancer 2000;82:220–6.
56. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, Mandelblatt JS, Yakovlev AY, Habbema JD, Feuer EJ, Cancer Intervention and Surveillance Modeling Network C. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med 2005;353:1784–92.
57. Groenewoud JH, Pijnappel RM, van den Akker-Van Marle ME, Birnie E, Buijs-van der Woude T, Mali WP, 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;90:383–92.
58. van Oortmarssen GJ, Habbema JD, van der Maas PJ, de Koning HJ, Collette HJ, Verbeek AL, Geerts AT, Lubbe KT. A model for breast cancer screening. Cancer 1990;66:1601–12.

Effects and risks of the current breast cancer screening programme



Digital mammography screening: Weighing reduced mortality against increased overdiagnosis

Rianne de Gelder, Jacques Fracheboud, Eveline A.M. Heijnsdijk, Gerard den Heeten, André L.M.Verbeek, Mireille J.M. Broeders, Gerrit Draisma, Harry J. de Koning

Prev Med. 2011 Sep 1;53(3):134–40. Epub 2011 Jun 21

ABSTRACT

Objective

Digital mammography has been shown to increase the detection of ductal carcinoma *in situ* (DCIS) compared to screen-film mammography. The benefits and risks of such an increase were assessed.

Methods

Breast cancer detection rates were compared between 502,574 screen-film and 83,976 digital mammograms performed between 2004 and 2006 among Dutch screening participants. The detection rates were then modeled using a baseline model and two extreme models that respectively assumed a high rate of progression and no progression of preclinical DCIS to invasive cancer. With these models, breast cancer mortality and overdiagnosis were predicted.

Results

The DCIS detection rate was significantly higher at digital mammography (1.2 per 1000 mammograms (95% C.I. 1.0–1.5)) than at screen-film mammography (0.7 per 1000 mammograms (95% C.I. 0.6–0.7)). Consequently, 287 (range progressive- non progressive model: 1–598) extra breast cancer deaths per 1,000,000 women (a 4.4% increase) were predicted to be prevented. An extra 401 (range: 165–2271) cancers would be overdiagnosed (a 21% increase).

Conclusion

Modeling predicted that digital mammography screening would further reduce breast cancer mortality by 4.4%, at a 21% increased overdiagnosis rate. The consequences of digital screening, however, are sensitive to underlying assumptions on the natural history of DCIS.

2.1 INTRODUCTION

Because of its improved contrast resolution, digital mammography has the potential to improve test accuracy compared to screen-film mammography.[1–4] Digital mammography in women aged 45–69 resulted in significantly higher referral and cancer detection rates than screen-film mammography.[5] Other trials, however, showed that an improved accuracy was limited to women under the age of 50, women with dense breasts and pre- or peri-menopausal women.[2–3] The long term benefits and risks of digital mammography in population-based screening have not yet been assessed.

In 2004, a feasibility study of screening women with digital mammography was started in the Netherlands. Here, the results were modeled and used to predict the benefits and risks of implementing digital mammography as compared to screen-film mammography. Because various digital mammography studies observed an increased detection of ductal carcinoma *in situ*[1] (DCIS) and micro-calcifications frequently related to DCIS,[6–7] our study focused on the benefits and risks of an increased detection of DCIS. However, the extent to which such lesions have the potential to become invasive cancers remains uncertain.[8] Detecting DCIS may prevent progression to invasive cancer, but may also imply that a lesion is diagnosed that would not have progressed to invasive cancer during the woman's lifetime (i.e. 'overdiagnosis'). The main purpose of this study is to assess the consequences of an increased detection of DCIS by digital mammography relative to screen-film mammography screening. Three scenarios for the natural history of DCIS were considered, that assumed that 1) a part would progress and another part would not, 2) all cases would progress to invasive cancer, or 3) all would be overdiagnosed.

2.2 METHODS

Since 1990, all women in the Netherlands aged 49–69 (since 1998: aged 49–74) are offered biennial screening. In 2004, digital mammography screening was implemented in three screening units in Utrecht, Drechtsteden and Heerenveen. During the same period, screen-film mammography was performed in 20 other screening units in these regions. In 10 of these screening units, the mammograms were interpreted by the same radiologists who read the digital mammograms. We only included those in our analysis. Referral and TNM stage-specific[9] detection rates at digital mammography screening between 2004 and 2006 were compared to the referral and detection rates at screen-film mammography during the same period. To test statistical differences between digital mammography and screen-film mammography, 95% confidence intervals around the

mean referral and detection rates were calculated. The rates were age-standardized to the Dutch female population in 2000.

Screen-film and digital mammography differed in the number of views made at subsequent screening examinations. Standard practice at subsequent screen-film mammograms was a medio-lateral oblique (MLO) view, with an additional cranio-caudal (CC) view on indication (in approximately 30%–40% of the examinations). Utrecht and Heerenveen used the same protocol for subsequent digital mammograms, but subsequent digital examinations in Drechtsteden all included a MLO and CC view. The digital screening units in Utrecht and Heerenveen used Full Field Digital Mammography, while the digital unit in Drechtsteden used digital phosphor storage plate mammography. In Heerenveen and Utrecht, prior screen-film mammograms were digitized and directly available at reading for comparison with the new digital images. In Drechtsteden, only screen-film mammograms were available at reading.

The benefits and risks of screen-film and digital mammography screening were predicted using the MIcro-simulation Screening Analysis model MISCAN.[10] With MISCAN, individual life histories of women are generated, and the impact of screening on these life histories is assessed. The simulated life histories together represent the observed female population aged 0–100 years in 1989. Some women in the simulated population develop breast cancer. Its natural course is modeled as a Markov-like progression through the successive preclinical invasive tumor stages T1a, T1b, T1c and T2+ (≤5 mm, 6–10 mm, 11–20 mm, >20 mm, respectively). T1a may or may not be preceded by preclinical screen-detectable DCIS (Figure 2.1a). A fraction of preclinical DCIS may also regress spontaneously. Each preclinical tumor may progress into the next stage, become clinically diagnosed, or become screen-detected. Transition probabilities, durations of tumor stages and test sensitivity were estimated using data from the Dutch cancer registry and the nationwide screening program.[11–12] These data included the age-, stage-, and calendar year specific incidence of clinically diagnosed and screen-detected breast cancer (at screen-film mammography), breast cancer detection rates, and interval cancer rates between 1990 and 2006.

In the 'baseline model', the fraction of breast tumors that has a screen-detectable preclinical DCIS stage was estimated to be 18%. Of these lesions, 11% progress to invasive cancer, 5% is clinically diagnosed and 2% regress. The mean duration of preclinical screen-detectable DCIS was estimated to be 5.2 years; the mean duration of preclinical invasive breast cancer 2.6 years. The estimated test sensitivity of screen-film mammography, defined as the fraction of screen-detectable tumors that become screen-detected,

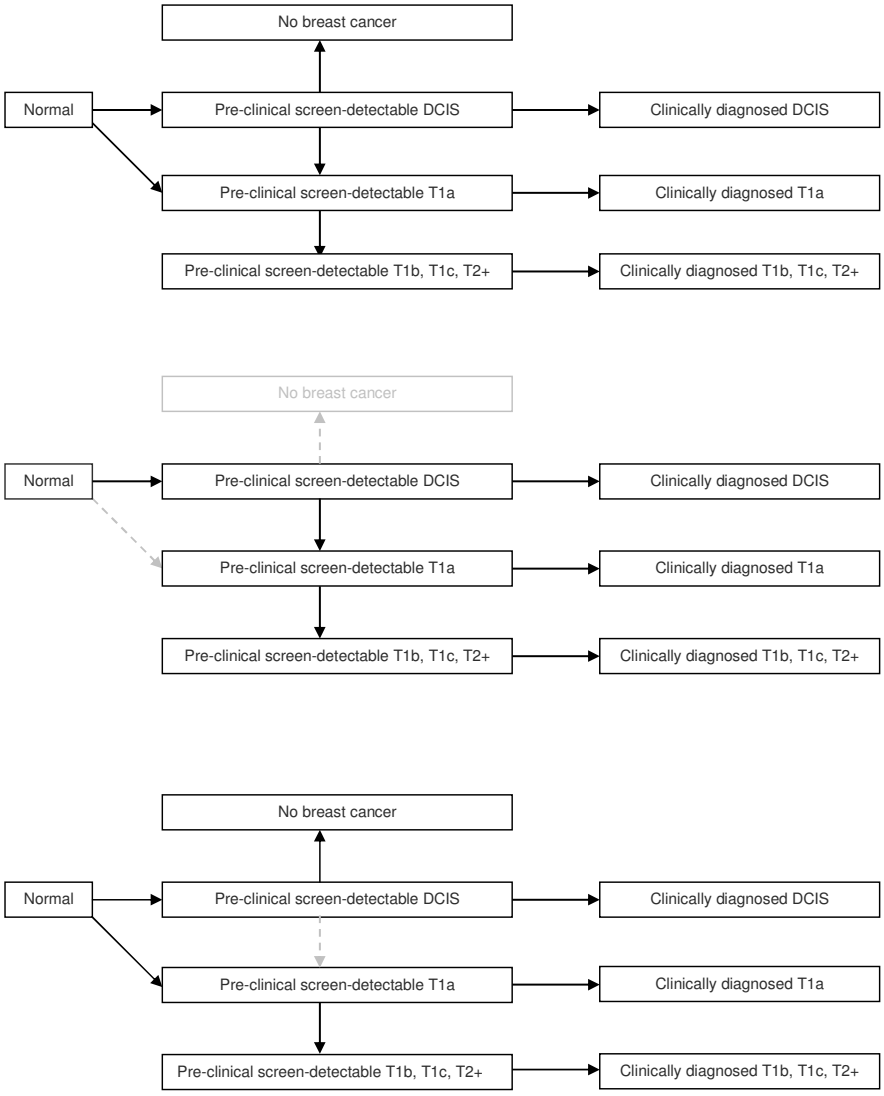


Figure 2.1a–c Preclinical tumor stage transitions in the absence of mammography screening in the baseline (a), the progressive (b) and the non-progressive (c) screen-film and digital mammography models. In the state 'no breast cancer', a woman does not develop a breast malignancy during her lifetime anymore. In the progressive model variant (b), no transition between normal and preclinical screen-detectable T1a, and between preclinical screen-detectable DCIS and 'no breast cancer' occurs. In the non-progressive model variant (c), no transition between preclinical screen-detectable DCIS and preclinical screen-detectable T1a occurs.

was 47%, 47%, 62%, 90% and 8% for DCIS, T1a, T1b, T1c and T2+ respectively (Appendix 2.1).

Because breast cancer is a heterogeneous disease for which the development from preclinical lesion to invasive cancer is unclear,[13] and because model predictions depend on the assumed natural history of the disease, two extreme alternatives to the baseline model were also explored:

1. The 'progressive model' in which all breast tumors pass through a preclinical screen-detectable DCIS stage, none of which regress. An estimated 96% of all screen-detectable DCIS would progress to invasive cancer if no screening would take place, and 4% would be clinically diagnosed. (Figure 2.1b)
2. The 'non-progressive model', in which none of the invasive breast tumors pass through a preclinical screen-detectable DCIS stage, and in which the majority of preclinical DCIS would regress. An estimated 2% of preclinical DCIS would be clinically diagnosed (Figure 2.1c).

The parameters of the two extreme models described above were estimated in the same way as the baseline model parameters, using incidence and detection rates from the period in which screen-film mammography was only used. To assess the consequences of digital screening, variants to these models were developed with a higher test sensitivity for DCIS than their screen-film counterparts, using detection rates of DCIS and invasive breast cancer at digital mammography screening between 2004 and 2006. In the baseline model, the sensitivity of digital mammography for DCIS was estimated to be 100%. In the progressive and non-progressive model, the sensitivity was estimated to be 94% and 72%, respectively (Appendix 2.1).

Using these models, we predicted the number of prevented breast cancer deaths and overdiagnosed breast cancers after a 30 year period of biennial screening, starting in 1990. An 82% participation rate was assumed, corresponding to the current participation rate in the Netherlands.[12] The number of prevented breast cancer deaths was calculated by comparing the predicted breast cancer mortality in the presence and absence of screening. Overdiagnosis was calculated similarly, by comparing the predicted breast cancer incidence in the presence and absence of screening. All effects were calculated for 1,000,000 women aged 0–100 in 1989 with at least one screening examination between 1990 and 2020, measured during the lifespan of this population.

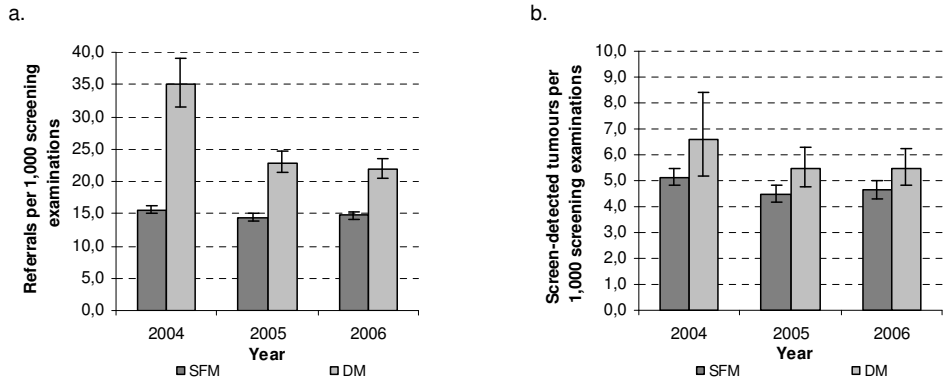


Figure 2.2a–b Observed referral (a) and breast cancer detection rates (b), by year and screening modality

2.3 RESULTS

Screen-film and digital mammography screening

Between 2004 and 2006, 83,976 digital and 502,574 screen-film mammograms were made. Referral rates were significantly higher at digital than at screen-film mammography screening: 14.9 per 1000 screen-film (95% C.I.: 14.6–15.3) *versus* 23.8 per 1000 digital mammograms (95% C.I.: 22.8–24.9) (Table 2.1). Referral rates peaked at the start of digital screening in 2004 to a rate twice that of screen-film mammography, but decreased from 2005 on (Figure 2.2a). The breast cancer detection rate at digital mammography remained stable over the years (Figure 2.2b). The average detection rate between 2004 and 2006 was significantly higher at digital than at screen-film mammography: 4.8 per 1000 screen-film (95% C.I.: 4.6–5.0) compared to 5.6 per 1000 digital examinations (95% C.I.: 5.1–6.1) (Table 2.1). The relative increase in detection rates was similar between first and subsequent screening examinations (Table 2.1).

The increase in breast cancer detection was mainly attributable to an 80% increase in the detection of DCIS, from 0.7 per 1000 screen-film (95% C.I.: 0.6–0.7) to 1.2 per 1000 digital mammograms (95% C.I.: 1.0–1.5). The detection rate of invasive cancers did not significantly differ between screen-film and digital mammography (3.9 per 1000 screen-film (95% C.I.: 3.8–4.1) *versus* 4.4 per 1000 digital mammograms (95% C.I.: 3.9–4.8). DCIS constituted 14% of all tumors detected at screen-film mammography, and 22% of all tumors detected at digital mammography screening. As a consequence, the stage distribution of digitally detected tumors was slightly more favorable than that of tumors detected at screen-film mammography, with 84% against 78% of all detected breast cancers being non-invasive or 20 mm or smaller. No differences were observed between

Table 2.1 Referral advices, screen-detected in situ and invasive carcinomas between 2004 and 2006, by type of mammography

2004–2006	First screening examinations			Subsequent screening examinations (<2,5 years interval)			All screening examinations		
	Age 49–54			Age 50–74			Age 49–74		
	Screen-film	Digital	+/- (%)	Screen-film	Digital	+/- (%)	Screen-film	Digital	+/- (%)
Screening examinations (n)	**61,038**	**12,078**		**434,743**	**70,635**		**502,574**	**83,976**	
Referral advice (n)	**1815**	**553**		**5411**	**1421**		**7445**	**2023**	
Age-standardized referral rate per 1000 examinations (95% C.I.)	30.2 (28.8–31.6)	49.4 (45.5–53.7)	+64[a]	12.6 (12.2–12.9)	20.3 (19.3–21.4)	+63[a]	14.9 (14.6–15.3)	23.8 (22.8–24.9)	+61[a]
Screening carcinomas[b] (n)	**301**	**74**		**1998**	**381**		**2365**	**463**	
Age-standardized detection rate per 1000 examinations (95% C.I.)	5.1 (4.6–5.7)	6.8 (5.5–8.5)	+33	4.6 (4.4–4.8)	5.4 (4.9–6.0)	+18[c]	4.8 (4.6–5.0)	5.6 (5.1–6.1)	+19[c]
Screen-detected DCIS[b] (n)	**47**	**19**		**278**	**83**		**335**	**102**	
Age-standardized detection rate per 1000 examinations (95% C.I.)	0.8 (0.6–1.1)	1.7 (1.1–2.6)	+103	0.6 (0.6–0.7)	1.2 (0.9–1.4)	+81[d]	0.7 (0.6–0.7)	1.2 (1.0–1.5)	+80[d]
Screen-detected invasive[b] (n)	**238**	**54**		**1654**	**296**		**1946**	**358**	
Age-standardized detection rate per 1000 examinations (95% C.I.)	4.1 (3.6–4.6)	5.0 (3.9–6.5)	+23	3.8 (3.6–4.0)	4.2 (3.8–4.7)	+11	3.9 (3.8–4.1)	4.4 (3.9–4.8)	+12

[a] p-Value < 0.0001

[b] The difference between the number of screening carcinomas and the number of DCIS and invasive cancers is the number of unclassified cancers and cancers with unknown TNM stage.

[c] p-Value < 0.02

[d] p-Value < 0.001

screen-film mammography and digital mammography in interval cancers and the sensitivity and specificity of mammography (data not shown).

Model validation and parameter estimates

Each model predicted the incidence of clinically diagnosed DCIS and invasive breast cancers, the stage- and age- specific tumor detection rates, interval cancer rates and breast cancer mortality reasonably well. Figure 2.3 shows the observed and modeled detection rates of DCIS and invasive cancer.

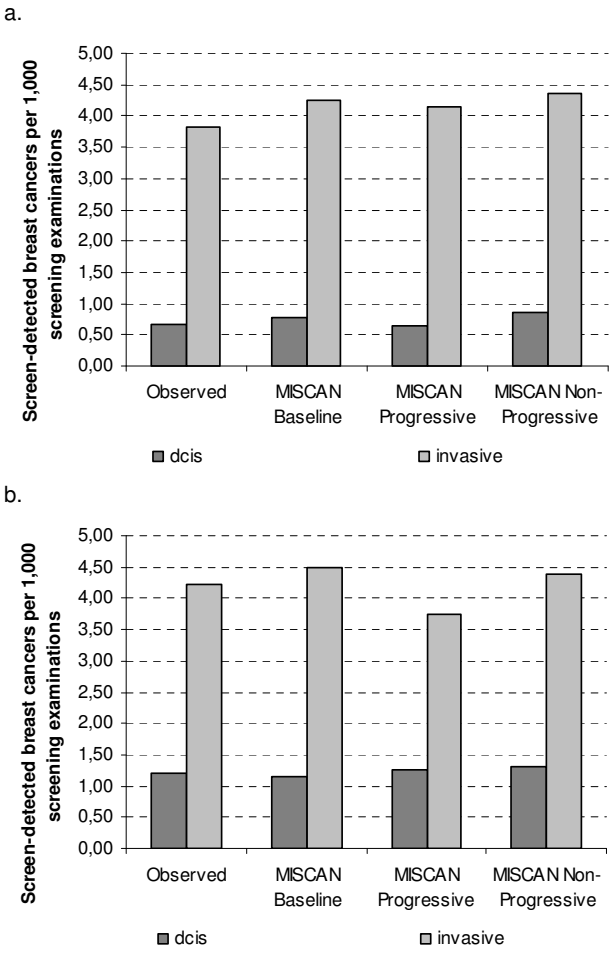


Figure 2.3a–b Observed and predicted detection rates of DCIS and invasive breast cancer in the period 2004–2006, at screen-film mammography screening (a) and digital mammography screening (b)

Table 2.2 Predicted breast cancer incidence, overdiagnosis and breast cancer deaths at screen-film and digital mammography screening

Baseline model	Screen-film mammography	Digital mammography
Population aged 0–100 in 1990, with at least 1 screening examination between 1990 and 2020 (n)	1,000,000	1,000,000
Screening examinations, 1990–2020 (n)	5,632,810	5,628,930
Breast cancers, measured during the whole lifespan of the population (n)	92,413	92,801
Screen-detected breast cancers (n)	26,720	28,258
Overdiagnosed breast cancers (n)	1926	2327
Fraction of overdiagnosed breast cancers of all diagnosed breast cancers (%)	*2.1%*	*2.5%*
Fraction of overdiagnosed breast cancers of all screen-detected breast cancers (%)	*7.2%*	*8.2%*
Predicted breast cancer deaths, without screening, during the whole lifespan of the population (n)	28,971	28,971
Predicted breast cancer deaths, with screening, during the whole lifespan of the population (n)	22,394	22,106
Reduction in breast cancer deaths (n, %)	*6577 (22.7%)*	*6864 (23.7%)*
Progressive model		
Population aged 0–100 in 1990, with at least 1 screening examination between 1990 and 2020 (n)	1,000,000	1,000,000
Screening examinations, 1990–2020 (n)	5,637,820	5,635,370
Breast cancers, measured during the whole lifespan of the population (n)	94,517	94,676
Screen-detected breast cancers (n)	25,279	26,674
Overdiagnosed breast cancers (n)	1168	1333
Fraction of overdiagnosed breast cancers of all diagnosed breast cancers (%)	*1.2%*	*1.4%*
Fraction of overdiagnosed breast cancers of all screen-detected breast cancers (%)	*4.6%*	*5.0%*
Predicted breast cancer deaths, without screening, during the whole lifespan of the population (n)	29,668	29,668
Predicted breast cancer deaths, with screening, during the whole lifespan of the population (n)	22,914	22,316
Reduction in breast cancer deaths (n, %)	*6753 (22.8%)*	*7351 (24.8%)*
Non-progressive model		
Population aged 0–100 in 1990, with at least 1 screening examination between 1990 and 2020 (n)	1,000,000	1,000,000
Screening examinations, 1990–2020 (n)	5,628,200	5,620,920
Breast cancers, measured during the whole lifespan of the population (n)	96,504	98,778
Screen-detected breast cancers (n)	27,838	30,209
Overdiagnosed breast cancers (n)	5336	7607
Fraction of overdiagnosed breast cancers of all diagnosed breast cancers (%)	*5.5%*	*7.7%*
Fraction of overdiagnosed breast cancers of all screen-detected breast cancers (%)	*19.2%*	*25.2%*
Predicted breast cancer deaths, without screening, during the whole lifespan of the population (n)	28,758	28,758
Predicted breast cancer deaths, with screening, during the whole lifespan of the population (n)	22,777	22,776
Reduction in breast cancer deaths (n, %)	*5982 (20.8%)*	*5983 (20.8%)*

Breast cancer mortality

In a population of 1,000,000 women aged 0–100 in 1989 with at least 1 screening examination, 5,632,810 screen-film mammograms were predicted to be performed between 1990 and 2020 (Table 2.2, using the baseline model). Screening would prevent 6577 breast cancer deaths in this population. Digital mammography was predicted to prevent 287 more breast cancer deaths than screen-film mammography: a relative increase of 4.4%.

In the progressive and non-progressive models, a predicted 6753 and 5982 breast cancer deaths per 1,000,000 women were prevented by screen-film mammography screening (Table 2.2). Digital mammography screening would prevent 598 more breast cancer deaths in the progressive model. In the non-progressive model, however, digital screening had no additional benefits over screen-film mammography screening.

Overdiagnosis

The baseline model predicted that 1926 tumors per 1,000,000 screened women would be overdiagnosed at screen-film mammography (i.e. 2.1% of all diagnosed breast cancers in screened women, or 7.2% of all screen-detected cancers; Table 2.2). Digital mammography screening would increase the number of overdiagnosed tumors by 21%, to 2527 per 1,000,000 screened women (2.5% of all diagnosed breast cancers in screened women or 8.2% of all screen-detected tumors).

Using the progressive model, digital screening would increase the number of overdiagnosed tumors by 14% (from 1168 cases at screen-film to 1333 cases at digital mammography). In the non-progressive model, on the contrary, digital screening would raise the number of overdiagnosed tumors by 43% (from 5336 cases at screen-film to 7607 cases at digital mammography).

2.4 DISCUSSION

Digital mammography led to a statistically significant 80% increase in the number of screen-detected DCIS cases compared to screen-film mammography, as observed in other studies.[6–7] Using our baseline model we found an increased detection of DCIS could raise the number of prevented breast cancer deaths by 4.4%, and the number of overdiagnosed tumors by 21%. Such predictions, however, depend on the assumed progression and regression rate and the mean duration of preclinical screen-detectable DCIS.

DCIS is a heterogeneous disease and its natural history is poorly understood. The prevalence of DCIS observed in autopsy studies (0.2%–18%) suggests that not all cases become invasive.[14] Most argue that all DCIS have the potential to progress, but that some cases are destined to grow faster and are associated with recurrence after local excision,[15–16] invasion of the basement membrane,[17] distant metastases after recurrence,[18] and poor survival.[18–19] Studies in the Netherlands found that between 47% and 54% of screen-detected DCIS are the more progressive, poorly differentiated 'high grade' type;[20–22] somewhat less than the 61% (11%/18%) progression rate used in our baseline model. The most direct evidence for DCIS progression comes from the follow-up of under-treated DCIS initially misdiagnosed as benign, in which 11%–60% women develop invasive cancer within 10–20 years.[14] Further indication for DCIS progression consists of microscopic studies on DCIS, that showed basement membrane invasion in 15%–28% of the cases, and microscopy on invasive lesions that showed that DCIS was present in 20%–30% of the carcinomas.[17] Because genetic and histological similarities were found between DCIS and recurrent invasive cancer,[15, 18] it is thought that they share a common etiology. Because our two alternative models with extreme assumptions on progression and regression fit equally well to the observed breast cancer incidence and detection rates as the baseline model, our study could not provide information about the 'true' natural history of breast cancer.

Our predicted overdiagnosis rate of 2.1% at screen-film mammography and 2.5% at digital mammography (baseline model) was in line with estimates of between 1% and 3% from previous modeling studies,[23–24] but much lower than recent estimates of around 50%.[25–26] Differences between overdiagnosis estimates may be explained by the length follow-up to allow for lead time, or by the denominator that is used to define the population at risk.[27–28]

We may have underestimated the consequences of digital screening, because the 12% increase in the detection of invasive cancers, mostly attributable to T1a and T1b tumors, was not accounted for. However, this increase was non-significant. Recent data (including detection rates in 2007) showed that the detection of invasive cancers, particularly T1a, increased significantly at digital screening compared to screen-film mammography.[12] If we would include this increase, the predicted number of breast cancer deaths would be 13% lower than at screen-film mammography, but the number of overdiagnosed cases 26% higher. Our findings may also have been affected by the additional CC views at subsequent digital mammograms in Drechtsteden. The breast cancer referral and detection rates at this unit were higher than in Utrecht or Heerenveen. Because additional CC views at subsequent examinations are increasingly indicated in the Netherlands, the breast cancer mortality reduction and overdiagnosis rate are unlikely to be strongly

over-estimated. In all three regions, referral rates peaked in 2004, while detection rates remained stable. This suggests that our estimates are not affected by a 'learning curve' effect. Our model was based on Dutch incidence and mortality data from 1990 to 2006, a period in which breast cancer patients may be treated by adjuvant systemic therapy (see Jatoi, this issue). This may have affected the screening effects to some extent.

Our analysis focused on the consequences of digital mammography among the targeted age group of 49–74 years old women. The Digital Mammographic Imaging Screening Trial (DMIST) suggested that digital mammography was more accurate than screen-film mammography for pre- and peri-menopausal women younger than 50 years with dense breasts.[3, 29] Younger women may therefore benefit more from digital mammography than the age group that was studied here. Presently, the effectiveness of screen-film mammography screening for women below age 50 is not sufficiently supported by scientific evidence (see Moss, this issue), but digital screening might change this.

Conclusion

The increased detection of DCIS by digital mammography screening could reduce breast cancer mortality by 4.4%, at a 21% increased overdiagnosis rate, but this is sensitive to assumptions on progression of DCIS.

Acknowledgements

The authors thank Ank de Roo and the screening organization 'Stichting Bevolkingsonderzoek Borstkanker Zuid-West-Nederland' in Rotterdam, Thijs Beekman and the screening organization 'Stichting Bevolkingsonderzoek Midden-West' in Utrecht, Clyde Elgin and the screening organization 'Stichting Bevolkingsonderzoek Borstkanker Noord-Nederland' in Groningen, and Linda de Munck and the Comprehensive Cancer Centre North East for providing data on digital mammography screening.

REFERENCES

1. Lewin JM, Hendrick RE, D'Orsi CJ, Isaacs PK, Moss LJ, Karellas A, Sisney GA, Kuni CC, Cutter GR. Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4,945 paired examinations. Radiology 2001;218:873–80.
2. Lewin JM, D'Orsi CJ, Hendrick RE, Moss LJ, Isaacs PK, Karellas A, Cutter GR. Clinical comparison of full-field digital mammography and screen-film mammography for detection of breast cancer. AJR Am J Roentgenol 2002;179:671–7.
3. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, Conant EF, Fajardo LL, Bassett L, D'Orsi C, Jong R, Rebner M, *et al.* Diagnostic performance of digital *versus* film mammography for breast-cancer screening. N Engl J Med 2005;353:1773–83.
4. Skaane P, Young K, Skjennald A. Population-based mammography screening: comparison of screen-film and full-field digital mammography with soft-copy reading--Oslo I study. Radiology 2003;229:877–84.
5. Skaane P, Hofvind S, Skjennald A. Randomized trial of screen-film *versus* full-field digital mammography with soft-copy reading in population-based screening program: follow-up and final results of Oslo II study. Radiology 2007;244:708–17.
6. Del Turco MR, Mantellini P, Ciatto S, Bonardi R, Martinelli F, Lazzari B, Houssami N. Full-field digital *versus* screen-film mammography: comparative accuracy in concurrent screening cohorts. AJR Am J Roentgenol 2007;189:860–6.
7. Vigeland E, Klaasen H, Klingen TA, Hofvind S, Skaane P. Full-field digital mammography compared to screen film mammography in the prevalent round of a population-based screening programme: the Vestfold County Study. Eur Radiol 2008;18:183–91.
8. National Institutes of Heath State-of-the-Science Conference Statement. Diagnosis and Management of Ductal Carcinoma *in situ* (DCIS). U.S. Department of Health & Human Services - National Institutes of Health: Bethesda, Maryland. http://consensus.nih.gov/2009/dcisstatement.htm. September 24, 2009.
9. International Union against Cancer. TNM classification of malignant tumours, 4 ed. Berlin: Springer-Verlag, 1987.
10. 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. International journal of cancer 2004;110:756–62.
11. Vereniging Integrale Kankercentra. Incidence data. Utrecht. http://www.ikcnet.nl/. April 6, 2010.
12. National Evaluation Team for Breast cancer screening (NETB), National evaluation of breast cancer screening in the Netherlands 1990–2007. Twelfth evaluation report. Erasmus MC University Medical Center Rotterdam, Radboud University Nijmegen Medical Centre, 2009.
13. Vargo-Gogola T, Rosen JM. Modelling breast cancer: one size does not fit all. Nature reviews 2007;7:659–72.
14. McCann J, Treasure P, Duffy S. Modelling the impact of detecting and treating ductal carcinoma *in situ* in a breast screening programme. Journal of medical screening 2004;11:117–25.
15. Erbas B, Provenzano E, Armes J, Gertig D. The natural history of ductal carcinoma *in situ* of the breast: a review. Breast cancer research and treatment 2006;97:135–44.
16. Tsikitis VL, Chung MA. Biology of ductal carcinoma *in situ* classification based on biologic potential. Am J Clin Oncol 2006;29:305–10.

17. Feig SA. Ductal carcinoma *in situ*. Implications for screening mammography. Radiol Clin North Am 2000;38:653–68, vii.
18. Jones JL. Overdiagnosis and overtreatment of breast cancer: progression of ductal carcinoma *in situ*: the pathological perspective. Breast Cancer Res 2006;8:204.
19. Evans AJ, Pinder SE, Ellis IO, Wilson AR. Screen detected ductal carcinoma *in situ* (DCIS): overdiagnosis or an obligate precursor of invasive disease? Journal of medical screening 2001;8:149–51.
20. de Roos MA, van der Vegt B, de Vries J, Wesseling J, de Bock GH. Pathological and biological differences between screen-detected and interval ductal carcinoma *in situ* of the breast. Annals of surgical oncology 2007;14:2097–104.
21. Meijnen P, Peterse JL, Oldenburg HS, Woerdeman LA, Rutgers EJ. Changing patterns in diagnosis and treatment of ductal carcinoma *in situ* of the breast. Eur J Surg Oncol 2005;31:833–9.
22. Schouten van der Velden AP, Van Dijck JA, Wobbes T. Variations in treatment of ductal carcinoma *in situ* of the breast: a population-based study in the East Netherlands. Eur J Surg Oncol 2007;33:424–9.
23. Duffy SW, Agbaje O, Tabar L, Vitak B, Bjurstam N, Bjorneld L, Myles JP, Warwick J. Overdiagnosis and overtreatment of breast cancer: estimates of overdiagnosis from two trials of mammographic screening for breast cancer. Breast Cancer Res 2005;7:258–65.
24. de Koning HJ, Draisma G, Fracheboud J, de Bruijn A. Overdiagnosis and overtreatment of breast cancer: microsimulation modelling estimates based on observed screen and clinical data. Breast Cancer Res 2006;8:202.
25. Jorgensen KJ, Gotzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. BMJ (Clinical research ed 2009;339:b2587.
26. Morrell S, Barratt A, Irwig L, Howard K, Biesheuvel C, Armstrong B. Estimates of overdiagnosis of invasive breast cancer associated with screening mammography. Cancer Causes Control;21:275–82.
27. de Gelder R, Heijnsdijk EA, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. Epidemiol Rev 2011;33:111–21.
28. Puliti D, Miccinesi G, Paci E. Overdiagnosis in breast cancer: design and methods of estimation in observational studies. Preventive medicine 2011;53:131–3.
29. Pisano ED, Hendrick RE, Yaffe MJ, Baum JK, Acharyya S, Cormack JB, Hanna LA, Conant EF, Fajardo LL, Bassett LW, D'Orsi CJ, Jong RA, *et al.* Diagnostic accuracy of digital *versus* film mammography: exploratory analysis of selected population subgroups in DMIST. Radiology 2008;246:376–83.

Appendix 2.1 Estimated model parameters[a]

	Baseline model	Progressive model	Non-progressive model
Fraction of tumors with a preclinical screen-detectable DCIS stage (%)	18%	100%	52%
that progress to preclinical invasive cancer in the absence of screening (%)	11%	96%	0%
that will be clinically diagnosed in the absence of screening (%)	5%	4%	2%
that regress in the absence of screening (%)	2%	0%	50%
Fraction of tumors without a preclinical screen-detectable DCIS stage (%)	82%	0%	48%
Duration of preclinical breast cancer per stage at age 50 (year)			
DCIS	5.2	0.4	0.5
Invasive	2.6	2.6	2.6
Test sensitivity at screen-film mammography (%)			
DCIS	47	47	47
T1a	47	47	47
T1b	62	62	62
T1c	90	90	90
T2+	98	98	98
Test sensitivity at digital mammography (%)			
DCIS	100	94	72
T1a	47	47	47
T1b	62	62	62
T1c	90	90	90
T2+	98	98	98

[a] Fraction of preclinical screen-detectable DCIS that progress, regress, or become clinically diagnosed in the absence of screening, and the estimated mean preclinical stage durations and test sensitivities at screen-film and digital mammography. The stage durations and the fraction of preclinical screen-detectable DCIS that progress to invasive cancer, the fraction that is clinically diagnosed and the fraction that regress are age-dependent. The parameters that are presented here are calculated for age 50.

Interpreting overdiagnosis estimates in population-based mammography screening

Rianne de Gelder, Eveline A.M. Heijnsdijk, Nicolien T. van Ravesteyn, Jacques Fracheboud, Gerrit Draisma, Harry J. de Koning

Epidemiol Rev. 2011;33(1):111–21. Epub 2011 Jun 27.

ABSTRACT

Estimates of overdiagnosis in mammography screening range from 1% to 54%. This review explains such variations using gradual implementation of mammography screening in the Netherlands as an example. Breast cancer incidence without screening was predicted with a micro-simulation model. Observed breast cancer incidence (including ductal carcinoma *in situ* and invasive breast cancer) was modeled and compared with predicted incidence without screening during various phases of screening program implementation. Overdiagnosis was calculated as the difference between the modelled number of breast cancers with and the predicted number of breast cancers without screening. Estimating overdiagnosis annually between 1990 and 2006 illustrated the importance of the time at which overdiagnosis is measured. Overdiagnosis was also calculated using several estimators identified from the literature. The estimated overdiagnosis rate peaked during the implementation phase of screening, at 11.4% of all predicted cancers in women aged 0–100 years in the absence of screening. At steady-state screening, in 2006, this estimate had decreased to 2.8%. When different estimators were used, the overdiagnosis rate in 2006 ranged from 3.6% (screening age or older) to 9.7% (screening age only). The authors concluded that the estimated overdiagnosis rate in 2006 could vary by a factor of 3.5 when different denominators were used. Calculations based on earlier screening program phases may overestimate overdiagnosis by a factor 4. Sufficient follow-up and agreement regarding the chosen estimator are needed to obtain reliable estimates.

3.1 INTRODUCTION

Mammography screening has been shown to be effective in reducing breast cancer mortality,[1–4] but the magnitude of the harms of screening is less well established. One of the harms of screening is overdiagnosis: detection of breast cancers that would not have become symptomatic during a woman's lifetime if no screening had taken place.

Screening is expected to increase the observed incidence of breast cancer among women in the targeted age group, partly because of overdiagnosis but also because of the advanced diagnosis of breast cancer. At prevalence screening, mammography may detect breast cancers from a pool of preclinical tumors that exist in a population, which increases the observed incidence. At subsequent screens, future incidence trends – breast cancer incidence in industrialized countries increases by calendar time[4] – are brought forward in time, which may also lead to an excess of breast cancers compared with a situation without screening.

Theoretically, those tumors for which the diagnosis is advanced by screening will not be diagnosed when they would have been if no screening had taken place. The period during which the diagnosis is advanced is called 'lead time'. When the lead time has elapsed, the incidence among previously screened women is expected to fall to a level below that predicted without screening ('deficit incidence'). The deficit in incidence in the previously screened age group is expected to balance out the excess in incidence in the screening ages. In practice, this is not entirely the case[5–7] and is referred to as overdiagnosis.

The frequency at which overdiagnosis occurs is a topic of strong debate. In a meta-analysis of overdiagnosis in randomized breast cancer screening trials and various population-based screening programs, Biesheuvel *et al.*[8] found that estimates ranged between –4% and 54% of all expected cancers (including invasive breast cancers only). More recent analyses estimated the rate of overdiagnosis to be 52%–54% of all expected cancers without screening in women of the screening age, meaning that 1 of 3 cancers in a screened population is overdiagnosed.[5, 9] Modeling studies, on the contrary, estimated overdiagnosis to be between 1% of all diagnosed breast cancers in a screened population[6] and 3% of all predicted breast cancers in the total population.[10]

The question thus arises regarding why overdiagnosis estimates differ to such an extent between those studies. In this analysis, we discuss differences between key studies of overdiagnosis. Using the gradual implementation of nationwide breast cancer screening of more than 1 million Dutch women as an example, we focus on the importance of

the time at which overdiagnosis is measured. To enable a comparison between various estimates of overdiagnosis, we calculate rates using several different denominators.

3.2 MATERIALS AND METHODS

MIcro-simulation SCreening ANalysis model (MISCAN)

Biennial mammography screening in the Netherlands started in 1990 and was gradually implemented in the whole country between 1990 and 1997 ('implementation phase'), targeting women in the age group 49–69 years. Between 1998 and 2001, the screening program was extended to women aged 49–74 years ('extension phase'). A 'steady-state phase' of screening was reached in 2002, when the number of first screening examinations and subsequent examinations with an interval of more than 30 months from the previous screening remained stable. Annually, more than 1 million women are invited to participate in the program; 82% of the invited population attends screening.

The implementation, extension, and steady-state phases of the Dutch screening program were modeled in MISCAN,[11, 12] designed to assess the effects of screening on a population. The model consisted of 2 parts: a part in which 1) the individual life histories of women in a nonscreened population were simulated, and 2) a part in which a screening program was modeled and the influence of mammography screening on these life histories was determined. We modelled the Dutch female population aged 0–100 years in 1989. In the model, some of these women may develop preclinical invasive breast cancer during their lives, which may or may not be preceded by screen-detectable ductal carcinoma *in situ* (DCIS). In the absence of screening, preclinical screen-detectable DCIS may progress to invasive cancer, become clinically diagnosed, or regress. Preclinical invasive tumors may grow into a successively larger preclinical stage of disease, as a Markov-like stage transition process. They also may become symptomatic and consequently diagnosed. If screening takes place, preclinical lesions can also become screen detected, depending on their size and the sensitivity of the test.

The rates at which these transitions occur, the mean duration of preclinical DCIS and invasive cancer, and the sensitivity of screening mammography were estimated by using data from the Comprehensive Cancer Centers[13] and screening organizations in the Netherlands.[14] These data included the observed age-specific incidence in the Netherlands between 1990 and 2006; the age-specific and stage-specific incidence of clinically diagnosed and screen-detected cancer; the age-, stage-, and screening round-specific cancer detection rates; and interval cancer rates since the start of screening. By minimizing the deviance between observed and modeled breast cancer incidence, screen detection, and

interval cancer rates, the optimal model parameters were chosen. A chi-square test was used to test goodness of fit. With the model, the observed incidence in the Netherlands and various other countries could be reproduced reasonably.[15–17] We assumed that no mammography screening took place outside the organized program because screening participation is high, especially among women who have previously attended (82% in the target population, 95% of previous attendees).[14] Furthermore, a survey of women older than the screening age showed no evidence for opportunistic screening.[18]

Model parameters

In the best-fitting model, we estimated the duration of preclinical DCIS to be Weibull distributed, with a mean of 2.6 years at all ages. The duration of preclinical invasive cancer has an exponential distribution with a mean estimated to increase by age, from 1.0 year at age 20 to 3.9 years from age 65 onward. The sensitivity of mammography was 72% for DCIS, 47% for stage T1a, 62% for stage T1b, 90% for stage T1c, and 95% for stage T2+, and it was assumed not to change over time. We further estimated that 18% of all tumors had a screen-detectable DCIS stage, of which 11% progressed to invasive breast cancer, 5% was clinically diagnosed, and 2% regressed.

Overdiagnosis calculations

To calculate the rate of overdiagnosis in the Netherlands, we modeled breast cancer incidence in the presence of screening. The model that fitted best to the observations was then compared with predicted breast cancer incidence in the absence of screening. With this approach, screened and nonscreened populations were exactly the same, with a similar background risk of developing breast cancer. Both DCIS and invasive cancer were included in the overdiagnosis estimate, because, in the model, overdiagnosis can occur when

1. preclinical DCIS, detected by screening, would have regressed if no screening had taken place (Figure 3.1a);
2. preclinical DCIS, detected by screening, would not have progressed to invasive cancer during a woman's lifetime if no screening had taken place (Figure 3.1b);
3. preclinical DCIS, detected by screening, would have progressed to invasive cancer but would not have become symptomatic during a woman's lifetime if no screening had taken place (Figure 3.1c); and
4. preclinical invasive breast cancer, detected by screening, would not have become symptomatic during a woman's lifetime if no screening had taken place (Figure 3.1d).

The estimator for the overdiagnosis rate was $(E–D)/T_{0,\ \text{age 0–100 years}}$. E represents the number of excess breast cancers in women of screening age, calculated as the difference in the

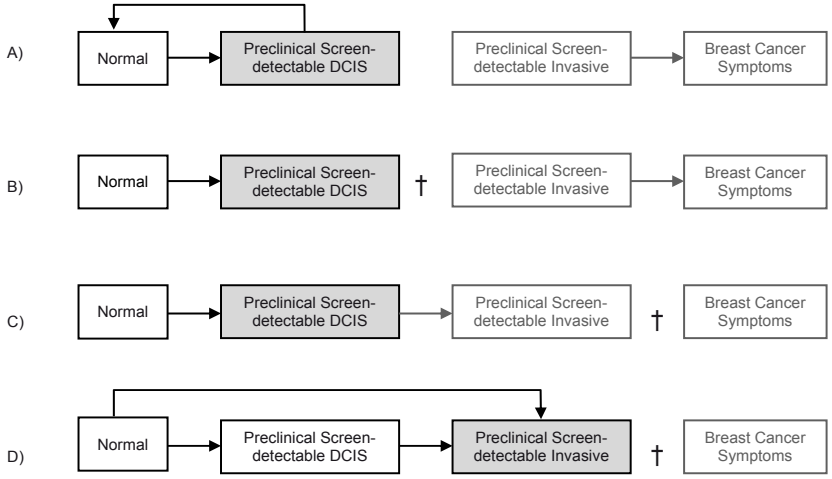


Figure 3.1a–d Stages at which overdiagnosis can occur: a) When preclinical ductal carcinoma in situ (DCIS), detected by screening, would have regressed if no screening had taken place; b) when preclinical DCIS, detected by screening, would not have progressed to invasive cancer during a woman's lifetime if no screening had taken place; c) when preclinical DCIS, detected by screening, would have progressed to invasive cancer but would not have become symptomatic during a woman's lifetime if no screening had taken place; and d) when preclinical invasive breast cancer, detected by screening, would not have become symptomatic during a woman's lifetime if no screening had taken place. Dot-filled boxes; stage at which a breast cancer is screen detected; grey-shaded boxes: stages of the natural history of the tumor averted by screen detection; crosses: death from causes other than breast cancer.

modelled number of breast cancers with and the predicted number of cancers without screening. D is the number of deficit breast cancers in the age groups exceeding the screening limit, calculated as the difference in the predicted number of breast cancers without and the modelled number of cancers with screening. $T_{0,\ \text{age 0–100 years}}$ represents the total number of breast cancers predicted in a population aged 0–100 years without screening.

To illustrate the extent to which overdiagnosis estimates are influenced by the denominator used, various estimators are applied to the modeled and predicted breast cancer incidence in the Netherlands. We searched the PubMed literature to identify alternative estimators to calculate overdiagnosis. With the query ''('breast neoplasms'[MeSH Terms] OR ('breast'[All Fields] AND 'cancer'[All Fields]) OR 'breast cancer'[All Fields]) AND ('overdiagnosis'[All Fields] OR 'over-diagnosis'[All Fields] OR 'overdetection'[All Fields] OR 'over-detection'[All Fields]),'' a total of 158 titles were obtained. Only primary research or review articles in English that gave explicit estimates of overdiagnosis in breast cancer screening trials and population-based mammography screening were

considered relevant. Using these criteria, we included a total of 15 papers. On the basis of the literature references in these articles, 1 other paper was also included. Data on the denominator used to define the population at risk, the time period of screening, and the length of follow-up after screening ended were extracted from each study. An overview of the 16 studies obtained, with their estimates of overdiagnosis, is presented in Table 3.1. The studies were grouped by the estimator used to calculate overdiagnosis:

1. $(E–D)/T_{0,\ age\ 0–100\ years}$, which is the relative increase in breast cancers due to overdiagnosis (E–D) compared with the predicted number of breast cancers in the female population aged 0–100 years in a situation without screening. This estimator was used by de Koning *et al.*, [10] who estimated the overdiagnosis rate to be 3%.
2. $(E–D)/T_{0,\ screening\ age\ and\ older}$, which is the relative increase in breast cancers due to overdiagnosis (E–D) compared with the predicted number of breast cancers in women of the screening age and older in a situation without screening. Three previous studies used this estimator, with overdiagnosis estimates ranging between 1% and 30.5%.[7, 19, 20]
3. $(E–D)/T_{0,\ screening\ age}$, which is the relative increase in breast cancers due to overdiagnosis compared with the predicted number of breast cancers in women of the screening age in a situation without screening. This method was used in 3 studies,[5, 21–23] with overdiagnosis estimates varying between 4.6% and 52%.
4. $(E–D)/T_{1,\ screening\ age}$, which is the fraction of overdiagnosed cancers of all diagnosed breast cancers in women of the screening age in a situation with screening. The 3 studies that used this estimator assessed overdiagnosis to be 1%–12%.[6, 24, 25]
5. (E–D)/SD, which is the fraction of all screen-detected (SD) cancers that is overdiagnosed. Welch and Black[26] recalculated the results of the Swedish Malmö trial[7] with this denominator and estimated the overdiagnosis rate to be 24%.
6. $T_{1,\ screening\ age}/T_{0,\ screening\ age}$, which is the relative risk of breast cancer for women of the screening age in a situation with screening compared with the predicted number of breast cancers in women of the same age in a situation without screening. The estimator can be corrected for lead time, for instance, by shifting the predicted incidence without screening forward in time. Three studies used this method[9, 27, 28]; their overdiagnosis estimates ranged between –4% and 54%.
7. $T_{1,\ screening\ age}/(T_{1,\ screening\ age,\ corrected})$, which is the relative risk of breast cancer for women of the screening age in a situation with screening compared with the predicted number of tumors in a situation with screening if no overdiagnosis would take place ($T_{1,\ screening\ age,\ corrected}$). This method was used by Martinez-Alonso *et al.*, [29] who estimated the overdiagnosis rate to range between 0.4% and 46.6%.

Table 3.1 Estimators for Overdiagnosis and Follow-up Time to Correct for Lead Time, as Reported in the Literature

	Estimator	Method used: First Author, Year Reference No.)	Follow-up allowed to Correct for Lead Time	Overdiagnosis Estimate
1.	(E–D)/ $T_{0,\text{ age 0–100 years}}$	de Koning, 2006[10]	Modeled follow-up during remaining lifetime, in the steady-state phase of the screening program	3%
2.	(E–D)/ $T_{0,\text{ screening age and older}}$	Moss, 2005[19]	5–13 years of follow-up after randomization	–5.8% to 30.5%
		Zackrisson, 2006[7]	15 years of follow-up after the trial ended	10%
		Puliti, 2009[20]	5–10 years of follow-up past the screening age	1% to 13%
3.	(E–D)/ $T_{0,\text{ screening age}}$	Paci, 2006[21]	Modeled follow-up during remaining lifetime, in the first 5 years of the screening program	4.6%
		Jorgensen, 2009 (4 countries)[5]	Follow-up of 7–9 years after full implementation of screening or 10–11 years after the program started	52%
		Jorgensen, 2009 (Denmark)[22]	Follow-up of 2–10 years after full implementation of program	33%
4.	(E–D)/ $T_{1,\text{ screening age}}$	Duffy, 2005[6]	Modeled follow-up during remaining lifetime, at the end of the screening trial	1% to 2%
		Olsen, 2006[24]	Modeled follow-up during remaining lifetime, in the first 2 screening rounds	4.8%
		Duffy, 2010 (Sweden)[25]	Modeled follow-up during remaining lifetime, at the end of the screening trial	12%
5.	(E–D)/SD	Welch, 2010[26]	15 years of follow-up after the trial ended	24%
6.	$T_{1,\text{ screening age}}$/ $T_{0,\text{ screening age}}$	Zahl, 2004[27]	1–4 years of follow-up after full implementation of the screening program.	45% to 54% (excluding DCIS)
		Jonsson, 2005[28]	7–15 years of follow-up since screening started	–4% to 54% (excluding DCIS)
		Morrell, 2009[9]	4–6 years of follow-up after full implementation of the program	30% to 42% (excluding DCIS)
7.	$T_{1,\text{ screening age}}$/ $T_{1,\text{ screening age, corr}}$	Martinez-Alonso, 2010[29]	Modeled follow-up during remaining lifetime, since screening started	0.4% to 46.6%

Abbreviations: D, number of deficit breast cancers in the age groups exceeding the screening limit, calculated as the difference in the number of breast cancers without and with screening; DCIS, ductal carcinoma in situ; E, number of excess breast cancers in the screening ages, calculated as the difference in the number of breast cancers with and without screening; SD, number of screen-detected cancers; T_0, predicted number of breast cancers in the absence of screening; T_1, modeled total number of breast cancers in the presence of screening; $T_{1,\text{ corr}}$, total number of breast cancers in the presence of screening minus the number of overdiagnosed cancers

Overdiagnosis was calculated by applying estimators 1–6 to the modeled and predicted numbers of breast cancers with and without screening in the Netherlands. Doing so demonstrates the impact of using different denominators on the estimated overdiagnosis rate. The overdiagnosis rate using estimator 6 ($T_{1,\text{ screening age}}/T_{0,\text{ screening age}}$) was calculated without a correction for lead time. A lead-time correction – for instance, by shifting the expected incidence without screening 2.5 years forward in age (comparable to the studies by Morrell *et al.*[9] and Jonsson *et al.*[28] – should result in an estimate in between those of estimators 3 and 6. Estimator 7 was not used, because it is not possible in MISCAN to model the incidence of breast cancer without assuming some degree of overdiagnosis. The outcomes were compared with the overdiagnosis rate obtained by using estimator 1 $(E–D)/T_{0,\text{ age 0–100 years}}$.

To illustrate the importance of the time at which overdiagnosis is estimated, the rate was calculated for each year between 1990 and 2006, during the implementation, extension, and steady-state phases of the screening program. Only in a steady state will the estimators provide an unbiased estimate of overdiagnosis.

3.3 RESULTS

From the moment that screening started in the Netherlands, observed breast cancer incidence among women of the screening ages increased (Figure 3.2a–j). Related to the growing number of women screened and the relatively high proportion of prevalence screens during the implementation phase of the program, the difference between the observed incidence rate of women of the screening ages and the invitation rate in the population no longer increased and excess incidence remained stable. Because part of the women aged 50–54 years reached the lower age limit for screening and had a prevalence screening, their incidence was higher than that of women aged 55–59 years, of whom the majority as invited for a subsequent screening round at this time.

In 1998, the upper age limit for the screening program was extended to women aged 70–74 years. At the peak of this extension phase, in 1999, the number of invited women and screening examinations with an interval of more than 2.5 years from the previous screening examination rose strongly, resulting in a higher detection rate. Consequently, excess breast cancer incidence among women of the screening ages also increased sharply (Figure 3.2f). The excess dropped again when all women aged 70–74 years had been reinvited to screening at least once (in 2002, Figure 3.2h) and a 'steady-state' phase of the screening program was reached. From 2002 onward, the excess incidence among women of the screening ages remained fairly constant.

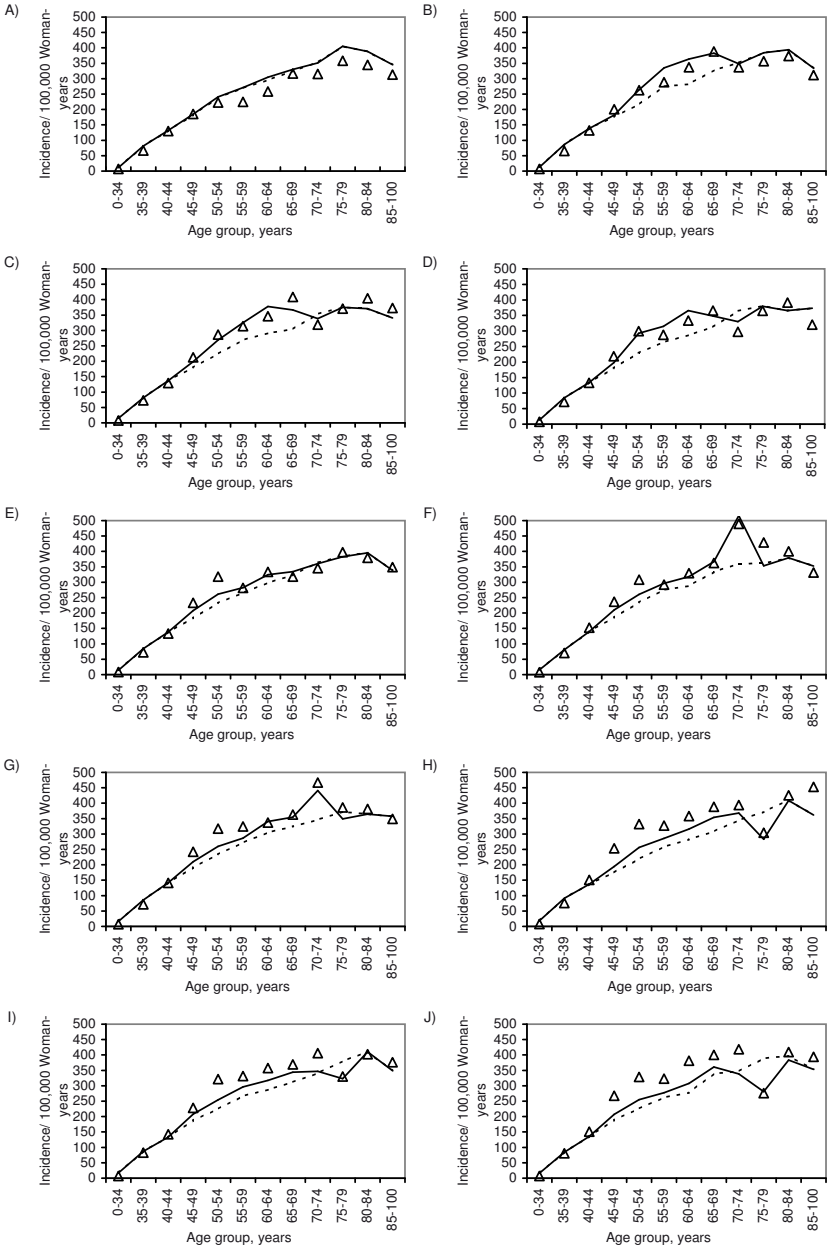


Figure 3.2 Observed and modeled breast cancer incidence per 100,000 woman-years in the presence and absence of screening between 1990 and 2006 (values after years indicate percentage of the target population aged 49–69 years invited, fraction of prevalent screenings). a) 1990: 9.2%, 74%; b) 1992: 47.4%, 77%; c) 1994: 74.3%, 49%; d) 1996: 92.0%, 39%; e) 1998: 80.8%, 20%; f) 1999: 1.8%, 19%; g) 2000: 94.4%, 18%; h) 2002: 96.1%, 14%; i) 2004: 95.8%, 14%; j) 2006: 92.2%, 13%. Solid lines, modeled with screening; dashed lines, modeled without screening; triangles, observed.

In the age groups that passed the screening age (≥70 years between 1990 and 1997, ≥75 years from 1998 onward), the observed breast cancer incidence dropped to a level lower than the predicted incidence without screening (Figure 3.2). Because of lead time, generally estimated to be between 2 and 4 years,[20, 30] the drop in incidence among women no longer screened is predicted to occur 2–4 years later than the increase in the screening ages, when all tumors would have been clinically diagnosed if no screening had taken place. Indeed, from 1994 onward, a deficit in breast cancer incidence was observed. From the moment that the majority of women were invited for a subsequent screening, in 1996, the deficit reached its maximum. The deficit in the incidence rate almost disappeared in the year the screening program was extended to include women aged 70–74 years. This extension phase lasted until 2001; the deficit in incidence among women aged 75 years or older was expected to be observed between 2003 and 2005. Indeed, the deficit increased during these years.

Our overdiagnosis estimates were based on the modelled incidence of breast cancer and the predicted incidence without screening. Overall, the model reproduced the observed incidence reasonably well. Between 1990 and 1993, however, the simulated incidence among women of the screening ages (50–69 years) was higher than observed, whereas, between 2001 and 2006, the modeled incidence was lower. When the modeled breast cancer incidence in a screening situation was compared with the predicted incidence without screening, the estimated overdiagnosis rate in the total population during the implementation phase of screening increased from 1.0% of all predicted breast cancers in 1990 to 11.4% in 1993 (Table 3.2). In 1993, the modelled excess in breast cancers peaked (17.1% of all predicted cancers in women aged 50–69 years), while the modelled deficit in incidence among women no longer screened was 0.8% of all predicted cancers in that age group. The estimate of overdiagnosis decreased the more women had subsequent screens, to 5.6% in 1997.

During the extension phase, the overdiagnosis estimate increased to 10.0% in 1999, after which it decreased to 4.7% in 2001. During the steady-state phase of screening, the estimate first increased to 4.9% in 2003 but then dropped to 2.8% of all predicted breast cancers in 2006. In 2006, the excess of breast cancers in the age group was 7.0%; the deficit was 11.7% (Table 3.2). Most of the deficit was expected directly when screening ceased: in the age group 70–74 years before 1998 and in the age group 75–79 years from 1998 onward. In 2006, a small deficit was also predicted among women aged 80–84 years.

Depending on the denominator used to define the population at risk, the overdiagnosis estimate at steady-state screening may increase to 8.9% if the rate is calculated as a

Table 3.2 Predicted Excess and Deficit in Breast Cancers and Overdiagnosis in the Netherlands[a]

Phase and Years	$T_{0,\ \text{age 0–69/74 years}}$	$T_{1,\ \text{age 0–69/74 years}}$	$E_{\text{age 0–69/74 years}}$, %	$T_{0,\ \text{age 69/74–100 years}}$	$T_{1,\ \text{age 69/74–100 years}}$	$D_{\text{age 69/74–100 years}}$, %	$E_{\text{age 0–69/74 years}} - D_{\text{age 69/74–100 years}}$	$(E-D)/T_{0,\ \text{age 0–100 years}}$, %
Implementation phase								
1990–1991	15,237	15,481	1.6	7207	7197	0.1	234	1.0
1991–1992	15,646	17,065	9.1	7201	7184	0.2	1402	6.1
1992–1993	15,606	17,719	13.5	7240	7214	0.4	2087	9.1
1993–1994	15,695	18,381	17.1	7458	7400	0.8	2628	11.4
1994–1995	16,039	18,490	15.3	7499	7405	1.3	2357	10.0
1995–1996	16,149	18,550	14.9	7821	7669	1.9	2249	9.4
1996–1997	16,235	18,608	14.6	7877	7628	3.2	2124	8.8
1997–1998	16,646	18,291	9.9	7958	7686	3.4	1373	5.6
Extension phase								
1998–1999	19,506	20,746	6.4	5404	5392	0.2	1228	4.9
1999–2000	19,779	22,368	13.1	5488	5433	1.0	2534	10.0
2000–2001	20,043	22,108	10.3	5675	5517	2.8	1907	7.4
2001–2002	20,375	21,892	7.4	5841	5560	4.8	1236	4.7
Steady-state phase								
2002–2003	20,371	21,961	7.8	5892	5538	6.0	1236	4.7
2003–2004	20,601	22,336	8.4	5965	5533	7.2	1303	4.9
2004–2005	20,471	22,127	8.1	5908	5377	9.0	1125	4.3
2005–2006	20,984	22,741	8.4	5857	5288	9.7	1188	4.4
2006–2007	21,087	22,569	7.0	6136	5421	11.7	767	2.8

[a] The percentage of excess (E) breast cancers in the age group 0–69/74 years was calculated as $(T_{1,\ \text{age 0–69/74 years}} - T_{0,\ \text{age 0–69/74 years}})/T_{0,\ \text{age 0–69/74 years}}$. T_1, modeled number of breast cancers in the presence of screening; T_0, predicted number of breast cancers in the absence of screening. The percentage of deficit (D) breast cancers was calculated as $(T_{0,\ \text{age 69/74–100 years}} - T_{1,\ \text{age 69/74–100 years}})/T_{0,\ \text{age 69/74–100 years}}$. Overdiagnosis was then calculated as the number of excess cancers in the age group 0–69/74 years minus the number of deficit cancers in the age group 69/74–100 years divided by the total number of breast cancers in the absence of screening in women aged 0–100 years.

fraction of all screen-detected cancers (Table 3.3, estimator 5, 2006). This rate is 3.2 times higher than the estimate that uses all predicted breast cancers in women aged 0–100 years in the denominator (Table 3.3, estimator 1, 2006) but has the same numerator. If calculated as a fraction of all diagnosed tumors among women of the screening age in a screening situation (Table 3.3, estimator 4, 2006), the estimate would be 4.6%. The estimated rates of overdiagnosis calculated as a relative increase among women of the screening age and older (Table 3.3, estimator 2, 2006) or women of the screening age only (Table 3.3, estimator 3, 2006) were 3.6% and 5.0%, respectively. Without an adjustment for lead time, the overdiagnosis rate calculated for women of the screening age

Table 3.3 Overdiagnosis Estimates in the Netherlands Using Various Estimators[a]

Phase and Years	Estimator					
	1: $(E–D)/T_{0,\text{ age 0–100 years}}$, %	2: $(E–D)/T_{0,\text{ age 49–100 years}}$, %	3: $(E–D)/T_{0,\text{ age 49–69/74 years}}$, %	4: $(E–D)/T_{1,\text{ age 49–69/74 years}}$, %	5: $(E–D)$/SD, %	6: $T_{1,\text{ 49–69/74 years}}/T_{0,\text{ 49–69/74 years}}$, %
Implementation phase						
1990–1991	1.0	1.4	2.4	2.3	35.4	2.3
1991–1992	6.1	8.2	14.1	12.4	67.4	14.3
1992–1993	9.1	12.2	21.3	17.5	61.5	21.6
1993–1994	11.4	15.2	26.7	21.0	54.7	27.3
1994–1995	10.0	13.3	23.2	18.7	44.5	24.0
1995–1996	9.4	12.4	21.8	17.7	38.2	23.3
1996–1997	8.8	11.6	20.3	16.5	32.6	22.7
1997–1998	5.6	7.3	12.7	11.0	22.1	15.2
Extension phase						
1998–1999	4.9	6.5	9.0	8.3	18.9	9.1
1999–2000	10.0	13.1	18.2	15.4	30.4	18.6
2000–2001	7.4	9.7	13.6	11.8	23.0	14.7
2001–2002	4.7	6.1	8.7	7.8	15.4	10.6
Steady-state phase						
2002–2003	4.7	6.1	8.6	7.7	15.2	11.1
2003–2004	4.9	6.3	8.9	8.0	15.6	11.9
2004–2005	4.3	5.5	7.7	6.9	13.2	11.4
2005–2006	4.4	5.7	7.9	7.0	13.6	11.6
2006–2007	2.8	3.6	5.0	4.6	8.9	9.7

[a] *E–D* is the number of excess breast cancers (*E*) minus the number of deficit breast cancers (*D*). The excess is calculated as the difference between the modeled number of breast cancers with (T_1) and the predicted number of breast cancers without screening (T_0) in the screened age group, the deficit is calculated as the difference in the predicted number of breast cancers without and the modeled number of cancers with screening in the age groups past the screening age.
Abbreviation: SD, number of screen-detected cancers

only would be 9.7%: 3.5 times higher than the baseline estimate (Table 3.3, estimator 6, 2006). Overdiagnosis also depended on the year it was measured. Calculations based on years in which a screening program was not yet fully implemented were 4 times higher than estimates based on steady screening (Table 3.3, estimator 1, 1993 *vs.* 2006). The estimated overdiagnosis rate by year of measurement and by estimator is shown in Table 3.3.

3.4 DISCUSSION

The estimated overdiagnosis rate peaked during the implementation phase of screening at 11.4% of all predicted cancers in women aged 0–100 years in the absence of screening. Five years after implementation was completed, in 2006, this estimate had decreased to 2.8%. If different estimators were used, the overdiagnosis rate in 2006 would range between 3.6% (screening age and older) and 9.7% (screening age only). The estimate of overdiagnosis is thus strongly dependent on the time it was calculated and the denominator used to define the population at risk.

Our findings seem to strongly differ from those in some recent publications, with estimated overdiagnosis rates up to approximately 50%.[5, 8, 9, 27] This paper may perhaps not resolve the controversy, but it does explain why reported epidemiologic estimates may differ to such an extent. Using gradual implementation of nationwide breast cancer screening of more than 1 million Dutch women, we illustrated that a steady-state screening situation and sufficient follow-up to allow for lead time are crucial to observing a deficit in breast cancer incidence and to calculating overdiagnosis correctly. In several studies, the first years of the screening program were included in the overdiagnosis estimate.[21, 23, 29] A relatively large proportion of women will have a prevalence screen in these years, which will increase the number of excess breast cancers. In the worst-case scenario, this could have resulted in overestimation of the overdiagnosis rate by a factor of 4 (Table 3.3, estimator 1, 1993 *vs.* 2006). Of course, first (prevalent) screening rounds of women who reach the lower age limit for screening should be included in an overdiagnosis estimate. However, the proportion of women reaching this age will be stable only during steady-state screening.

Several studies based their analyses on the period after implementation of screening but still may not have fully accounted for lead time[5, 9, 22, 27] because they calculated overdiagnosis by using average breast cancer incidence during this phase. However, even during the steady-state phase, overdiagnosis may further drop by a factor of 1.7 (Table 3.3, estimator 1, 2003 *vs.* 2006) as the number of women contributing to the deficit in incidence still increases. A compensatory drop in incidence will reach its maximum only if all women in the age group past the screening age had been invited to screening when they were eligible. Moreover, some tumors may have a lead time longer than 5 years. On the basis of the estimated distribution of the lead time of breast cancer in our study (the best-fitting model assumed a Weibull-distributed lead time with a median of 2 years and a mean lead time of 3.7 years), approximately 20% of all tumors will have a lead time of more than 5 years, and 5% will have a lead time of more than 10 years. Ideally, the lead

time of these tumors should be accounted for by calculating overdiagnosis several years after screening has reached the steady-state phase.

Overdiagnosis estimates will be affected by the denominator used to define the population at risk. Various estimators were used in this study, resulting in overdiagnosis estimates differing by a factor of 3.5. By calculating overdiagnosis for women of all ages, we also included women who will never be screened and will not be at risk of overdiagnosis. If overdiagnosis is calculated as a relative risk for women of the screening ages only,[5, 21, 22] overdiagnosis could be 1.8 times higher than if women of all ages are included (Table 3.3, estimator 3 *vs.* estimator 1, 2006). However, the impact of a screening program is sometimes observed at a later age; by limiting the denominator to the screened age group, the lifetime effect of screening is not given justice. Alternatively, the risk of overdiagnosis can also be calculated for women of the screening age and older.[7, 19, 20, 23] In this case, the overdiagnosis estimate would be 1.3 times higher than when women of all ages are included (Table 3.3, estimator 2, 2006). Calculated as the fraction of all breast cancers diagnosed in women of the screening age in a situation with screening, comparable to the estimates by Duffy *et al.*[6, 25] and Olsen *et al.*,[24] the estimate in the Netherlands would be 4.6% (Table 3.3, estimator 4, 2006). If overdiagnosis is calculated as a fraction of screen-detected cancers, the estimate would increase by a factor of 3.2 (Table 3.3, estimator 5, 2006). A comparable finding was shown by Welch and Black,[26] who demonstrated that the overdiagnosis rate in the Malmö trial, previously estimated to be 10%,[7] would be 24% if only screen-detected cancers were taken into account. The choice of the denominator will likely depend on the purpose of the overdiagnosis estimate. If the population risks of different screening regimens – for instance, with varying starting ages for screening – are compared, the denominator that includes all diagnosed breast cancers in women aged 0–100 years may be useful. If the main purpose is to inform individual women of their risk of being overdiagnosed, the denominator that includes women of the screening age and older may be more useful. The fraction of screen-detected cancers overdiagnosed may be relevant in evaluating the performance of a particular screening program or in treatment decisions for DCIS.

Varying overdiagnosis estimates could also be explained by differences in screening characteristics. For instance, the more women are screened, the more likely that an irrelevant tumor is detected. Thus, shorter screening intervals and higher attendance rates may increase the overdiagnosis rate. Overdiagnosis could also be affected by referral or recall practice. If the threshold for diagnostic assessment of small or obscure lesions is higher, fewer of these tumors may be detected or overdiagnosed. The fraction of tumors that are noninvasive may also influence overdiagnosis. In the United States, for instance,

17%–34% of all screen-detected cancers are DCIS,[31] whereas, in the Netherlands, this fraction is somewhat lower (16%).[14]

Overdiagnosis estimates will also be affected by the age of the screened group. In the Malmö trial, for instance, the study group was 45–69 years of age at randomization. At 15 years of follow-up, they will be 60–84 years of age. Because tumors grow slower at older ages, and because mortality from causes other than breast cancer increases with age, such trial-based estimates will be higher than overdiagnosis estimates based on ongoing screening programs that have a constant inflow of women in the lower age limit of screening. Another factor that might bias the overdiagnosis rate is screening of women in the age groups no longer eligible for screening[5, 9] or screening in the control group of a trial.[7] For instance, an estimated 24% of women in the control group of the Malmö trial were thought to be screened.[19]

Use of mathematical modeling to calculate overdiagnosis has certain limitations. Overdiagnosis estimates will be affected by model assumptions about the natural history of breast cancer. Previous studies showed that model parameters, such as test sensitivity, mean duration of the preclinical phase of cancer, and probability of preclinical DCIS to progress to invasive cancer or to regress, can be interchanged to some extent[32] (R. de Gelder, Erasmus MC, Department of Public Health, unpublished manuscript). This means that, for instance, a model with a higher progression and lower regression rate of preclinical DCIS could simulate observed breast cancer incidence equally well as a model with a lower progression and higher regression rate, provided that test sensitivity, mean duration, and onset rate of breast cancer are adjusted accordingly. This might affect the overdiagnosis rate.

In the present study, the probability of DCIS progression at age 50 years was estimated to be 61% and the probability of regression to be 11%. If we instead assumed that 0% of preclinical DCIS would progress and 96% would regress, the predicted overdiagnosis rate would be 8.1% of all predicted cancers in 2006 (data not shown). This rate is still considerably lower than the overdiagnosis estimates published elsewhere.[5, 9, 27] Because the natural history of DCIS is unobservable, no direct evidence exists on the 'true' progression and regression rate.

However, indirect evidence suggests that the assumed progression rate in our model is plausible. For instance, follow-up of undertreated DCIS initially misdiagnosed as benign shows that 11%–60% of all DCIS recurs as invasive cancer within 10–20 years.[32] Furthermore, basement membrane invasion has been observed in DCIS cases, and microscopy on invasive lesions showed that DCIS was present in 20%–30% of the carcinomas.[33] Lit-

erature provides no evidence on the fraction of preclinical DCIS that regresses. Because of structural model uncertainties, it would be useful to assess the overdiagnosis rate in one particular screening situation with various collaborating models, such as those in the Cancer Intervention and Surveillance Modeling Network (CISNET).[34] These models share the same input but vary in their structure and assumptions, which reflects the uncertainties about the natural history of breast cancer. Only some of the CISNET models incorporated DCIS or assigned low malignant potential to a fraction of the tumors,[34] which would of course affect overdiagnosis estimates.

The present study is limited by the fact that from 2002 onward, the modeled breast cancer incidence is lower than the observed incidence among women of the screening ages (Figure 3.2). However, increasing the sensitivity of mammography in the model did not result in a substantial increase in the modeled incidence, without affecting the predicted number of interval cancers or clinically diagnosed cancers.[14]

The difference between observed and modeled incidence rates should therefore be explained by an increasing trend in background incidence. This could happen when breast cancer incidence increases because of, for instance, a rising prevalence of risk factors such as lower parity, older age at birth of the first child, or obesity. Increasing incidence trends have been observed before implementation of the Dutch screening program and in unscreened women.[35, 36] Future modeling efforts should take such background trends into account. Because increases in background incidence are likely to occur in both screened and unscreened women at an approximately similar rate, it is unlikely that our overdiagnosis estimate, which did not include the secular increase in incidence, was affected by poor model fit in recent years. If we conservatively assume that improvement in the sensitivity of mammography would have increased the excess incidence by 10%, the overdiagnosis estimate at steady-state screening would be 3.4%–5.6%. Although the modeled incidence was lower than the actual observed incidence in recent years, we modeled the observed incidence between 1990 and 2002 fairly well. Moreover, the model reproduced the observed incidence decline in women who are no longer screened accurately for all observation years.

Despite several limitations, our approach to calculating overdiagnosis has certain advantages. By using a model, screened and nonscreened populations could be exactly the same, with a similar background risk of developing breast cancer. Studies that, for instance, compare screened and historical comparison groups have the disadvantage that temporal incidence trends may have affected one group but not the other (e.g., by the use of hormone replacement therapy in the late 1990s). Moreover, our approach has the advantage that it is based on observed data from a long-running, population-based

mammography screening program with high participation rates (82% in the target population, 95% of previous attendees) that annually targets more than a million women. The observed data on which the model was based include clinically diagnosed breast cancers, screen-detection rates, and interval cancer rates. Natural history parameters, such as lead time, could be estimated from these data. The observations show that the incidence of breast cancer strongly decreases after women have reached the upper age limit for screening. This finding strongly suggests that the risk of overdiagnosis must be smaller than recent estimates of approximately 50%.[5, 8, 9, 27]

In conclusion, our estimates of overdiagnosis are substantially lower than those published in recent literature. This discrepancy is most likely related to methodological differences between studies and lack of sufficient follow-up, and partly to differences in screening characteristics and performance. In 2006, the estimated risk of overdiagnosis in the Netherlands ranged between 2.8% of all predicted cancers in women aged 0–100 years in the absence of screening and 9.7% of all predicted cancers in women of the screening age only.

REFERENCES

1. de Koning HJ. Mammographic screening: evidence from randomised controlled trials. Ann Oncol. 2003;14(8):1185–1189.
2. International Agency for Research on Cancer. Volume 7: breast cancer screening. In: Vainio H, Bianchini F, eds. IARC Handbooks of Cancer Prevention. Lyon, France: IARC Press; 2002.
3. Allgood PC, Warwick J, Warren RM, *et al.* A case-control study of the impact of the East Anglian breast screening programme on breast cancer mortality. Br J Cancer. 2008;98(1):206–209.
4. Waller M, Moss S, Watson J, *et al.* The effect of mammographic screening and hormone replacement therapy use on breast cancer incidence in England and Wales. Cancer Epidemiol Biomarkers Prev. 2007;16(11):2257–2261.
5. Jørgensen KJ, Gøtzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. BMJ. 2009;339:b2587. (doi: 10.1136/bmj.b2587).
6. Duffy SW, Agbaje O, Tabar L, *et al.* Overdiagnosis and overtreatment of breast cancer: estimates of overdiagnosis from two trials of mammographic screening for breast cancer. Breast Cancer Res. 2005;7(6):258–265.
7. Zackrisson S, Andersson I, Janzon L, *et al.* Rate of overdiagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. BMJ. 2006;332(7543):689–692.
8. Biesheuvel C, Barratt A, Howard K, *et al.* Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. Lancet Oncol. 2007;8(12):1129–1138.
9. Morrell S, Barratt A, Irwig L, *et al.* Estimates of overdiagnosis of invasive breast cancer associated with screening mammography. Cancer Causes Control. 2010;21(2):275–282.
10. de Koning HJ, Draisma G, Fracheboud J, *et al.* Overdiagnosis and overtreatment of breast cancer: microsimulation modelling estimates based on observed screen and clinical data. Breast Cancer Res. 2006;8(1):202. (doi:10.1186/bcr1369).
11. de Koning HJ, Boer R, Warmerdam PG, *et al.* Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. J Natl Cancer Inst. 1995;87(16):1217–1223.
12. Groenewoud JH, Pijnappel RM, van den Akker-Van Marle ME, *et al.* Cost-effectiveness of stereotactic large-core needle biopsy for nonpalpable breast lesions compared to open-breast biopsy. Br J Cancer. 2004;90(2):383–392.
13. Vereniging Integrale Kankercentra VIKC. Incidence data. Utrecht, the Netherlands: VIKC; 2010. (http://www.ikcnet.nl/). (Accessed April 6, 2010).
14. National Evaluation Team for Breast Cancer Screening (NETB). National Evaluation of Breast Cancer Screening in The Netherlands 1990–2007. Twelfth Evaluation Report. Rotterdam, Nijmegen, the Netherlands: Erasmus MC, Radboud University Nijmegen Medical Centre; 2009.
15. van den Akker-van Marle ME, Reep-van den Bergh CM, Boer R, *et al.* Breast cancer screening in Navarra: interpretation of a high detection rate at the first screening round and a low rate at the second round. Int J Cancer. 1997;73(4):464–469.
16. Rijnsburger AJ, van Oortmarssen GJ, Boer R, *et al.* Mammography benefit in the Canadian National Breast Screening Study-2: a model evaluation. Int J Cancer. 2004;110(5):756–762.
17. de Gelder R, Bulliard JL, de Wolf C, *et al.* Cost-effectiveness of opportunistic *versus* organised mammography screening in Switzerland. Eur J Cancer. 2009;45(1):127–138.

18. Beemsterboer PM, de Koning HJ, Looman CW, *et al.* Mammography requests in general practice during the introduction of nationwide breast cancer screening, 1988–1995. Eur J Cancer. 1999;35(3):450–454.
19. Moss S. Overdiagnosis and overtreatment of breast cancer: overdiagnosis in randomised controlled trials of breast cancer screening. Breast Cancer Res. 2005;7(5):230–234.
20. Puliti D, Paci E. The other side of technology: risk of overdiagnosis of breast cancer with mammography screening. Future Oncol. 2009;5(4):481–491.
21. Paci E, Miccinesi G, Puliti D, *et al.* Estimate of overdiagnosis of breast cancer due to mammography after adjustment for lead time. A service screening study in Italy. Breast Cancer Res. 2006;8(6):R68. (doi:10.1186/bcr1625).
22. Jørgensen KJ, Zahl PH, Gøtzsche PC. Overdiagnosis in organised mammography screening in Denmark. A comparative study. BMC Womens Health. 2009;9:36. (doi:10.1186/1472-6874-9-36).
23. Paci E, Warwick J, Falini P, *et al.* Overdiagnosis in screening: is the increase in breast cancer incidence rates a cause for concern? J Med Screen. 2004;11(1):23–27.
24. Olsen AH, Agbaje OF, Myles JP, *et al.* Overdiagnosis, sojourn time, and sensitivity in the Copenhagen mammography screening program. Breast J. 2006;12(4):338–342.
25. 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(1):25–30.
26. Welch HG, Black WC. Overdiagnosis in cancer. J Natl Cancer Inst. 2010;102(9):605–613.
27. Zahl PH, Strand BH, Maehlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study. BMJ. 2004;328(7445):921–924.
28. Jonsson H, Johansson R, Lenner P. Increased incidence of invasive breast cancer after the introduction of service screening with mammography in Sweden. Int J Cancer. 2005;117(5):842–847.
29. Martinez-Alonso M, Vilaprinyo E, Marcos-Gragera R, *et al.* Breast cancer incidence and overdiagnosis in Catalonia (Spain). Breast Cancer Res. 2010;12(4):R58. (doi:10.1186/bcr2620).
30. Duffy SW, Lynge E, Jonsson H, *et al.* Complexities in the estimation of overdiagnosis in breast cancer screening. Br J Cancer. 2008;99(7):1176–1178.
31. Kerlikowske K. Epidemiology of ductal carcinoma *in situ.* J Natl Cancer Inst Monogr. 2010;2010(41):139–141.
32. 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(3):117–125.
33. Feig SA. Ductal carcinoma *in situ.* Implications for screening mammography. Radiol Clin North Am. 2000;38(4):653–668, vii.
34. Cronin KA, Feuer EJ, Clarke LD, *et al.* Impact of adjuvant therapy and mammography on U.S. mortality from 1975 to 2000: comparison of mortality results from the CISNET breast cancer base case analysis. J Natl Cancer Inst Monogr. 2006;36:112–121.
35. Louwman WJ, Voogd AC, van Dijck JA, *et al.* On the rising trends of incidence and prognosis for breast cancer patients diagnosed 1975–2004: a long-term population-based study in southeastern Netherlands. Cancer Causes Control. 2008;19(1):97–106.0
36. Otten JD, Broeders MJ, Fracheboud J, *et al.* Impressive timerelated influence of the Dutch screening programme on breast cancer incidence and mortality, 1975–2006. Int J Cancer. 2008;123(8):1929–1934.

Breast cancer screening: Evidence for false reassurance?

Rianne de Gelder, Elisabeth van As, Madeleine M.A Tilanus-Linthorst,
Carina C.M. Bartels, Rob Boer, Gerrit Draisma, Harry J. de Koning

Int J Cancer. 2008 Aug 1;123(3):680–6.

ABSTRACT

Tumour stage distribution at repeated mammography screening is, unexpectedly, often not more favourable than stage distribution at first screenings. False reassurance, i.e., delayed symptom presentation due to having participated in earlier screening rounds, might be associated with this, and unfavourably affect prognosis. To assess the role of false reassurance in mammography screening, a consecutive group of 155 breast cancer patients visiting a breast clinic in Rotterdam (The Netherlands) completed a questionnaire on screening history and self-observed breast abnormalities. The length of time between the initial discovery of breast abnormalities and first consultation of a general practitioner ('symptom–GP period') was compared between patients with ('screening group') and without a previous screening history ('control group'), using Kaplan–Meier survival curves and log-rank testing. Of the 155 patients, 84 (54%) had participated in the Dutch screening programme at least once before tumour detection; 32 (38%) of whom had noticed symptoms. They did not significantly differ from control patients (n = 42) in symptom–GP period (symptom–GP period ≥30 days: 31.2% in the symptomatic screened group, 31.0% in the control group; p = 0.9). Only 2 out of 53 patients (3.8%) with screen-detected cancer had noticed symptoms prior to screening, reporting symptom–GP periods of 2.5 and 4 years. The median period between the first GP- and breast clinic visit was 7.0 days (95% C.I. 5.9–8.1) in symptomatic screened patients and 6.0 days (95% C.I. 4.0–8.0) in control patients. Our results show that false reassurance played, at most, only a minor role in breast cancer screening.

4.1 INTRODUCTION

Breast cancer screening programmes are effective in reducing mortality from breast cancer.[1,2] By detecting malignancies earlier, relatively smaller tumours are diagnosed and reductions in late stage disease are expected to be found. Consequently, at repeated screening, tumour stage distribution will be more favourable than at prevalent screening. In practice, stage distribution is not, or only slightly, more favourable.[3–6] 'False reassurance' is one of the explanations put forward for this lack of improvement at repeated screening relative to first screening rounds.[6] False reassurance is defined as diagnostic delay due to having participated in screening. It occurs after a negative screen result, when a patient or doctor, perceiving the risk of developing cancer to be small, is consequently less alert to present symptoms or the need for further evaluation. Prolonged patient delays, in literature often defined as intervals of more than 12 weeks, have been associated with increased tumour sizes,[7,8] more positive lymph nodes[7] and with decreased long-term survival.[7,9] False reassurance may therefore influence tumour stage and prognosis. Although various authors have suggested the possibility of false reassurance as a negative consequence of screening,[10–12] as far as we know, no study has actually examined the phenomenon of delayed symptom presentation after mammography screening.

In our study, the impact of breast cancer screening on the time of presentation of breast cancer symptoms and the moment of breast assessment was examined. The length of the period between the first symptom(s), first medical consultation and breast assessment is compared between a group of breast cancer patients who were regularly screened and a group of patients who were not. Predictive factors that could underlie a prolonged delay are further investigated.

4.2 METHODS

Study participants were recruited from the breast clinics at hospitals in Rotterdam (The Netherlands). Recruitment initially took place at the Erasmus MC–Daniel den Hoed Cancer Clinic, but was later extended to the breast clinics at 4 teaching hospitals in Rotterdam. In Rotterdam, all women with breast complaints suspicious for breast cancer are referred to a breast clinic. The main function of the breast clinics is to evaluate breast abnormalities and determine the course of treatment. During 20 months at the Erasmus MC-Daniel den Hoed Cancer Clinic, and during 13 months at the teaching hospitals, all consecutive women with breast abnormalities highly suspicious for breast malignancy who visited the breast clinic for the first time were invited to participate in the false reassurance study. Within this time frame, a study population large enough to assess the role of false

reassurance in breast cancer screening with reasonable certainty was expected to be included. However, because false reassurance was not examined before, it was not known how many screened breast cancer patients would be symptomatic before diagnosis, and how long these women would delay to present symptoms. We were therefore unable to make well-founded assumptions about the sample size that would be needed to obtain sufficient power.

After their first visit to the breast clinic, patients were given written and oral information about the study. Patients who had cognitive impairments, who did not speak Dutch, or who had had breast cancer in the past were excluded from participation. Eligible patients, who gave their informed consent to participate in the study, were asked to complete a questionnaire at home, after the first visit to the breast clinic. Based on the answers to questionnaire items regarding participation in the nation-wide breast cancer screening programme in the past 5 years, women were either included in a screening group or a control group of women who had not been screened (Figure 4.1). The 'screening group' consisted of women who were referred to the breast clinic after a recent positive mammography and who had participated in at least 1 screening round prior to that ('screen detected group'), and women who discovered abnormalities in the interval between 2 screening rounds ('interval group'). Patients regularly undergoing mammography for surveillance purposes outside the screening programme were not included in the analysis. Because the main objective of our study was to assess the length of the period between noticing the first symptom(s) and seeking medical advice, only those patients who themselves detected breast abnormalities were further analysed. In the 'screen detected group', this meant that prior to the detection of an abnormality by mammography, the patient had already noticed (a) symptom (-s). The screen detected group and the interval group together composed the 'symptomatic screened group'. The 'control group' included women who either fell outside the age limits for a screening invitation (<50 or >75 years old) or had never attended screening. Women who were, at a subsequent visit to the breast clinic, not diagnosed with breast cancer were retrospectively excluded from the analysis. Only those patients whose symptoms were most likely related to the breast cancer diagnosis were further analysed.

The questionnaire focused on breast abnormalities and their discovery, seeking help, diagnosis and start of treatment. We asked participants whether they had observed 1 or more of the following breast abnormalities, changes or symptoms during the past 12 months: a lump in the breast, nipple discharge, scaling/eczema/retraction of the nipple, dimple in the breast/skin retraction, pain in the breast, other abnormalities or no symptoms. The date on which the (first) abnormality was detected, the subsequent visit to the general practitioner (GP) and referral and first visit to the breast clinic were recorded. The

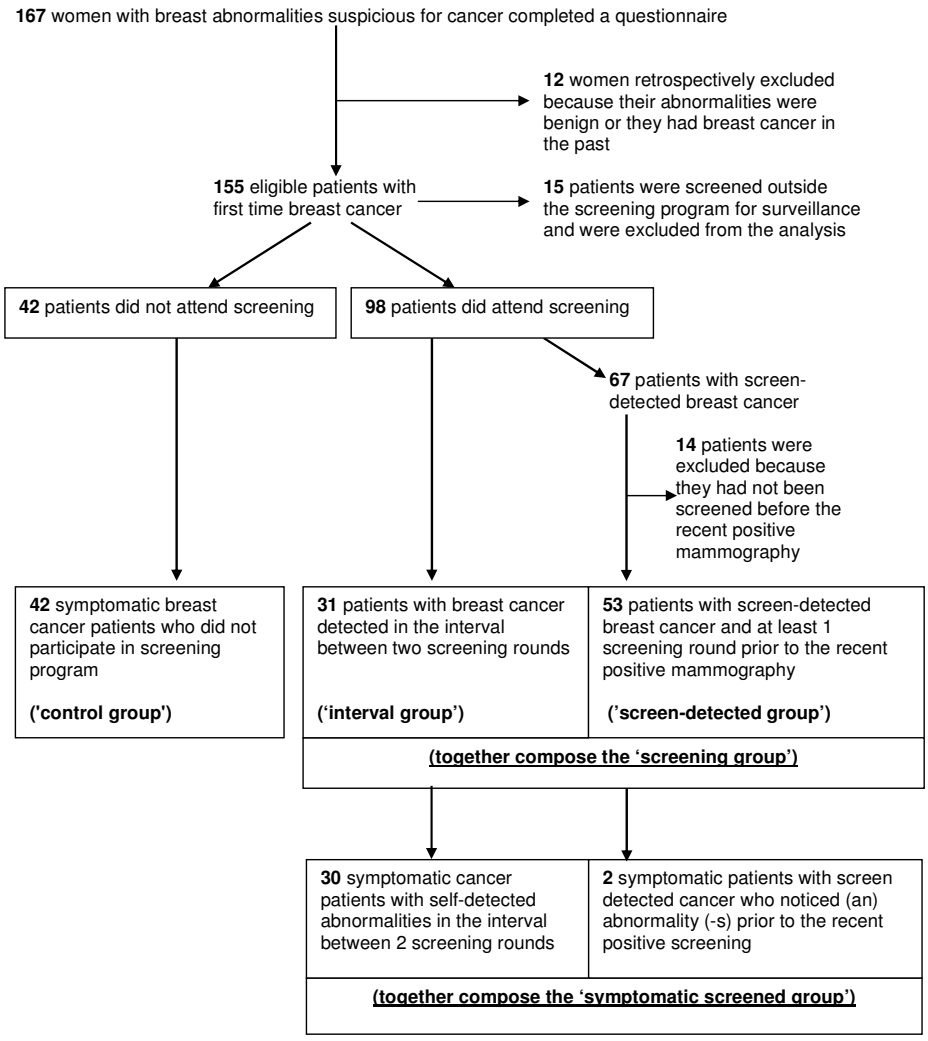


Figure 4.1 Inclusion of study participants.

nature of present and past breast-related symptoms was enquired, as well as general health and preventive behaviour patterns and demographic characteristics. The second part of the questionnaire covered a number of psychological aspects involved, including the attitude towards screening and the reasons women decide, or decide not to consult a GP. Knowledge and beliefs about breast cancer and screening were also measured.

Women were further asked to participate in a telephone interview that was designed to validate the data mentioned in the questionnaire. The interview lasted half an hour and was carried out by a qualified researcher (EvA). The questionnaire data were verified on

the basis of hospital records for women who allowed 30 days or more to elapse between detection of a breast abnormality and visiting the doctor, women with a relatively long period between the GP- and breast clinic consult, women with missing questionnaire data and women with screen-detected cancer. Before the study was initiated, a small pilot survey among breast cancer patients at the Erasmus MC-Daniel den Hoed was conducted. The Medical Ethical Committee of the Erasmus Medical Centre approved of the study protocol (MEC 185.919/1999/193).

Statistical analysis

To compare demographic characteristics and the frequency of breast symptoms between the symptomatic screened and the control group, the chi-square and Mann–Whitney U test were used. The percentage of patients who delay to present symptoms for ≥30 days and ≥90 days was compared between both groups, using a chi-square test. Within the control group, differences in the time of presenting symptoms between women who were not screened because they fell outside the age limits for screening and women who did not attend screening were assessed, using a Student's t-test. With a Kaplan–Meier survival model and 2-sided log-rank test, the median time between the first symptom(s) and the first visit to the GP ('symptom–GP period') was compared between the symptomatic screened and control patients. Because the end point in this analysis was the moment of seeking medical care, detection by screening (in the 'screen detected group') was considered a censoring event. Furthermore, the hazard ratio and confidence interval of the symptom–GP period between control and screened patients was calculated, and the influence of potential covariates on the symptom–GP period was estimated using a Cox regression model. The length of time between the first GP and first breast clinic visit ('GP–breast clinic period') was also calculated, using a Kaplan–Meier survival model and log rank testing. A p-value of 0.05 was considered to be significant. Additionally, a 2-sided non-parametric bootstrap procedure was used to calculate 95% confidence intervals around the median symptom–GP and GP–breast clinic period in the 2 groups. A total of 1,000 resamples were therefore randomly drawn with replacement from the original dataset, and the 2.5 and 97.5 percentiles were used to obtain a bootstrap confidence interval.

Results were analysed using SPSS statistical software, version 11.0 and S-Plus, version 6.

4.3 RESULTS

Of the 210 patients who were invited to participate in the study, 167 women (79.5%) gave their informed consent and completed the questionnaire. The most common reason

cited for not participating in the study was 'having too much on my mind'. Forty-seven patient record reviews and 113 patient interviews were additionally conducted. Twelve patients (7.2%) were retrospectively excluded because they had benign abnormalities (n = 10) or had had cancer in the past (n = 2). A total of 155 eligible breast cancer patients were thus further analysed.

Table 4.1 Demographic characteristics of participants

Patient group	Participated in screening program ('screening group') (n = 84)	Symptomatic screened group (n = 32)	Not attending screening program ('control group') (n = 42)	p-Value of difference between control group and screening group	p-Value of difference between control group and symptomatic screened group
Age in years (M, SD and range) at time of completing the questionnaire	61.0 ± 6.8 (51–75)	59·0 ± 7.5 (51–57)	49.3 ± 14.4 (30–86)	<0.001[a]	<0.001[a]
Attended breast clinic (n, %)					
Erasmus MC/ Daniel den Hoed	50 (59.5)	21 (65.6)	20 (47.6)	0·2[b]	0.1[b]
Teaching hospital	34 (40.5)	11 (34.4)	22 (52.4)		
Marital status (n, %)					
Married	54 (64.3)	21 (65.6)	26 (61.9)	0.2[b]	0.4[b]
Not married	6 (7.1)	2 (6.3)	8 (19.0)		
Divorced	10 (11.9)	5 (15.6)	4 (9.5)		
Widowed	14 (16.7)	4 (12.5)	4 (9.5)		
Education (n, %)					
Primary school	16 (19.0)	6 (18.8)	8 (19.0)	<0.05[b]	0.06[b]
Lower vocational school	28 (33.3)	12 (37.5)	4 (9.5)		
Intermediate secondary school	25 (29.8)	8 (25.0)	12 (28.6)		
Intermediate vocational school	7 (8.3)	2 (6.3)	8 (19.0)		
Higher vocational school	7 (8.3)	4 (12.5)	9 (21.4)		
University	1 (1.2)	0 (0)	1 (2.4)		
Insurance (n, %)					
Public insurance	50 (59.5)	16 (50.0)	28 (66.7)	0.3[b]	0.06[b]
Private insurance	30 (35.7)	15 (46.9)	10 (23.8)		
Civil service insurance	3 (3.6)	0 (0)	3 (7.1)		
No insurance	1 (1.2)	1 (3.1)	0 (0)		
Unknown	0 (0)	0 (0)	1 (2.4)		

[a] Mann-Whitney U test; [b] Chi-square test

Abbreviation: n = Total number included

Women who were screened for surveillance (n = 15) and women with screen-detected cancer who participated in the nation-wide screening programme for the first time (n = 14) were not included in the analysis. The total study population thus consisted of 126 patients; 42 of them not attending screening ('control group') and 84 who did participate in screening ('screening group') (Figure 4.1). The women in the screening arm were significantly older (Mean age in screening group was 61.0 *vs.* 49.3 in the control group, Table 4.1) and less well educated than the control patients. The 2 groups did not differ with regard to marital status and insurance (Table 4.1).

Out of the 84 patients in the screening group, 53 were detected by screening ('screen detected group') and 31 were diagnosed in the interval between 2 screening examinations, after their previous screening test (<5 years ago) was negative ('interval group'). Of those women with screen-detected cancer, 51 (96%) reported not to have noticed any breast cancer abnormalities prior to the positive mammogram. These patients were not further analysed. Two women with screen-detected cancer (3.8%) did have previous symptoms. Of the 31 patients who were diagnosed in the interval between 2 screening rounds, 1 was not further analysed because her symptoms were not self-detected, but initially discovered by a GP (Figure 4.1). The remaining 30 patients and the 2 symptomatic patients with screen-detected cancer together composed the 'symptomatic screened

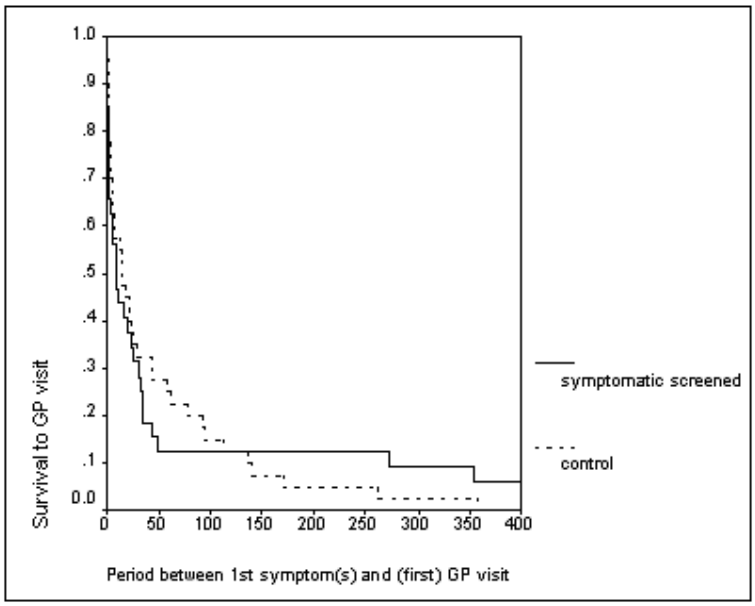


Figure 4.2 Period between discovering 1st symptoms and (1st) GP visit (days). The table was cut off at 400 days; 1 screened patient had a period of 890 days and 1 screened patient had a period of 1,453 days between the first symptoms and the first GP consult.

Table 4.2 Symptoms and the time between discovery and visiting a doctor

Patient group	Symptomatic screened group (n = 32)	Control group (n = 42)	p-Value
Time in days between discovery of the (first) symptom and the first GP visit			
(Median, 95% C.I.)	7.0 (0.0–15.3)	13.5 (7.3–19.7)	0.9[a]
(≥30 days: n, %)	10 (31.2)	13 (31.0)	0.9[b]
(≥90 days: n, %)	4 (12.5)	8 (19.0)	0.4[b]
Time in days between first GP visit and first breast clinic visit			
(Median, 95% C.I.)	7.0 (5.9–8.1)	6·0 (4.0–8.0)	0.9[a]
(≥10 days: n, %)	7 (21.9)	11 (26.2)	0.6[b]
Kind of symptoms discovered in the previous year (n, %)			
Breast lump	26 (81.3)	34 (81.0)	1.0[b]
Breast pain	6 (18.8)	5 (12.2)	0.4[b]
Dimple in breast /skin retraction	2 (6.3)	4 (9.8)	0.6[b]
Nipple discharge	0 (-)	0 (-)	0.7[b]
Scaling, eczema or retraction of nipple	3 (9.4)	1 (2.4)	0.2[b]
Other symptoms	3 (9.4)	6 (14.6)	0.5[b]
Time in days between discovery of the symptom and first visit to breast clinic for that particular symptom (Median, SD)			
Breast lump	11.0 (14.7)	23.3 (69.0)	0.05[c]
Breast pain	37.0 (584.2)	56.0 (107.7)	-
Dimple in breast /skin retraction	181.0 (256.0)	12.0 (9.6)	-
Nipple discharge	-	-	-
Scaling, eczema or retraction of nipple	59.0 (486.7)	178.0 (-)	-
Other symptoms	44.0 (144.1)	16.0 (106.4)	-

[a] Log-rank test; [b] Chi-square test; [c] Student's t-test.

Abbreviation: n = Total number included

group'. The symptomatic screened women differed significantly from the control patients with regard to age (The mean age of symptomatic screened women was 59.0, Table 4.1).

The median symptom–GP period in the symptomatic screened group, after censoring patients who were referred after a screening visit, was 7.0 days (Kaplan–Meier, 95% C.I. 0.0–15.3), whereas this period was 13.5 days (Kaplan–Meier, 95% C.I. 7.3–19.7) in the control group (Figure 4.2, Table 4.2). The log-rank test demonstrated that the 2 groups did not differ significantly regarding the length of time between discovery of the (first) breast

Table 4.3 Comparability of study groups

Covariate	Hazard Ratio	95% Confidence Interval	p-Value
Control/ Symptomatic screened group	1.03	0.62 – 1.71	0.91
Age	1.00	0.98 – 1.02	0.75
Education	0.96	0.80 – 1.16	0.66

Table 4.4 Symptom–GP period and tumour size

		Tumour diameter	
		≤20 mm (n, %)	>20 mm (n, %)
Period between observing (a) first symptom(s) and (first) GP consultation[1]	<30 days	25 (58.1%)	18 (41.9%)
	≥30 days	7 (33.3%)	14 (66.7%)
Period between observing (a) first symptom(s) and (first) GP consultation[2]	<90 days	27 (50.9%)	26 (49.1%)
	≥90 days	5 (45.5%)	6 (54.5%)

[a]p = 0.06 (10 missing values; of 2 women the symptom–GP period was not known and of 8 women the tumour size was not available). [b]p = 0.74 (10 missing values; of 2 women the symptom–GP period was not known and of 8 women the tumour size was not available).

abnormality and the first visit to a GP (p = 0.9). The hazard ratio of the symptom–GP period between control and screened patients was 1.03 (95% C.I. 0.62–1.71). The differences in age and education were not associated with the period between the appearance of the first symptoms and visiting a GP (Table 4.3).

Two women in the symptomatic screened group had relatively long symptom–GP periods of 890 and 1,453 days. They presented with retraction of the nipple and breast pain. These were the 2 patients with screen-detected cancer who indicated having noticed abnormalities prior to undergoing screening. No abnormalities were found during the previous screening round 2 years earlier. The estimated tumour stage at diagnosis of these patients was T1C (n unknown) and T1CN1 (a tumour size of 13 and 16 mm).

In the control group, 34 patients did not participate in screening because they fell outside the age limits for breast cancer screening; 8 women were eligible for screening but did not attend. The groups did not significantly differ with regard to the median symptom–GP period (Median symptom–GP period of women outside the age limits: 13.5 days, median

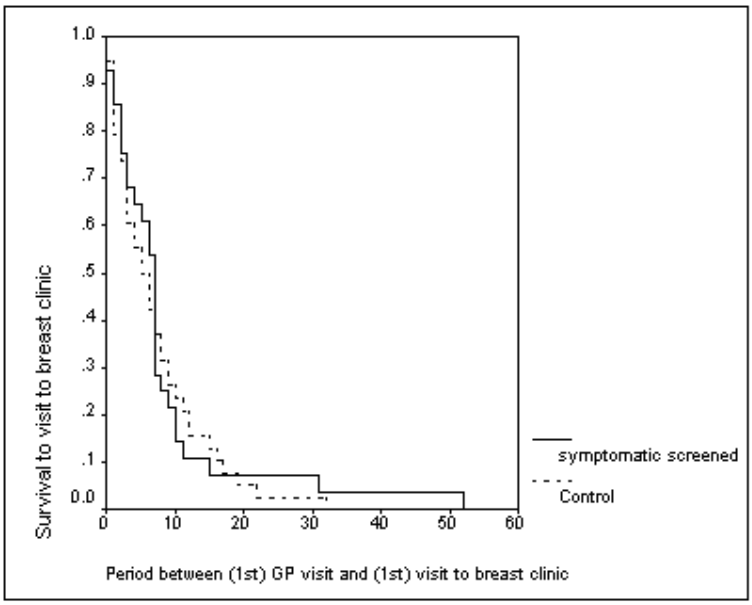


Figure 4.3 Period between (1st) GP visit and (1st) visit to breast clinic (days)

symptom–GP period of non-attendees: 23.0 days, p = 0.96). Neither did the group of non-attendees differ significantly from the symptomatic screened group (p = 0.6).

Of the symptomatic screened patients, 31.2% (n = 10) had a symptom–GP period of 30 or more days, and 12.5% (n = 4) had a symptom–GP period of 90 or more days. Of the control patients, 31.0% (n = 13) allowed 30 or more days to elapse between first symptoms and visiting a GP, and 19.0% (n = 8) had a symptom–GP period of 90 or more days. No significant differences were found between symptomatic and control patients in the percentages of patients with symptom–GP periods of ≥30 and ≥90 days (chi-square test: p = 0.9 and p = 0.4). Tumour size differed borderline significantly between women with symptom–GP periods <30 or ≥30 days (p = 0.06). No statistically significant tumour size differences were found among patients waiting ≥90 days from the time of finding a first symptom to the moment of consulting a GP, as compared to women with a symptom–GP period of <90 days (p = 0.4) (Table 4.4).

The length of the period between the first GP and first breast clinic visit did not differ significantly between both study arms (Figure 4.3, log-rank: p = 0.9). In the symptomatic screened group, the median GP–breast clinic period was 7.0 days (Kaplan–Meier, 95% C.I. 5.9–8.1). In the control group, this was 6.0 days (Kaplan–Meier, 95% C.I. 4.0–8.0). The percentage of women who had a GP–breast clinic period of 10 days or more was 21.9% in the symptomatic screened group (n = 7) and 26.2% in the control group (n = 11) (Table

Table 4.5 Reasons to visit a GP and symptom-GP period

	Frequency (n, %)	Mean time between discovery of the symptom and first GP visit in days (Median, SD)
Considerations not to visit a GP (initially)		
Symptoms will disappear spontaneously	13 (17.6)[a]	42.0 (75.2)
Having reservations about visiting doctors	7 (9.5)[a]	42.0 (124.5)
Nothing was wrong at the last screening visit	6 (18.8)[b]	34.5 (11.3)
Expecting an upcoming screening invitation	2 (2.7)[a]	19.0 (5.7)
Afraid of bad news	5 (6.8)[a]	35.0 (133.5)
Lack of time	2 (2.7)[a]	186.8 (107.1)
Other reasons	12 (16.2)[a]	42.0 (48.7)
Considerations to (eventually) visit a GP		
Reassurance	21 (28.4)[a]	15.0 (63.3)
Worrying	35 (47.3)[a]	5.0 (66.7)
Wanting to know if it is serious or not	32 (43.2)[a]	15.5 (90.7)
Visit to a GP for something else	5 (6.8)[a]	42.0 (182.9)
Following the advice of others	5 (6.8)[a]	29.0 (150.9)
After media information	4 (5.4)[a]	22.0 (125.0)
Other considerations	8 (10.8)[a]	36.0 (99.0)

[a] Percentages were based on the study group of 74 women with self-observed breast abnormalities. [b] Percentage was based on the 32 women in the symptomatic screened group

4.2). The bootstrap 95% confidence intervals around the median symptom–GP and GP–breast clinic period were comparable with the intervals calculated in the Kaplan–Meier survival analysis.

The most commonly cited reason for patients not to immediately see a physician was: 'my complaints will disappear spontaneously' (17.6%) (Table 4.5). This resulted in a median symptom–GP period of 42.0 days. For 6 (18.8%) of the symptomatic women who had participated in screening, previous negative screening results were cited as a reason to postpone a visit to a GP, with a median resulting symptom–GP period of 34.5 days (range 15–48 days). Still, a majority of women (78.1%) indicated that a prior negative screen did not influence the decision to consult a GP for the current symptom(s) (data not shown). In addition, only 2 out of the 74 participants indicated that an upcoming screening invitation resulted in a delay of the GP-consult, by 15.0 and 23.0 days, respectively. The main reasons for ultimately deciding to see a GP were 'worrying' and 'wanting to know whether it was something serious, or not'. For 47.3 and 43.2 % of all patients, these

motives played a role in the decision to consult a GP. Five women eventually presented their symptoms while consulting the physician for other problems. This resulted in the longest symptom–GP period, with a median of 42 days (Table 4.5).

When the 74 patients with self-reported breast abnormalities were asked how they would estimate the general influence of screening programmes on the moment of visiting a doctor and the awareness of breast abnormalities, opinions varied. Almost half (46.0%) of the participants disagreed with the statement that women with negative mammography results tend to postpone a visit to a GP after the discovery of a change in the breast. On the other hand, 35.1% of all respondents indicated that they considered the notion of false reassurance to be more or less plausible. One-third (33.8%) of the women believed that the screening programme contributed to a certain extent to decreasing the attention paid in general by women to possible changes in the breast. In contrast, almost 50% of the women (47.3%) disagreed emphatically and less emphatically with this statement. The majority (86.5%) of women agreed that breast self-examination was useful, and should continue to be performed, along with the screening programme. Screened and non-screened women did not differ in their perception of the influence of screening programmes.

Knowledge about breast cancer and screening was reasonable. For instance, most participants (86.5%) knew that only women aged 50–75 years are invited for screening, and that mammography can detect changes in the breast that are not yet observable (75.7%). A majority of women (87.8%) also knew that breast cancer could occur without symptoms or without feeling ill, and that lumps or changes in the breast do not necessarily indicate malignancy (75.7 and 78.4%). Half of the respondents with self-detected abnormalities knew that the lifetime risk of developing breast cancer is 1 out of 10. Fifty per cent of the patients (58.1%) were not familiar with the fact that approximately three-fourth of all Dutch breast cancer patients are older than 50, but 63.5% knew that breast cancer is the most common cancer among women in this age category in the Netherlands. Approximately 20% of the respondents assumed that the nation-wide screening programme takes place once every 5 years, perhaps confusing breast cancer screening with the cervical cancer-screening programme.

4.4 DISCUSSION

A majority of the women in our study indicated that previous or upcoming screening visits did not affect their decision to present current symptoms. The analysis showed for the first time that the period between initially observing abnormalities, the first GP visit and the first breast clinic consult was similar between screened and non-screened women. Additionally, only 2 out of 53 patients (3.8%) with screen-detected cancer reported having noticed a breast abnormality prior to the recent positive screening. It is possible that their long symptom–GP period of 890 and 1,453 days is attributable to false reassurance, but other causes cannot be ruled out. If their delay is indeed attributable to false reassurance, these results indicate that false reassurance plays, at the most, only a minor role in breast cancer screening, and probably does not affect the tumour stage distribution found at repeated screening rounds. Nevertheless, a relatively high percentage (18.8%) of women in our study population indicated that the last negative screening result constituted a consideration to postpone a visit to the GP. Their median symptom–GP period was 34.5 days (range, 15–48 days), indicating that false reassurance generally does not lead to long delays in presenting symptoms. However, 31% of all symptomatic patients waited more than 30 days to consult a doctor about their symptoms, and 12.5% of the symptomatic screened and 19.0% of the control patients had a period of more than 90 days between experiencing the first symptoms and visiting a GP.

A delay in presenting symptoms might partially be attributed to wrong assumptions about the risk of breast cancer and the benefits and disadvantages of screening. Our survey indicated that the knowledge about screening was generally realistic, probably because all Dutch women who are invited for screening receive detailed information about breast cancer and screening. Other studies, on the contrary, have shown that, in other countries, knowledge about screening may be limited. For example, a high number of women assume that mammography prevents or reduces the risk of contracting breast cancer.[11,13] Others were not familiar with the fact screening may have false negative results.[11] Some screened women may therefore underestimate the risk of developing breast cancer or have fewer fears about cancer, which has been associated with delay.[14–16] A lack of knowledge about breast cancer symptoms may also cause delay.[11,17] Most women are familiar with breast lumps as a potential sign of breast cancer, but knowledge about other symptoms may be limited.[15] The long delay after non-lump symptoms, which was illustrated in our analysis and has been demonstrated in previous studies,[8,14,16,18,19] could have resulted from this. Providing women with valid information about breast cancer and mammography screening, including the risk of false-negative screening outcomes and the possibility that symptoms develop in the period between 2 screening examinations, is needed to minimize the risk of delay and advanced disease.

Some factors could have influenced our results. Recall bias was a risk in our particular study design. In some cases, first symptoms were observed long ago, which made accurate timing difficult. Moreover, screened patients may well differ from non-screened individuals in (not) attributing certain symptoms to cancer. For instance, after a negative screening result, a patient may forget about a symptom that was present before screening. However, after a positive screening, symptoms may be reported that would not have led the woman to consult her GP had she not been screened. Our approach, in which we interviewed the women about their symptoms, next to examining patient hospital records to check for the presence of breast symptoms, has led us to feel relatively confident that we have succeeded in obtaining fairly reliable results. Moreover, to avoid having the women feel compelled to give 'socially correct' answers, we downplayed the notion (in both the questionnaire and patient information) that a screening programme could possibly lead to delay. General terms were used when referring to symptoms and time periods. Another limitation of the analysis was the age and education difference: non-screened women were significantly younger and better educated than women who were screened. Although age[7,17,20] and education[14,21] have been associated with delay, there was no evidence that the age and education differences affected our results.

The backwards design of our study also constituted some limitation. The alternative, a follow-up study with registration of the symptoms, would be a complex and expensive method to assess the role of false reassurance in breast cancer screening. Such an approach, however, might be recommended in future research.

The study population of 155 consecutive patients, 74 of whom reporting symptoms, was considered to be of sufficient magnitude to detect significant differences between the 2 study groups. The hazard ratio of the symptom–GP period between control and screened patients of 1.03 showed that, at each point in time, there is a 3% difference between the 2 groups in the likeliness that they will consult a doctor for their symptoms. Only a small proportion of the screened women in our study postpone seeking medical consultation for breast symptoms for a long period. However, the 95% confidence interval of the hazard ratio of is 0.62–1.71, indicating uncertainty in the analysis. Nevertheless, the fact that only 2 out of the 53 patients with screen-detected breast cancer had symptoms before the positive mammogram indicates that the role of false reassurance in breast cancer screening is small. The extent to which false reassurance plays a role in breast cancer screening remains to be studied in greater depth. Questionnaires that are directed at assessing the presence of symptoms before a screening visit and detailed interviews about psychological responses to screening should provide more insight into the role of false reassurance in delay.

Conclusion

Patients who participated in breast cancer screening did not have a significantly longer period between noticing (a) first breast abnormality (-s) and the first time presenting this to a GP and breast clinic than patients who did not undergo screening. Our data show that false reassurance plays, at most, only a minor role in breast cancer screening.

Acknowledgements

The authors are grateful to Mrs. E. Boot, Dr. J.H. Driebeek-van Dam, Dr. W.S. Meyer, Dr. J.C.J. Wereldsma and Dr. M.A. Paul for their help in collecting the data. They also thank Mr. C.W. Looman for statistical advice and Mrs. N. van der Zouwe for her advice in designing the psychological questionnaire.

REFERENCES

1. 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. BMJ 2005;330:220.
2. Otto SJ, Fracheboud J, Looman CW, Broeders MJ, Boer R, Hendriks JH, Verbeek AL, de Koning HJ. National Evaluation Team for Breast Cancer S. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. Lancet 2003;361:1411–7.
3. Fracheboud J, de Koning HJ, Beemsterboer PM, Boer R, Hendriks JH, Verbeek AL, van Ineveld BM, de Bruyn AE, van der Maas PJ. Nation-wide breast cancer screening in The Netherlands: results of initial and subsequent screening 1990–1995. National Evaluation Team for Breast Cancer Screening. Int J Cancer 1998;75:694–8.
4. Paci E, Ponti A, Zappa M, Patriarca S, Falini P, Delmastro G, Bianchi S, Sapino A, Vezzosi V, Senore C, Crocetti E, Frigerio A, *et al.* Early diagnosis, not differential treatment, explains better survival in service screening. Eur J Cancer 2005;41:2728–34.
5. Hofvind S, Wang H, Thoresen S. Do the results of the process indicators in the Norwegian Breast Cancer Screening Program predict future mortality reduction from breast cancer? Acta Oncol 2004;43:467–73.
6. Boer R, de Koning H, van Oortmarssen G, Warmerdam P, van der Maas P. Stage distribution at first and repeat examinations in breast cancer screening. J Med Screen 1999;6:132–8.
7. Afzelius P, Zedeler K, Sommer H, Mouridsen HT, Blichert-Toft M. Patient's and doctor's delay in primary breast cancer. Prognostic implications. Acta Oncol 1994;33:345–51.
8. Burgess CC, Ramirez AJ, Richards MA, Love SB. Who and what influences delayed presentation in breast cancer? Br J Cancer 1998; 77:1343–8.
9. Richards MA, Smith P, Ramirez AJ, Fentiman IS, Rubens RD. The influence on survival of delay in the presentation and treatment of symptomatic breast cancer. Br J Cancer 1999;79: 858–64.
10. Drossaert CH, Boer H, Seydel ER. Does mammographic screening and a negative result affect attitudes towards future breast screening? J Med Screen 2001;8:204–12.
11. Webster P, Austoker J. Women's knowledge about breast cancer risk and their views of the purpose and implications of breast screening—a questionnaire survey. J Public Health (Oxf) 2006;28: 197–202.
12. Baines CJ. Are there downsides to mammography screening? Breast J 2005;11 (Suppl 1):S7–S10.
13. Domenighetti G, D'Avanzo B, Egger M, Berrino F, Perneger T, Mosconi P, Zwahlen M. Women's perception of the benefits of mammography screening: population-based survey in four countries. Int J Epidemiol 2003;32:816–21.
14. Friedman LC, Kalidas M, Elledge R, Dulay MF, Romero C, Chang J, Liscum KR. Medical and psychosocial predictors of delay in seeking medical consultation for breast symptoms in women in a public sector setting. J Behav Med 2006;29:327–34.
15. Grunfeld EA, Ramirez AJ, Hunter MS, Richards MA. Women's knowledge and beliefs regarding breast cancer. Br J Cancer 2002;86: 1373–8.
16. Burgess CC, Potts HW, Hamed H, Bish AM, Hunter MS, Richards MA, Ramirez AJ. Why do older women delay presentation with breast cancer symptoms? Psychooncology 2006;15: 962–8.

17. Facione NC, Miaskowski C, Dodd MJ, Paul SM. The self-reported likelihood of patient delay in breast cancer: new thoughts for early detection. Prev Med 2002;34:397–407.
18. MacArthur C, Smith A. Delay in breast cancer and the nature of presenting symptoms. Lancet 1981;1:601–3.
19. Feldman JG, Saunders M, Carter AC, Gardner B. The effects of patient delay and symptoms other than a lump on survival in breast cancer. Cancer 1983;51:1226–9.
20. Facione NC. Delay *versus* help seeking for breast cancer symptoms: a critical review of the literature on patient and provider delay. Soc Sci Med 1993;36:1521–34.
21. Montella M, CrispoA, D'Aiuto G, DeMarcoM, de Bellis G, Fabbrocini G, PizzorussoM, Tamburini M, Silvestra P.Determinant factors for diagnostic delay in operable breast cancer patients. Eur J Cancer Prev 2001;10:53–9.

Cost-effectiveness of opportunistic *versus* organised mammography screening in Switzerland

Rianne de Gelder, Jean-Luc Bulliard, Chris de Wolf, Jacques Fracheboud, Gerrit Draisma, Dorris Schopper, Harry J. de Koning

Eur J Cancer 2009 Jan;45(1):127–38. Epub 2008 Nov 27.

ABSTRACT

Background

Various centralised mammography screening programmes have shown to reduce breast cancer mortality at reasonable costs. However, mammography screening is not necessarily cost-effective in every situation. Opportunistic screening, the predominant screening modality in several European countries, may under certain circumstances be a cost-effective alternative. In this study, we compared the cost-effectiveness of both screening modalities in Switzerland.

Methods

Using micro-simulation modelling, we predicted the effects and costs of biennial mammography screening for 50–69 years old women between 1999 and 2020, in the Swiss female population aged 30–70 in 1999. A sensitivity analysis on the test sensitivity of opportunistic screening was performed.

Results

Organised mammography screening with an 80% participation rate yielded a breast cancer mortality reduction of 13%. Twenty years after the start of screening, the predicted annual breast cancer mortality was 25% lower than in a situation without screening. The 3% discounted cost-effectiveness ratio of organised mammography screening was €11,512 per life year gained. Opportunistic screening with a similar participation rate was comparably effective, but at twice the costs: €22,671–24,707 per life year gained. This was mainly related to the high costs of opportunistic mammography and frequent use of imaging diagnostics in combination with an opportunistic mammogram.

Conclusion

Although data on the performance of opportunistic screening are limited, both opportunistic and organised mammography screening seem effective in reducing breast cancer mortality in Switzerland. However, for opportunistic screening to become equally cost-effective as organised screening, costs and use of additional diagnostics should be reduced.

5.1 INTRODUCTION

Breast cancer mortality has decreased in several countries in the last decade. Mammography screening is one of the factors contributing to this decline.[1–3] Various organised mammography screening programmes have shown to be effective in reducing breast cancer mortality[4–6] at costs well below the WHO threshold[7] of cost-effectiveness.[8–10] However, mammography screening is not necessarily cost-effective in every situation. Cost-effectiveness depends on country-specific demographic and epidemiologic characteristics, breast cancer incidence, tumour stage distribution and breast cancer mortality before the initiation of screening, and the characteristics of screening, such as attendance, targeted screening ages and screening interval. The organisation of a health care system, and the costs of screening, diagnostics and treatment also determine whether mammography screening is cost-effective.

Under certain circumstances, 'opportunistic' mammography screening in asymptomatic women, the predominant form of screening in several European countries, may be a cost-effective alternative to programme-based mammography screening. The objective of the current study is to compare the cost-effectiveness of both screening modalities in Switzerland. Six French-speaking Swiss cantons have a biennial organised mammography screening programme (MSP), which coexists with opportunistic screening (OS). While OS is assumed to have started around the mid-eighties, programme screening in these areas started between 1999 and 2007, currently inviting approximately 25% of the 50–69 years old Swiss female population.[11] Other, mainly German-speaking, cantons only screen opportunistically, to a smaller or larger extent.

A previous analysis of the incremental cost-effectiveness of MSP relative to OS estimated that in Switzerland, MSP would yield a relevant reduction of breast cancer mortality at moderate additional costs.[12] In that study, however, cost-effectiveness was predicted using a single-cohort model, based on a conservative screening-associated breast cancer mortality reduction of 15% (randomised controlled trials showed reductions of 21–31%).[4–6,13,14] Cost savings related to a decreased use of palliative care were not included. In the present analysis, the effects and costs of MSP and OS were predicted at population level, using the internationally validated micro-simulation model 'MISCAN'.[15–18] Trial-based screening effects, specific Swiss demographic and epidemiologic data and MSP- and OS- specific screening characteristics were taken into account. To account for regional variations, five OS and MSP scenarios with varying screening participation rates were studied. Because data on the performance of OS are scarce, a sensitivity analysis on the test sensitivity of OS was performed.

5.2 METHODS

The 'MISCAN' model

With the micro-simulation screening analysis model 'MISCAN', the consequences of introducing a screening programme on individual life histories were assessed. In MISCAN, the natural history of breast cancer starts with a transition from 'no breast cancer' into pre-clinical screen-detectable breast cancer, in a certain percentage of the modelled population. Tumour development is modelled as a progression through the successive invasive disease stages T1a, T1b, T1c and T2+ (diameter ≤5 mm, 6–10 mm, 11–20 mm and >20 mm, respectively). Invasive cancer may or may not be preceded by pre-clinical DCIS. In each pre-clinical stage, a tumour may be clinically diagnosed or may grow into the next pre-clinical stage. If women are screened, the pre-clinical tumour may also be detected by screening (Figure 5.1). Screening ages, interval and attendance and the type of screening ('opportunistic' or 'organised'), as well as the sensitivity and specificity of mammography are defined in the model. MISCAN parameters are mean dwelling times, transition probabilities between pre-clinical stages and survival after clinical diagnosis or screen detection. These age- and stage-dependent parameters were estimated using breast cancer incidence and stage distribution of screened and unscreened populations. Transition probabilities, stage durations and survival after diagnosis were based on the outcomes of the Dutch nation-wide breast cancer screening programme[19] and the Dutch pilot studies in Nijmegen and Utrecht.[15,20–22] The survival after clinical diagnosis or screen detection was modelled using several international sources.[23–27] The improvement of prognosis after detection by screening (defined as 1 minus the ratio of the risk of dying of screen-detected cancer divided by the risk of dying when the cancer had been diagnosed in the absence of screening) was based on experience from the Swedish randomised trials.[14,24,28,29]

Model calibration

To adjust MISCAN for the Swiss situation, the model was calibrated with observed breast cancer data in the canton of Vaud, where a long-standing cancer registry including pre-screening years and the largest Swiss centrally organised screening programme operate.[30,31] For this purpose, we modelled the Vaud female population between 1974 and 2005. The Swiss life table of 1999 was used to model mortality from other causes than breast cancer. Breast cancer incidence before the introduction of screening was then modelled, using data from the Vaud Cancer Registry between 1974 and 1985. The mean duration of the pre-clinical tumour stages was estimated by fitting the model predictions to the stage distribution in the years before screening, and to the detection rates and interval cancer rates after the introduction of the screening programme. As data on the tumour stage distribution in pre-screening years were unavailable in Vaud, data from the

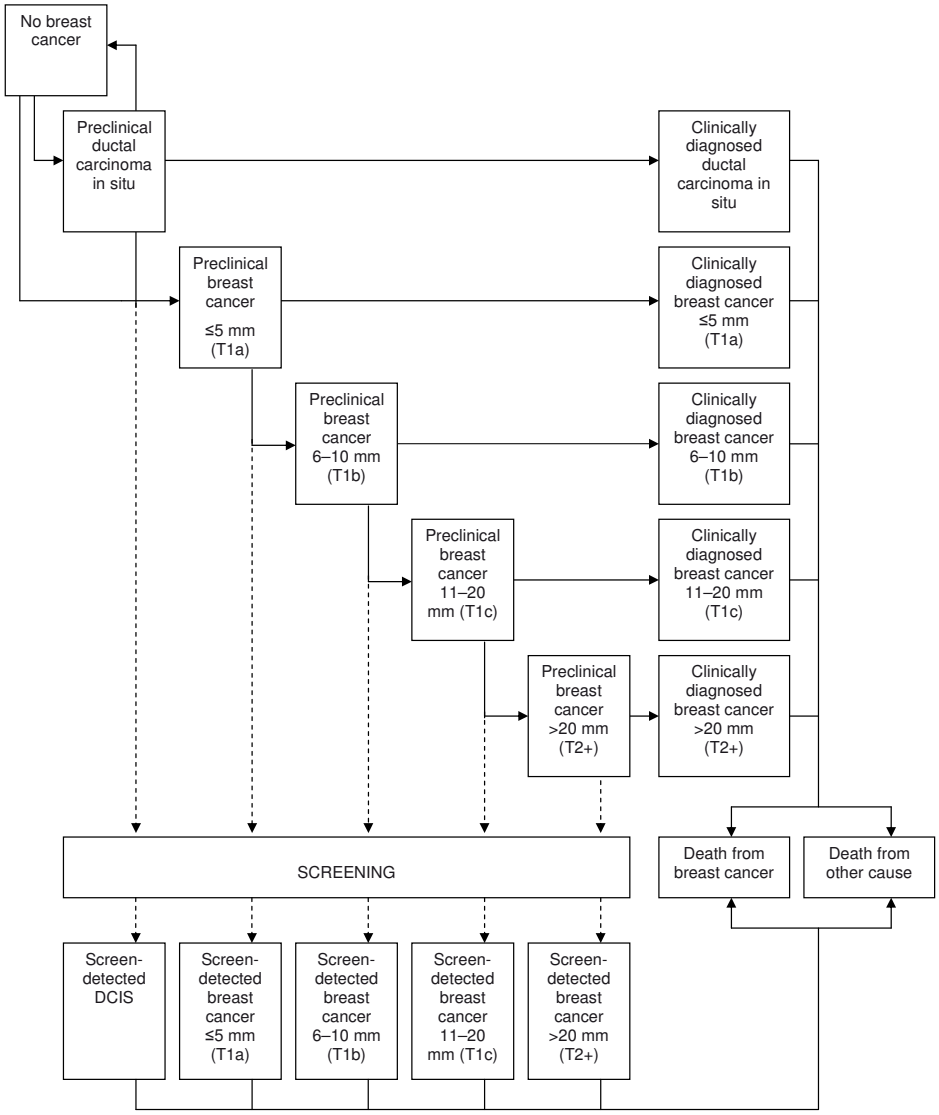


Figure 5.1 Transitions in the MISCAN mode

Geneva Cancer Registry between 1980 and 1984 were used. The breast cancer mortality was calibrated with data from the Vaud Cancer Registry between 1974 and 1985. The test sensitivity of screening was estimated by modelling screening characteristics of the Vaud screening programme[32] and calibrating the model to replicate the observed rates of interval cancers and screen-detected tumours by age, stage and screening round (first/subsequent). For the tumour stages DCIS, T1a, T1b, T1c and T2+, the test sensitivity was accordingly estimated to be 80%, 70%, 75%, 80% and 100%, respectively. The

breast cancer survival in the basic MISCAN model was slightly decreased to replicate the breast cancer mortality in Vaud. A χ^2 test applied to the deviance was used as a test of goodness-of-fit. The model parameters are shown in Appendix 5.1.

Screening assumptions

By comparing a screening scenario to a scenario without screening, cost-effectiveness was calculated. In the current analysis, we predicted the effects and costs of screening 50–69 years old Swiss women between 1999 and 2020, by MSP and/or OS. Computed effects were the number of breast cancer deaths prevented, life years gained and quality-adjusted life years gained per 1,000,000 women; the latter was calculated using utilities reported by de Haes *et al.*[33] Effects were calculated for the Swiss female population aged 30–70 in 1999, during the whole lifespan of this population. The breast cancer mortality reduction was also calculated for women aged 55–74, where the most relevant breast cancer mortality reduction was expected to be observed.[6] Additionally, the maximum reduction of the annual breast cancer mortality rate was computed. The modelled population had the same age distribution as observed in Vaud between 1999 and 2005.

Costs included all expenses on screening, (over)diagnosis, (over)treatment, follow-up and palliative care. Costs incurred during and after the screening period were included. Five hypothetical screening scenarios (Sc 1–5) were analysed:

Sc 1. 40% biennial OS (40% of the target population has an opportunistic mammogram every other year),
Sc 2. 80% biennial OS (80% of the target population has an opportunistic mammogram every other year),
Sc 3. 80% biennial MSP (all women in the target population are biennially invited to participate in an organised screening programme, and 80% participates),
Sc 4. 60% biennial MSP and 20% biennial OS (all women in the target population are biennially invited to participate in an organised screening programme, and 60% participates. Another 20% of the target population has an opportunistic mammogram every other year),
Sc 5. 40% biennial MSP and 40% annual OS (all women in the target population are biennially invited to participate in an organised screening programme, and 40% participates. Another 40% of the target population has an opportunistic mammogram every year).

We assumed that women who participated in a MSP were not screened opportunistically, and vice versa. Once a woman is screened, opportunistically or in a programme, she is assumed to have biennial (Sc 1–4 and Sc 5, MSP) or annual (Sc 5, OS) mammograms until

age 70. Women who are not screened were assumed never to have a mammogram. The 'intervals' for annual and biennial OS were assumed to be 0.75–1.25 years and 1.75–2.25 years, respectively. Scenarios 1, 4 and 5 represent mammography screening practice in parts of Switzerland; scenarios 2 and 3 were not regarded to be realistic representations of current or future mammography screening practice. However, scenario 3 approaches a screening situation such as observed in the North-European countries, where attendance rates around 80% have been observed.[34] Scenarios 2 and 3 enable a direct comparison of OS and MSP. As little is known about the performance of OS and the sensitivity of mammography may vary across Switzerland, a sensitivity analysis was performed, with three model variants of varying false-negative rates for OS compared to MSP:

A. An 'optimistic' variant, with a 25% lower false-negative rate of OS compared to MSP.
B. A 'baseline' variant, with similar false-negative rates of OS and MSP.
C. A 'pessimistic' variant, with a 25% higher false-negative rate of OS compared to MSP.

Costs and cost-effectiveness

The costs of diagnostics and treatment were calculated as the resource use multiplied with costs per unit. The number of diagnostic examinations done after a positive programme-based mammogram was estimated from referral rates in the Vaud screening programme.[32] The number of diagnostics used after an opportunistic mammogram was estimated with data from the largest Swiss health insurance company CSS, which provided individualised records on the use of health care services of women who had a mammogram between 2004 and 2006. The health insurance, however, did not register opportunistic mammography as a separate procedure, since it was reimbursed similarly as diagnostic mammography. Because the majority of these mammograms correspond to OS, all reimbursed mammograms were regarded as opportunistic mammograms. The number of clinical breast examinations performed for each mammogram was estimated by calculating the ratio between reimbursed clinical breast examinations and mammograms, using aggregated data on the use of breast cancer diagnostics from the health insurance organisation SantéSuisse. The number of diagnostic examinations needed to clinically diagnose 1 tumour outside the MSP and OS was, in absence of Swiss data, based on the results from the Dutch COBRA study.[35] The treatment use after the detection of a tumour in a MSP was derived from registrations in the Vaud screening programme.[32] Reliable data on the treatment of tumours detected by opportunistic screening and clinically diagnosed breast cancer were unavailable. Considering the Swiss health care structure, we assumed that a tumour detected by OS was treated similarly as a MSP-detected tumour. The probability of a specific treatment for a clinically diagnosed tumour was calculated as the same probability for a screen-detected tumour (as observed in Vaud) multiplied with the ratio of these probabilities between a screen-detected tumour and

a clinically diagnosed cancer, as observed in the Netherlands.[35] The probability that a specific treatment is used depends on tumour stage.

The costs of an opportunistic mammogram were calculated as the costs of a programme-based mammogram (€138[11]) multiplied by the ratio of CSS-reimbursed costs between a programme-based mammogram and an opportunistic mammogram. The costs of imaging and minimal invasive diagnostics were directly derived from CSS. Costs in Swiss Francs (CHF) were converted to Euro (€), using a currency exchange rate of 1.66 (June 2007). Since no reliable data were available on the costs of treatment in Switzerland, Dutch cost estimates were used.[36] A health care specific purchasing power parity (PPP) of 1.21 was applied to those costs, to account for the relatively higher health care costs in Switzerland[37] compared to the Netherlands. The costs of sentinel node procedures were estimated based on an analysis in the USA.[38] To account for time preference, both effects and costs were discounted at 3% per year,[39] from 1999 onwards.

5.3 RESULTS

Calibration

The calibrated model resembled the incidence and stage distribution of clinically diagnosed breast cancer, screen-detected breast cancer, interval cancers and the mortality due to breast cancer in Vaud reasonably well (Table 5.1). However, MISCAN predictions were optimistic with regard to the detection of T2+ tumours at subsequent screening examinations ($p<0.01$).

Effects

Effects are shown in Table 5.2. Biennial organised screening of 80% of the 50–69 years old female population ('Sc 3') was predicted to increase the total number of diagnosed breast cancers by 1.4%, compared to a situation without screening (80% MSP: 94,376 tumours; no screening: 93,036 tumours). Without screening, 42% of the tumours in the total simulated population were smaller than 20mm or non-invasive, *versus* 51% with screening (data not shown). For each screen-detected tumour, 222 screen-mammograms were performed, and for each prevented breast cancer death, 798 screen-mammograms were needed (data not shown). Opportunistic screening had similar outcomes.

Both MSP and OS were predicted to reduce breast cancer mortality. Biennial 80% MSP prevented 4921 breast cancer deaths per 1,000,000 women aged 30–70 in 1999, which is a breast cancer mortality reduction of 13% during the lifespan of this population, and a gain of 81,000 life years. Biennial 80% OS ('Sc 2', baseline variant) resulted in a breast

Table 5.1 Breast cancer incidence, mortality, detection rates and interval cancer rates as observed in the canton of Vaud, compared with MISCAN predictions

Parameter		Observed	MISCAN-predicted	p-Value
Clinical breast cancer incidence in pre-screening years, 1974–1985 (per 100,000 woman-years)	Age 0–100	104.5	106.4	0.3
Stage distribution of tumours in pre-screening years, 1980–1984 (%)				
DCIS	Age 0–100	3.9	4.1	0.9
T1a	Age 0–100	3.2	1.5	0.2
T1b	Age 0–100	10.4	6.5	0.2
T1c	Age 0–100	36.9	32.3	0.4
T2+	Age 0–100	45.6	55.6	0.2
Breast cancer mortality in pre-screening years, 1974–1985 (per 100,000 woman-years)	Age 0–100	39.1	37.8	0.3
Detection rates 1999–2005 (per 100,000 examinations)				
DCIS, first screening round	Age 50–69	130.4	155.3	0.2
DCIS, subsequent screening rounds	Age 50–69	85.8	85.8	1.0
T1a, first screening round	Age 50–69	78.3	66.4	0.4
T1a, subsequent screening rounds	Age 50–69	76.8	88.2	0.4
T1b, first screening round	Age 50–69	164.4	133.8	0.1
T1b, subsequent screening rounds	Age 50–69	135.5	131.2	0.8
T1c, first screening round	Age 50–69	255.7	283.8	0.3
T1c, subsequent screening rounds	Age 50–69	173.9	154.0	0.3
T2+, first screening round	Age 50–69	91.3	112.5	0.2
T2+, subsequent screening rounds	Age 50–69	74.5	35.4	<0.01
Interval cancers 1999–2004 (per 100,000 examinations)				
After a first screening round	Age 50–69	144.5	165.5	0.3
After a subsequent screening round	Age 50–69	135.1	102.2	0.1

Abbreviations: ductal carcinoma in situ (DCIS), diameter ≤5 mm (T1a), diameter 6–10 mm (T1b), diameter 11–20 mm (T1c), diameter >20 mm (T2+)

cancer mortality reduction of 13%, a prevention of 4876 breast cancer deaths and a gain of 80,400 life years. In the optimistic and pessimistic variant, the reductions were 14% and 13%, respectively (data not shown). Between the ages 55 and 74, the predicted breast cancer mortality reduction was 20% (80% MSP and 80% OS).We predicted that in 2018, 20 years after its start, MSP would reduce the breast cancer mortality rate by 25% (at population level) and 32% (age 55–74). For OS, these reductions were 23% (at population level) and 30% (age 55–74) (Table 5.2, Figure 5.2). The screening effects decreased proportionally with the fraction of women screened: 40% OS ('Sc 1') was predicted to reduce the breast cancer mortality over the whole lifespan of the population by 7%, and the annual breast cancer mortality rate in 2018 by 12% (at population level). The 40% biennial MSP/40% annual OS scenario ('Sc 5') was most effective, with a 15% breast cancer mortality reduction during the lifespan of the population, and a reduction of the breast cancer mortality rate in 2018 of 27% (at population level).

Costs

The unit costs for a programme-based mammogram were €138 and the costs for an opportunistic/diagnostic mammogram were €171. The costs of breast cancer diagnostics

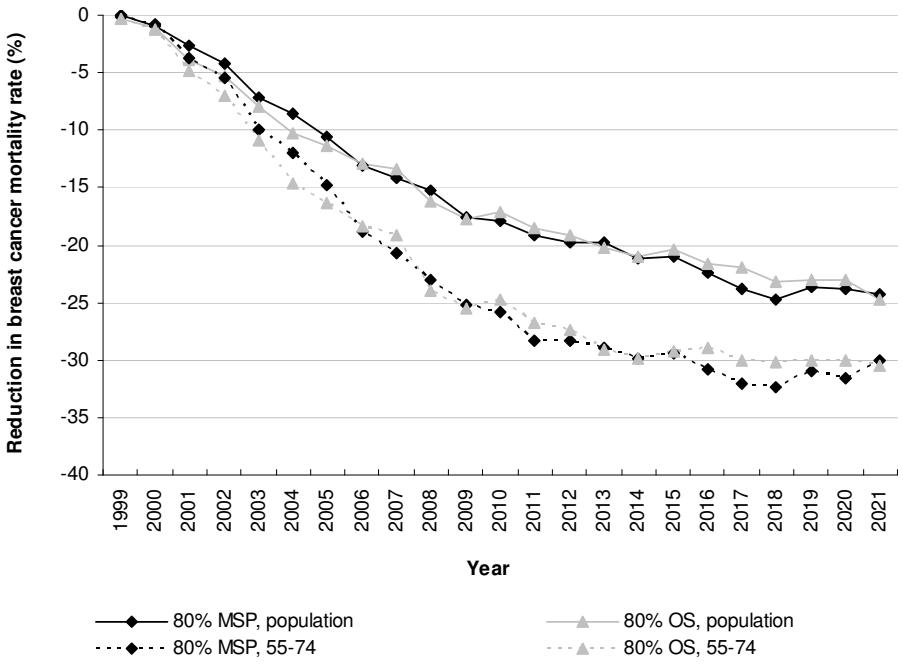


Figure 5.2 Annual reduction in the breast cancer mortality rate, relative to a situation without screening

Table 5.2 Predicted effects of opportunistic screening (OS) and mammography screening programme- (MSP) scenarios with varying participation rates, compared to a no-screening scenario, no discounting

	Sc 0	Sc 1	Sc 2	Sc 3	Sc 4	Sc 5
	No screening	40% OS	80% OS	80% MSP	60% MSP 20% OS	40% MSP 40% annual OS
Effects (no discounting)						
Breast cancers diagnosed, population level, during lifespan of population (N, %)	93,036	+670 (+0.7)	+1319 (+1.4)	+1340 (+1.4)	+1335 (+1.4)	+1430 (+1.5)
Breast cancer deaths, population level, during lifespan of population (N, %)	36,519	–2446 (–7)	–4876 (–13)	–4921 (–13)	–4909 (–13)	–5482 (–15)
Breast cancer deaths, age 55–74, during lifespan of population (N, %)	16.568	–1652 (–10)	–3298 (–20)	–3342 (–20)	–3331 (–20)	–3733 (–23)
Annual breast cancer mortality rate reduction in 2018, population level (%)	-	–12	–23	–25	–24	–27
Annual breast cancer mortality rate reduction in 2018, age 55–74 (%)	-	–15	–30	–32	–32	–35
Life years, population level, during lifespan of population (N)	36,390,700	+40,200	+80,400	+81,000	+80,825	+90,400
Quality adjusted life years, population level, during lifespan of cohort (N)	36,287,649	+38,305	+76,603	+77,176	+77,008	+86,174
Mammograms (N), population level	0	1,965,490	3,928,610	3,928,490	3,928,520	5,731,380

All effects and costs were rescaled to a population of 1,000,000 women. The results of the opportunistic screening scenarios were presented for the baseline model variant only. The maximal annual breast cancer mortality rate reduction was reached in 2018, 20 years after the start of screening. Abbreviations**:** opportunistic screening (OS), mammography screening programme (MSP)

Table 5.3 Costs of diagnostics and treatment per unit (€)

Screen invitation	1.2
Screen mammography	138
Diagnostic mammography	171
Opportunistic mammography	171
Imaging diagnostics	114
Minimal invasive diagnostics	653
Clinical breast examination	19
Sentinel node procedure	3313
Tumorectomy without radiotherapy	7126
Tumorectomy with radiotherapy	9791
Mastectomy without radiotherapy	3684
Mastectomy with radiotherapy	6410
Excision axillary lymph nodes	4904
Chemotherapy	1796
Hormonal therapy	989
Follow-up, first year	220
Follow-up, other years	156
Palliative treatment	21,417

Costs in Swiss Francs (CHF) were recalculated in Euro (€), using a currency exchange rate of 1.66 (June 2007).

Table 5.4 Number of diagnostics associated with 1 mammogram, per screening modality

1 mammogram, MSP	0 clinical breast examinations
	0.04 imaging diagnostics (MRI, ultrasound)
	0.02 minimal invasive diagnostic examinations
1 mammogram, OS	2.9 clinical breast examinations
	0.5 imaging diagnostics (MRI, ultrasound)
	0.004 minimal invasive diagnostic examinations
1 diagnostic mammogram, outside screening	5.0 clinical breast examinations
	0.2 imaging diagnostics (MRI, ultrasound)
	0.1 minimal invasive diagnostic examinations

Abbreviations**:** mammography screening programme (MSP), magnetic resonance imaging (MRI), opportunistic screening (OS)

and treatment are shown in Table 5.3. For each opportunistic mammogram, 2.9 clinical breast examinations were performed. In 53% of the opportunistic mammograms, an additional ultrasound or MRI was performed. In a MSP, on the contrary, only 4.3% of the mammograms were followed by an imaging examination, and no additional clinical breast examinations were done (Table 5.4). Consequently, the costs of diagnostics other than mammography in the baseline 80% OS scenario were predicted to be 305 million euros higher than in a scenario without screening, while in the 80% MSP scenario, the

Table 5.5 Predicted effects, costs and cost-effectiveness of opportunistic (OS) and mammography screening programme (MSP) scenarios with varying participation rates, compared to a no-screening scenario, 3% discounted

	Sc 0	**Sc 1**	**Sc 2**	**Sc 3**	**Sc 4**	**Sc 5**
	No screening	40% OS	80% OS	80% MSP	60% MSP 20% OS	40% MSP 40% annual OS
Effects						
Life years (n)	21,290,900	+16,900	+33,700	+34,000	+33,925	+37,950
QALYs (n)	21,239,159	+15,656	+31,161	+31,506	+31,547	+35,179
Costs (x €10^6)						
Screening	0	+250	+500	+406	+430	+680
Diagnostics	269	+153	+305	–5	+72	+304
Primary treatment	421	+17	+34	+34	+34	+36
Adjuvant treatment	41	–1	–2	–2	–2	–2
Follow-up	113	+7	+15	+15	+15	+16
Palliative care	395	–29	–57	–57	–57	–64
Total costs (x €10^6)	*1239*	*+398*	*+796*	*+391*	*+492*	*+971*
Cost-effectiveness (€)						
Costs per life year gained	-	23,547	23,617	11,512	14,507	25,584
Costs per QALY gained	-	25,418	25,541	12,424	15,601	27,599

To calculate cost-effectiveness, the difference in costs between the scenario without screening (scenario 0) and the scenarios with screening (1–5) were divided by the difference in effects. All effects and costs were rescaled to a population of 1,000,000 women, and calculated over the whole lifespan of the simulated cohort. The results of the opportunistic screening scenarios were presented for the baseline model variant only. Abbreviations**:** opportunistic screening (OS), mammography screening programme (MSP), quality-adjusted life year (QALY)

costs were 5 million euros lower (Table 5.5, scenarios 2 and 3 *versus* scenario 0, 3% discounted). Related to the higher costs of an opportunistic mammogram (Table 5.3), the costs of screening were higher for OS than for MSP (€500 million *versus* €406 million, Table 5.5). For both the 80% MSP and 80% OS scenarios, the relative costs of primary treatment increased compared to a situation without screening, by 34 million euros (without screening: €421 million). This was related to the improved stage distribution and to the fact that screen-detected cancers and smaller tumours were more commonly treated by (the more expensive) tumorectomy, while clinically diagnosed cancers and larger-sized tumours were more frequently treated by (the cheaper) mastectomy. Related to the increased cancer incidence and improved survival, the costs of follow-up also increased, by €15 million, in both the 80% OS and 80% MSP scenarios (without screening €113 million). Cost savings were predicted in palliative care: the more breast cancer deaths prevented, the lower the costs (up to a reduction of €57 million in the 80% OS and 80% MSP scenarios). The total costs increased proportionally with the fraction of screened women. Costs rose only slightly with an improved test performance: in the optimistic 80% OS variant, the total costs were €793 million higher than in a situation

Table 5.6 Predicted effects, costs and cost-effectiveness for an optimistic, baseline and pessimistic model variant of OS, assuming a 25% lower, similar and a 25% higher false-negative rate of OS compared to MSP, participation rate 80%, 3% discounted

Model Variant	No Screening	80% OS Optimistic	80% OS Baseline	80% OS Pessimistic
Effects				
Breast cancer deaths (n)	18,421	–2762	–2658	–2542
Life years (n)	21,290,900	+35,000	+33,700	+32,300
Quality-adjusted life years (N)	21,239,159	+32,430	+31,161	+29,889
Total costs (x € 10^6)	1239	+793	+796	+798
Cost-effectiveness (€)				
Costs per life year gained	-	22,671	23,617	24,707
Costs per quality-adjusted life years gained	-	24,467	25,541	26,700

To calculate cost-effectiveness, the difference in costs between the scenario without screening and the scenarios with screening (the optimistic, baseline and pessimistic model variant) were divided by the difference in effects. All effects and costs were rescaled to a population of 1,000,000 women, and calculated over the whole lifespan of the simulated cohort. Abbreviations**:** opportunistic screening (OS), mammography screening programme (MSP)

without screening, while in the pessimistic variant, the costs were €798 million higher (without screening: €1239 million, 3% discounted, Table 5.6).

Cost-effectiveness

Each life year gained (LYG) in the 80% MSP scenario is predicted to cost €11,512 (Table 5.5, 3% discounting). In the 80% OS scenario, the costs per LYG would increase to €22,671 in the optimistic model variant, and to €24,707 in the pessimistic variant (Table 5.6). A 5% discount rate raised the cost-effectiveness ratio (CER) of 80% MSP to €17,141 per LYG, and that of 80% OS to €34,318 per LYG (data not shown). Costs per life year gained increased proportionally with the fraction of women screened opportunistically. Participation rates did not strongly influence the CER. Quality-adjustment of gained life years resulted in less favourable CERs: €12,424 in the 80% MSP scenario and €25,541 in the 80% OS scenario (Table 5.5, 3% discounted).

5.4 DISCUSSION

This study showed that mammography screening in Switzerland is likely to be effective in reducing breast cancer mortality. A biennial organised mammography screening programme, covering 80% of the 50–69 years old Swiss female population, was predicted to reduce the breast cancer mortality by 13% at the population level and by 20% among women aged 55–74. Opportunistic screening with a similar participation rate was predicted to reach comparable results. The costs of OS per life year gained, however, were twice that of MSP.

The predicted breast cancer mortality reduction in this study seems somewhat lower than the reductions of 20–31% observed in randomised controlled trials and nationwide mammography screening programmes.[4–6,13,14,40] Several programme evaluations predicted reductions of 17–19% at the population level.[17,41,42] OS has been estimated to reduce breast cancer mortality by 8–23% in the USA population[43] (participation 70%[44]). Because screening effects were measured over the whole period that the simulated population is alive, rather than during the screening period only, and because a breast cancer mortality reduction among screened women gradually decreases once screening has ended, our predicted reduction is lower than in the above-mentioned studies. However, an analysis over the whole lifespan enables all potential effects and costs to be accounted for. A shorter period of analysis would increase the predicted breast cancer mortality reduction, up to a maximal reduction of 25% at the population level, and 32% among 55–74 years old women (80% MSP) in 2018, 20 years after the start of screening.

We predicted the breast cancer mortality reduction for women aged 30–70 years in 1999, because these women will at least once be targeted for screening. Including younger and older women would lower the predicted mortality reduction. The breast cancer mortality reduction may also have been overestimated by the fact that the predicted stage distribution at subsequent screening examinations was more favourable than actually observed. Because this was counteracted with a less favourably modelled stage distribution at first screening examinations, the inaccuracies in the mortality predictions due to lack of fit between modelled and observed breast cancer were likely to be small.

Our study supports the findings of Neeser *et al.*[12] that in Switzerland, MSP is cost-effective compared to OS. By assuming a larger screening-related breast cancer mortality reduction than Neeser *et al.* (25% annually, instead of 15%), and by taking into account cost-savings related to palliative care, we expect that the incremental cost-effectiveness of MSP opposed to OS would be substantially more favourable than €53,677 per LYG[a] (starting screening at age 50), as predicted in the above-mentioned study.

Nevertheless, the cost-effectiveness ratio (CER) of breast cancer screening in Switzerland was high compared to other countries. The 3% discounted CER of programme-based screening, targeting women aged 50–69 years, varied between €2207 per life year gained (LYG) in the Netherlands[9] and €13,458 per LYG in Finland.[a, 8] The predicted 3% discounted CER in Switzerland of €11,512 was in line with the Finnish estimate. The 5% discounted CERs of MSPs in various European countries and Australia ranged from €2650 to €8300 per LYG.[45] The corresponding Swiss prediction of €17,141 was higher than these estimates, even after correction of these estimates for inflation with (harmonised) consumer price indices[46,47] (corrected CERs: between €3557 and €11,962 per LYG). Opportunistic breast cancer screening in Switzerland was comparably cost-effective as decentralised breast cancer screening in the USA: €23,617 per LYG in Switzerland *versus* €24,140 per LYG[44] in the USA (the latter was quality-adjusted and based on screening women in the age of 50–75 years; both 3% discounted).

The relatively unfavourable cost-effectiveness ratio of mammography screening is related to the health care costs in Switzerland, which are among the highest in Europe. The estimated costs of a programme-mammogram, which were in line with the previous Swiss estimates,[12] were circa 2.5 times higher than, for instance, in the Netherlands. Reducing these costs to Dutch cost-levels (€50[19] instead of €138) improved the 3% discounted CER of MSP to €3967 per LYG, which is on the same level as the CER of mammography

[a] Costs in US dollars were recalculated in Euro using a currency exchange rate of 0.71 (October 2007)

screening in other western countries. A 50% reduction in the use of imaging examinations that are done in combination with an opportunistic mammogram could lower the CER of OS from €23,617 to €20,971 per LYG. A further 50% reduction of the costs of an opportunistic mammogram would decrease the CER to €13,550 per LYG.

Without data on breast cancer treatment and with limited data on OS, the assessment of the cost-effectiveness of screening in Switzerland includes uncertainties. The costs of clinical breast examinations related to OS may have been overestimated, because we used the ratio between all reimbursed breast examinations and mammograms to estimate resource use, regardless whether the clinical breast examination was indeed related to the mammogram. Costs of diagnostics may have been underestimated, because only health insurance-reimbursed costs could be included. Screening-related examinations done without seeking reimbursement and investigations performed in an in-patient setting could not be accounted for. The costs of treatment were based on Dutch data from 1991,[36] which likely have increased in later years. Indirect costs of screening, such as the additional health care costs made if a woman is saved by mammography, costs that would be necessary to increase screening participation and personal time costs were not included in the analysis.

As the model was calibrated with breast cancer and screening data from Vaud, the extrapolation of the results to the whole of Switzerland may involve some uncertainties. Mammography screening practice and quality, and possibly, breast cancer treatment varies across the country.[48] Nevertheless, it is unlikely that the natural history of breast cancer differs much between cantons, because such differences were also small between the Swiss model and, for instance, the Dutch screening evaluation model.[35] The baseline, optimistic and pessimistic scenarios for test sensitivity of OS reflected variations in opportunistic screening 'quality'. It might be argued that the performance of OS will be lower than MSP, related to less 'mammographic experience' of radiologists who work outside a programme, less possibility of discussion and feedback between screening, diagnostic and treatment disciplines, less (specific) training of radiologists and technicians, a lower number of readings and lower technical quality control.[16,49] A recent Danish study showed that programme-mammograms were considerably more sensitive than those performed opportunistically.[50] Improved performance of OS may be plausible as well, in particular when other imaging diagnostics are used in combination with mammography. For instance, a review of breast cancer cases detected by opportunistic screening in Austria showed a favourable tumour stage distribution, with T2+ tumours only comprising 10% of all breast cancer cases.[51] This could be related to the fact that women aged 35 and older are screened and to the average screening interval of 16 months, but also to additional diagnostics, such as breast ultrasound, that are performed

in combination with mammography in 64% of the cases. Studies comparing the performance of mammography between countries with organised screening and countries with opportunistic screening showed no differences between the two screening modalities in the detection of larger-sized tumours[49] or in the prognostic characteristics of invasive screen-detected and interval tumours.[52] A comparison between centralised and decentralised screening projects within the European Breast Cancer Network neither showed differences in performance indicators.[53] These studies, however, did not account for specific breast cancer and screening characteristics (e.g. screening interval and screening age) between the various countries. Within Switzerland, cantonal differences in screening outcomes indicate that the performance of mammography varies regionally. For example, in Valais, OS resulted in a slightly more favourable prognostic profile than MSP, which might indicate that the test sensitivity of OS was better than that of MSP.[32] On the contrary, in Vaud, no differences in prognostic profile were observed between MSP and OS, which might indicate that the two screening modalities have a similar test sensitivity in this canton.[32] Although the hypothetical model variants of a 25% lower, similar and a 25% higher false-negative rate of OS compared to MSP are chosen rather arbitrarily, these scenarios do reflect likely variations in screening performance across Switzerland and the way they influence screening effects. Higher or lower false-negative rates of OS might be possible, but the true performance of opportunistic screening is difficult to assess. Several screening scenarios with varying participation rates and screening intervals were analysed to account for differences in screening practice across Switzerland. An 80% population coverage by OS was not considered a likely representation of current or future screening practice, in particular in a decentralised health care setting.[53] Several organised programmes, however, reach participation rates around 80%.[34] The actual breast cancer mortality reduction in the population, to be obtained by OS, is therefore likely to be smaller than that of MSP, and lower than the 13% (80% OS) predicted in our study.

Although the analysed scenarios were less cost-effective than breast cancer screening in other European countries, they were cost-effective according to WHO guidelines, which defined a CER smaller than the GDP per capita as 'very cost-effective' and a CER that is one to three times the GDP per capita as 'cost-effective'[7] (GDP per capita in Switzerland approximately €35,000[54]). However, strong improvements can be obtained if the use of imaging diagnostics would be diminished and costs were decreased. A centralised screening centre would enhance the possibilities of multiple readings, discussion and feedback, which should increase the screening volume and performance. It would also enable a more effective use of equipment and a faster acquisition of work-up diagnostics. Furthermore, continuous quality control and evaluation of screening could ensure that maximum benefits are obtained at reasonable costs.

Conclusion

Both organised and opportunistic screening are predicted to be effective in reducing breast cancer mortality in Switzerland. However, the costs of opportunistic screening per life year gained were twice those of organised screening: € 23,617 *versus* € 11,512. For opportunistic screening to become equally cost-effective as organised screening, costs and use of additional diagnostics should be reduced.

Acknowledgements

The authors are grateful to Prof. Levi (Vaud Tumour Registry) for providing pre-screening cancer registry data and for his contribution to the study design, and to Dr. Bouchardy (Geneva Tumour Registry), for pre-screening cancer registry data. Dr. Beck, Mr. von Rotz (CSS Health Insurance Company) and Dr. Guetg (SantéSuisse) are also acknowledged for their most helpful advice and provision of health insurance data.

REFERENCES

1. Botha JL, Bray F, Sankila R, Parkin DM. Breast cancer incidence and mortality trends in 16 European countries. Eur J Cancer 2003;39(12):1718–29.
2. Karim-Kos HE, de Vries E, Soerjomataram I, Lemmens V, Siesling S, Coebergh JW. Recent trends of cancer in Europe: a combined approach of incidence, survival and mortality for 17 cancer sites since the 1990s. Eur J Cancer 2008;44(10):1345–89.
3. Hery C, Ferlay J, Boniol M, Autier P. Quantification of changes in breast cancer incidence and mortality since 1990 in 35 countries with Caucasian-majority populations. Ann Oncol 2008;19(6):1187–94.
4. Blanks RG, Moss SM, McGahan CE, Quinn MJ, Babb PJ. Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990–8: comparison of observed with predicted mortality. BMJ 2000;321(7262):665–9.
5. Olsen AH, Njor SH, Vejborg I, *et al.* Breast cancer mortality in Copenhagen after introduction of mammography screening: cohort study. BMJ 2005;330(7485):220.
6. Otto SJ, Fracheboud J, Looman CW, *et al.* Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. Lancet 2003;361(9367):1411–7.
7. World Health Organization. Macroeconomics and health: investing in health for economic development: report of the Commission on Macroeconomics and Health. Geneva; 2001.
8. Leivo T, Sintonen H, Tuominen R, Hakama M, Pukkala E, Heinonen OP. The cost-effectiveness of nationwide breast carcinoma screening in Finland, 1987–1992. Cancer 1999;86(4):638–46.
9. Groenewoud JH, Otten JD, Fracheboud J, *et al.* Cost-effectiveness of different reading and referral strategies in mammography screening in the Netherlands. Breast Cancer Res Treat 2007;102(2):211–8.
10. van Ineveld BM, van Oortmarssen GJ, de Koning HJ, Boer R, van der Maas PJ. How cost-effective is breast cancer screening in different EC countries? Eur J Cancer 1993;29A(12):1663–8.
11. Swiss Cancer League. Swiss Cancer League Request "Reimbursement of screening mammography as a service of basic health insurance" to the Federal Office of Health. Bern; 2007.
12. Neeser K, Szucs T, Bulliard J-L, Bachmann G, Schramm W. Cost-effectiveness analysis of a quality-controlled mammography screening program from the Swiss statutory health-care perspective: quantitative assessment of the most influential factors. Value Health 2007;10(1):42–53.
13. Tabar L, Fagerberg CJ, Gad A, *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 and Welfare. Lancet 1985;1(8433):829–32.
14. Nystrom 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(9310):909–19.
15. van Oortmarssen GJ, Habbema JD, van der Maas PJ, *et al.* A model for breast cancer screening. Cancer 1990;66(7): 1601–12.
16. Warmerdam PG, de Koning HJ, Boer R, *et al.* Quantitative estimates of the impact of sensitivity and specificity in mammographic screening in Germany. J Epidemiol Commun Health 1997;51(2):180–6.
17. Beemsterboer PM, de Koning HJ, Warmerdam PG, *et al.* Prediction of the effects and costs of breast-cancer screening in Germany. Int J Cancer 1994;58(5):623–8.

18. Rijnsburger AJ, van Oortmarssen GJ, Boer R, *et al.* Mammography benefit in the Canadian National Breast Screening Study-2: a model evaluation. Int J Cancer 2004;110(5):756–62.
19. National Evaluation Team for Breast cancer screening. National evaluation of breast cancer screening in the Netherlands – eleventh evaluation report. Rotterdam: Department of Public Health, Erasmus MC; 2005.
20. Collette HJ, Day NE, Rombach JJ, de Waard F. Evaluation of screening for breast cancer in a non-randomised study (the DOM project) by means of a case-control study. Lancet 1984;1(8388):1224–6.
21. Collette HJ, de Waard F, Rombach JJ, Collette C, Day NE. Further evidence of benefits of a (non-randomised) breast cancer screening programme: the DOM project. J Epidemiol Commun Health 1992;46(4):382–6.
22. Verbeek AL, Hendriks JH, Holland R, Mravunac M, Sturmans F, Day NE. Reduction of breast cancer mortality through mass screening with modern mammography. First results of the Nijmegen project, 1975–1981. Lancet 1984;1(8388):1222–4.
23. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. Cancer 1989;63(1):181–7.
24. Tabar L, Vitak B, Chen HH, *et al.* The Swedish two-county trial twenty years later. Updated mortality results and new insights from long-term follow-up. Radiol Clin North Am 2000;38(4):625–51.
25. Michaelson JS, Silverstein M, Wyatt J, *et al.* Predicting the survival of patients with breast carcinoma using tumor size. Cancer 2002;95(4):713–23.
26. Michaelson JS, Silverstein M, Sgroi D, *et al.* The effect of tumor size and lymph node status on breast carcinoma lethality. Cancer 2003;98(10):2133–43.
27. Sant M, Allemani C, Capocaccia R, *et al.* Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. Int J Cancer 2003;106(3):416–22.
28. de Koning HJ, Boer R, Warmerdam PG, Beemsterboer PM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. J Natl Cancer Inst 1995;87(16):1217–23.
29. Bjurstam N, Bjorneld L, Warwick J, *et al.* The Gothenburg breast screening trial. Cancer 2003;97(10):2387–96.
30. Bulliard J-L, De Landtsheer JP, Levi F. Results from the Swiss mammography screening pilot programme. Eur J Cancer 2003;39(12):1761–9.
31. Levi F, Te VC, la Vecchia C. Impact of mammography on breast cancer incidence in Vaud, Switzerland. J Natl Cancer Inst 1991;83(16):1181–2.
32. Bulliard J-L, Levi F. Evaluation épidémiologique du programme vaudois de dépistage du cancer du sein, 1999–2006. Lausanne: Raisons de santé ; 2007. p. 136. <http://www.iumsp.ch/ Publications/pub/rds136_fr.pdf>.
33. 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(4):538–44.
34. Lynge E, Olsen AH, Fracheboud J, Patnick J. Reporting of performance indicators of mammography screening in Europe. Eur J Cancer Prev 2003;12(3):213–22.
35. Groenewoud JH, Pijnappel RM, van den Akker-Van Marle ME, *et al.* Cost-effectiveness of stereotactic large-core needle biopsy for nonpalpable breast lesions compared to open-breast biopsy. Br J Cancer 2004;90(2):383–92.

36. de Koning HJ, van Ineveld BM, van Oortmarssen GJ, *et al.* Breast cancer screening and cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors. Int J Cancer 1991;49(4):531–7.
37. Organisation for Economic Co-operation and Development. Purchasing Power Parities Data, April 2007. <http://www.oecd.org/document/47/0,3343,en_2649_34357_36202863_1_1_1_1,00.html.
38. Jeruss JS, Hunt KK, Xing Y, *et al.* Is intraoperative touch imprint cytology of sentinel lymph nodes in patients with breast cancer cost effective? Cancer 2006;107(10):2328–36.
39. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. New York: Oxford University Press;1996.
40. Jonsson H, Nystrom L, Tornberg S, Lenner P. Service screening with mammography of women aged 50–69 years in Sweden: effects on mortality from breast cancer. J Med Screen 2001;8(3):152–60.
41. Tornberg S, Carstensen J, Hakulinen T, Lenner P, Hatschek T, Lundgren B. Evaluation of the effect on breast cancer mortality of population based mammography screening programmes. J Med Screen 1994;1(3):184–7.
42. van den Akker-van Marle ME, Reep-van den Bergh CM, Boer R, Del Moral A, Ascunce N, de Koning HJ. Breast cancer screening in Navarra: interpretation of a high detection rate at the first screening round and a low rate at the second round. Int J Cancer 1997;73(4):464–9.
43. Cronin KA, Feuer EJ, Clarke LD, Plevritis SK. Impact of adjuvant therapy, mammography on U.S. mortality from 1975 to 2000: comparison of mortality results from the cisnet breast cancer base case analysis. J Natl Cancer Inst Monogr 1975;2006(36):112–21.
44. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. J Natl Cancer Inst 2006;98(11):774–82.
45. De Koning HJ. Breast cancer screening; cost-effective in practice? Eur J Radiol 2000;33(1):32–7.
46. Eurostat. Harmonised indices of consumer prices, 2008. <http://epp.eurostat.ec.europa.eu/portal/page?_pageid=2714,1,2714_61582043&_dad=portal&_schema= PORTAL>.
47. Reserve Bank of Australia. Consumer Price Index, 2008. http://www.rba.gov.au/Statistics/Bulletin/G02hist.xls>.
48. Bulliard JL, La Vecchia C, Levi F. Diverging trends in breast cancer mortality within Switzerland. Ann Oncol 2006;17(1):57–9.
49. Smith-Bindman R, Ballard-Barbash R, Miglioretti DL, Patnick J, Kerlikowske K. Comparing the performance of mammography screening in the USA and the UK. J Med Screen 2005;12(1):50–4.
50. Bihrmann K, Jensen A, Olsen AH, *et al.* Performance of systematic and non-systematic ('opportunistic') screening mammography: a comparative study from Denmark. J Med Screen 2008;15(1):23–6.
51. Frede TE. Opportunistic breast cancer early detection in Tyrol, Austria 1996–2004. Is a mammography-screening program necessary? Eur J Radiol 2005;55(1):130–8.
52. Hofvind S, Vacek PM, Skelly J, Weaver DL, Geller BM. Comparing screening mammography for early breast cancer detection in Vermont and Norway. J Natl Cancer Inst 2008;100(15):1082–91.
53. Broeders MJ, Scharpantgen A, Ascunce N, *et al.* Comparison of early performance indicators for screening projects within the European Breast Cancer Network: 1989–2000. Eur J Cancer Prev 2005;14(2):107–16.
54. International Monetary Fund. World Economic and Financial Surveys, World Economic Outlook Database, 2007. <http://www.imf.org>.

Appendix 5.1. Model parameters on natural history of breast cancer and screening

Mean duration (years) of screen-detectable preclinical stage by age	Age			
Stage	40	50	60	70
Preclinical DCIS	5.2	5.2	5.2	5.2
Preclinical T1a (diameter ≤5 mm)	0.2	0.3	0.4	0.5
Preclinical T1b (diameter 6–10 mm)	0.4	0.6	0.9	1.1
Preclinical T1c (diameter 11–20 mm)	1.1	1.4	2.1	2.6
Preclinical T2+ (diameter >20 mm)	0.6	0.8	1.2	1.5

Long-term relative survival by clinical stage and age	Stage								
Age	DCIS	T1aN–	T1aN+	T1bN–	T1bN+	T1cN–	T1cN+	T2+N–	T2+N+
40	1.000	0.853	0.676	0.809	0.595	0.712	0.435	0.508	0.195
50	1.000	0.866	0.701	0.826	0.623	0.734	0.467	0.54	0.22
60	1.000	0.851	0.671	0.807	0.588	0.707	0.425	0.5	0.181
70	1.000	0.854	0.676	0.810	0.594	0.712	0.432	0.507	0.187

Sensitivity of mammography by stage, age ≥50 years	Study variant			
Stage	MSP	OS, optimistic model variant A	OS, baseline model variant B	OS, pessimistic model variant C
Preclinical DCIS	80%	85%	80%	75%
Preclinical T1a (diameter ≤5 mm)	70%	77.5%	70%	62.5%
Preclinical T1b (diameter 6–10 mm)	75%	81.25%	75%	68.75%
Preclinical T1c (diameter 11–20 mm)	80%	85%	80%	75%
Preclinical T2+ (diameter >20 mm)	100%	100%	100%	100%

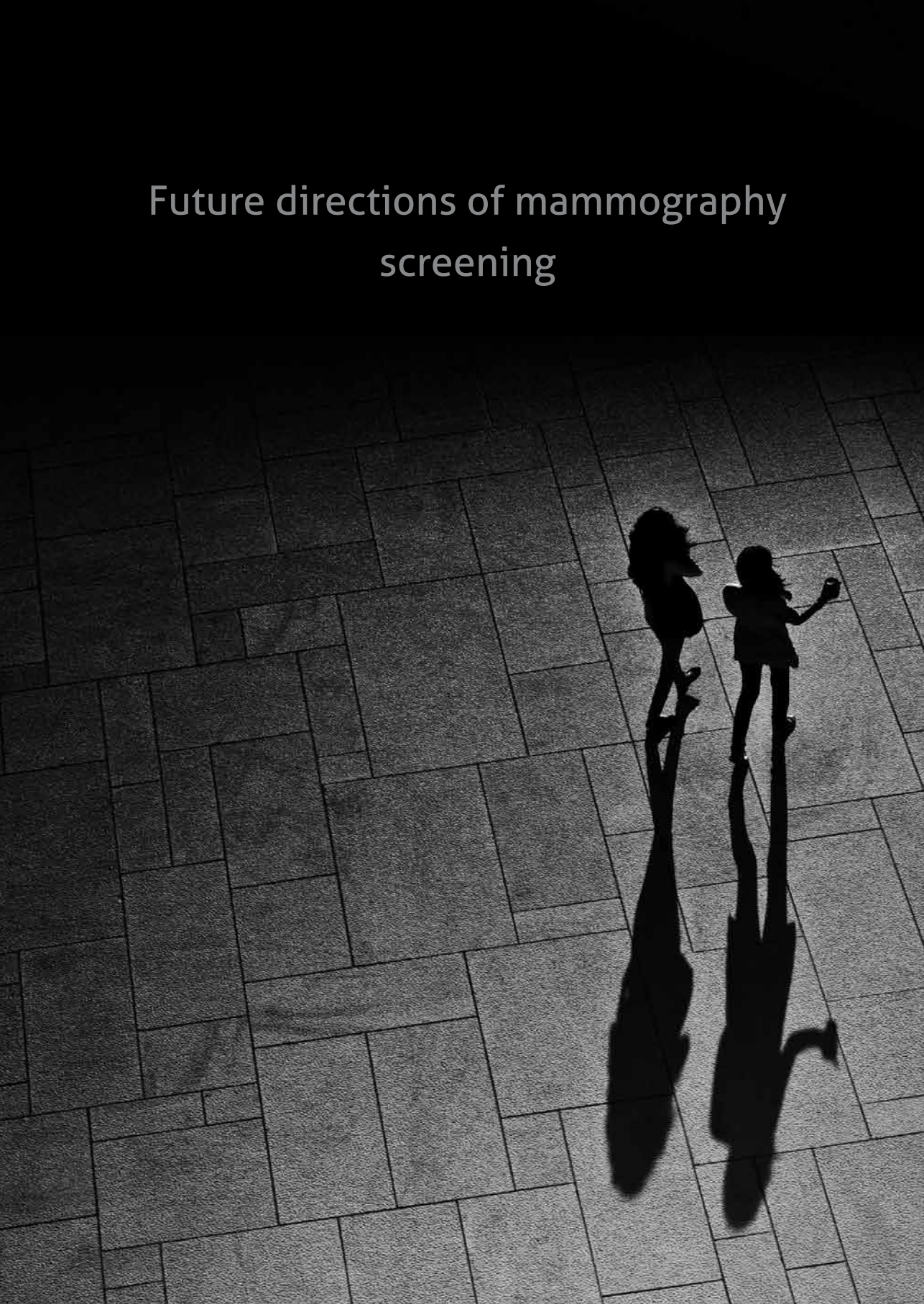
Reduction in risk of dying of breast cancer by stage at age 50		
Stage	Reduction in risk, N–	Reduction in risk, N+
Preclinical DCIS	100%	100%
Preclinical T1a (diameter ≤5 mm)	86.60%	70.10%
Preclinical T1b (diameter 6–10 mm)	82.60%	62.30%
Preclinical T1c (diameter 11–20 mm)	73.40%	46.70%
Preclinical T2+ (diameter >20 mm)	54.00%	22.00%

Quality of life; durations and utilities

Health stage	Duration[30]	Utility[30]
Terminal illness	1 month	0.712
Palliative therapy+ chemotherapy	4 months	0.469
Palliative therapy+ radiotherapy	1 month	0.419
Palliative therapy+ surgical therapy	5 weeks	0.383
Palliative therapy+ hormonal therapy	14 months	0.337
Initial chemotherapy	6 months	0.283
Initial radiotherapy	2 months	0.197
Initial surgery	2 months	0.133
Initial hormonal therapy	2 years	0.18
Excision lymph nodes	1 month	0.1
Sentinel node procedure	1 month	0.1
Disease-free 2 months - 1 year after mastectomy	10 months	0.156
Disease-free 2 months - 1 year after breast saving therapy	10 months	0.086
Disease-free >1 year after mastectomy	1 year	0.053
Disease-free >1 year after breast saving therapy	1 year	0.04
Screening attendance	1 week	0.006

Abbreviations**:** ductal carcinoma *in situ* (DCIS), diameter ≤5 mm lymph node negative (T1aN–), diameter 6–10 mm lymph node negative (T1bN–), diameter 11–20 mm lymph node negative (T1cN–), diameter >20 mm lymph node negative (T2+N–), diameter ≤5 mm lymph node positive (T1aN+), diameter 6–10 mm lymph node positive (T1bN+), diameter 11–20 mm lymph node positive (T1cN+), diameter >20 mm lymph node positive (T2+N+)

Future directions of mammography screening



The effects of population-based mammography screening starting between age 40 and 50 compared to the effects of adjuvant systemic therapy

Rianne de Gelder, Eveline A.M. Heijnsdijk, Jacques Fracheboud,
Gerrit Draisma, Harry J. de Koning

Submitted.

ABSTRACT

Background

Adjuvant systemic therapy has been shown to be effective in reducing breast cancer mortality. The additional effect of mammography screening remains uncertain, in particular for women aged 40–49 years. We therefore assessed the effects of screening starting between age 40 and 50, as compared to the effects of adjuvant systemic therapy.

Methods

The use of adjuvant endocrine therapy, chemotherapy and the combination of endocrine- and chemotherapy, as well as the uptake of mammography screening in the Netherlands was modelled using micro-simulation. With the model, the effects of 1) adjuvant therapy, 2) biennial screening between age 50 and 74 (current situation) in the presence of adjuvant therapy, and 3) extending the current screening programme with 1–10 extra examinations between age 40 and 50 were assessed, by comparing breast cancer mortality in women aged 0–100 years in scenarios with and without these interventions.

Results

In 2008, adjuvant treatment was estimated to have reduced the breast cancer mortality rate in the simulated population by 13.9%, compared to a situation without treatment. Biennial screening between age 50 and 74 further reduced the mortality rate by 15.7%. Extending screening to age 48 would lower the mortality rate by 1.0% compared to screening from age 50; 10 additional screening rounds between age 40 and 49 would reduce this rate by 5.1%.

Conclusions

Adjuvant systemic therapy reduced breast cancer mortality by 13.9%; mammography screening additionally decreased mortality by 15.7%. Expanding the lower age limit of screening would further reduce breast cancer mortality.

6.1 INTRODUCTION

Since clinical trials showed that adjuvant systemic treatment has the potential to reduce breast cancer mortality,[1] the use of adjuvant breast cancer therapies has increased substantially. Around the same time, randomized controlled trials showed that mammography screening could reduce breast cancer mortality by 25%–30% in women aged 50–69.[2–3] Consequentially, breast cancer screening also became more common. Most of the screening trials, however, started in a period in which adjuvant systemic therapy was not widely used, and there is strong debate whether the present-day use of adjuvant therapy would have affected the effects of screening (and vice versa).[4–5]

Moreover, due to ongoing developments in their dissemination and effectiveness, the effects of both adjuvant systemic treatment and screening may change. New types of adjuvant therapy have been developed or have more effectively been used in specific target groups.[6] The uptake of mammography screening has also increased,[7] and new developments such as digital mammography may increase its effectiveness.[8] Furthermore, the breast cancer incidence in various countries is on the rise, due to a higher prevalence of known breast cancer risk factors.[9–10] This paper therefore assesses the effects of adjuvant systemic therapy and mammography screening in a current health care system.

Particular attention is paid to screening below age 50. Currently most countries do not recommend screening below age 50. The outcomes of the UK Age Trial, however, showed a borderline significant 17% reduction in breast cancer mortality in screened women aged 40–48 years.[11] This raises the question whether women below the age of 50 should be targeted for screening, and from which age. Similar to older women, the benefits of mammography screening for women younger than 50 may be affected by the use of adjuvant therapy and rising incidence. We therefore assess the effects of mammography screening at various starting ages between age 40 and 50, given the changing trends in incidence and the changing use of adjuvant systemic therapy.

6.2 METHODS

MISCAN model

The effects of adjuvant treatment and screening were calculated using micro-simulation modelling. Micro-simulation models can include observed trends in breast cancer incidence and mortality, and assess how certain interventions, such as screening or adjuvant therapy, affect such trends. Models have the advantage that they can extrapolate findings from screening and adjuvant treatment trials to actual populations, allow for comparison

of a variety of intervention strategies, and can separate screening effects from adjuvant treatment effects.[12–13]

The MIcrosimulation SCreening ANalysis model MISCAN has originally been designed to study the effects of screening.[14–15] First, a population of individual life biographies is simulated based on specific demographic characteristics of the population under study. The natural history of breast cancer without screening is then modelled. Some women in the simulated population may develop breast cancer, which develops from a small preclinical lesion to a symptomatic cancer, possibly leading to breast cancer death. In MISCAN, this progression is modelled as a semi-Markov process through the successive preclinical invasive stages T1a, T1b, T1c and T2+. A fraction of preclinical invasive cancer is preceded by preclinical ductal carcinoma *in situ* (DCIS) (Appendix 6.1). In each stage, a lesion may grow to the next stage or become clinically diagnosed because of symptoms. The natural course of the disease may be interrupted by screening, at which a preclinical lesion can become screen-detected. Screen-detection can result in the detection of smaller tumours than that might have been diagnosed if no screening had taken place, which may entail a survival benefit. Each screen-detected or clinically diagnosed tumour may be treated with adjuvant systemic therapy, which also improves survival.

Model parameters

The mean duration of preclinical screen-detectable cancer, the probability of a transition between the stages, and the sensitivity of mammography were estimated using data from the screening organizations and Comprehensive Cancer Centres.[7, 16] (Appendix 6.1) These data included the incidence of breast cancer between 1975 and 1990, which is the period before screening was implemented (only data from the Eindhoven Cancer Registry were available for this period); the incidence of clinically diagnosed and screen-detected cancers between 1990 and 2008; and the rates of screen-detected and interval cancer, available for the periods 1990–2007 and 1990–2005, respectively. All data were specified by age, calendar year, tumour stage and screening round (first/ subsequent). Based on various screening trials and programmes targeting younger women, the test sensitivity for women under age 50 was estimated to be 25% lower than that of older women.[2] Other screening characteristics in the model, such as screening ages, intervals and attendance were similar to those observed in the Netherlands between 1990 and 2009.[7] From 1990, women aged 50–69 years, and from 1998, women aged 50–74 years were biennially invited to participate. Between the start of screening and 2009, the attendance rate increased between 75% and 82%.[7]

Data on adjuvant systemic therapy in the model (Appendix 6.2a–c) were derived from the Dutch screening organizations. These data included the age-, tumour stage-, de-

tection method- and calendar year specific probability of being treated with adjuvant endocrine therapy, chemotherapy or a combination of endocrine and chemotherapy, between 1990 and 2004. For the period 1975–1990, the probability of using adjuvant therapy was based on the Eindhoven Cancer Registry.[17] The combined use of endocrine and chemotherapy was not registered at that time; we assumed that no such treatments were given before 1990. No details were available in our databases on the specific type and duration of the adjuvant therapy that is used, nor on the estrogen receptor (ER) or progesterone receptor (PgR) status of these patients. In agreement with Dutch treatment guidelines,[18] we therefore assumed that all patients on endocrine treatment were ER positive, and were treated with 5 years of tamoxifen. Patients treated with chemotherapy were assumed to have used it for a period of 6–12 months.

Stage- and age- specific survival rates after clinical breast cancer diagnosis were modelled using several international sources.[19–24] In a fraction of cases, early detection by screening prevents death from cancer because of stage-shifting. Cure and survival rates after screen-detection, including those for 40–49 years old women, were based on the Swedish randomized controlled trials.[3, 23, 25–26] The risk of dying from breast cancer after adjuvant treatment was modelled using the published rate ratios from the EBCTCG meta-analysis.[1] Parameter estimates are shown in Appendix 6.1.

Model predictions

With MISCAN, breast cancer mortality was estimated for three scenarios:

A. Assuming no adjuvant therapy or screening,
B. Assuming adjuvant therapy such as practiced in the Netherlands, but no screening,
C. Assuming adjuvant therapy and screening such as practiced in the Netherlands.

The effects of adjuvant systemic therapy and screening were calculated as the reduction in the estimated annual breast cancer mortality rate in 2008, calculated for women aged 40–49, 50–74 and 75–84 years, and for the total simulated population aged 19–100 years in 2008. The rate reduction attributable to adjuvant therapy was assessed by comparing scenario B to A, the rate reduction attributable to screening by comparing scenario C to B, and the mortality rate reduction related to the combination of adjuvant systemic therapy and screening by comparing scenario C to A. The effects of screening before age 50 were calculated by advancing the starting age of screening with two-year intervals: at age 48, 46, [...] and 40. The effects of annual screening between age 40 and 49 were also assessed.

Sensitivity analysis

To account for the uncertainties in our model, effects were also calculated under the assumption that:

- The baseline survival of breast cancer (without adjuvant therapy and screening) does not improve,
- the effectiveness of endocrine therapy is 25% lower than in the baseline model, which could be the case if treatment adherence is lower than in the EBCTCG meta-analysis. Low treatment adherence, for instance, has been observed by van Herk-Sukel *et al.* (2010),[27]
- the test sensitivity for women below age 50 is the same as that for women older than 50,
- the effectiveness of screening women below age 50 is 25% lower than in the baseline model,
- the effectiveness of screening women below age 50 is 25% higher than in the baseline model.

6.3 RESULTS

Model validation

During the period before mammography screening was implemented (1975–1990) and in age groups that were not yet invited, breast cancer incidence increased. By assuming an annual percent change of 1.4% in the background breast cancer incidence (without screening), these trends were modelled well (Figure 6.1a). The observed breast cancer mortality in the pre-screening period was modelled reasonably, but from 1990 onwards the observed breast cancer mortality in all age groups decreased more steeply than simulated. The assumed effects of screening and adjuvant systemic therapy, which were based on randomized controlled screening and treatment trials, were thus insufficient to simulate breast cancer mortality correctly, and it was necessary to model an additional survival improvement from unknown causes of 1.3% per year to obtain a good fit (Figure 6.1b).

Estimated effects of adjuvant systemic treatment in the absence of screening

In all age groups, the observed breast cancer mortality rate decreased between 1975 and 2008 compared to the estimated situation without screening. The estimated contribution of adjuvant systemic treatment to the reduction in breast cancer mortality was shown in Figure 6.2a–c. In the age group 40–49 years, adjuvant therapy reduced the breast cancer mortality rate in 2008 from 34 to 25 deaths per 100,000 woman-years: a decrease

of 27.3%. In women aged 50–74 years, the breast cancer mortality reduction related to adjuvant therapy was 15.3% (a decrease from 100 to 85 breast cancer deaths per 100,000 woman-years). In the population aged 75–84 years, adjuvant systemic therapy lowered the breast cancer mortality rate in 2008 from 179 to 163 deaths per 100,000 woman-years: a reduction of 9.0% compared to a situation without such treatment. In the total simulated population aged 19–100 years in 2008, the estimated breast cancer mortality rate reduction in 2008 attributable to adjuvant therapy was 13.9% (Table 6.1).

The effects of screening of women aged 50–74 years in the presence of adjuvant systemic therapy

In a situation in which breast cancer patients can be treated with adjuvant systemic therapy, the current screening programme, targeting women aged 50–74 years, was estimated to reduce breast cancer mortality from 1990 onwards. In 2008, biennial mammography was estimated to reduce the breast cancer mortality rate in this age group to 67 deaths per 100,000 woman-years: a reduction of 20.9% compared to a situation without screening (Figure 6.2b). In women aged 75–84, screening reduced breast cancer mortality from 1998 onwards, related to the extension of the upper age limit for screening. Ten years later, in 2008, the screening-related reduction of the breast cancer mortality rate compared to a situation without screening was estimated to be 19.4% (Figure 6.2c), resulting in an estimated breast cancer mortality rate of 131 deaths per 100,000 woman-years. In the total simulated population, screening lowered the breast cancer mortality rate in 2008 by 15.7% (Table 6.1).

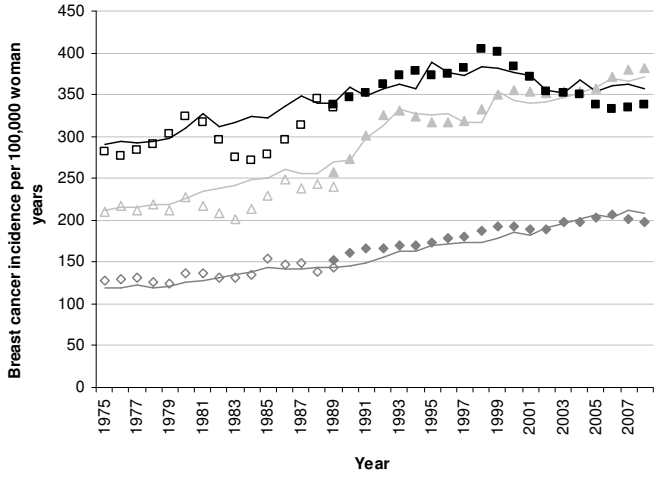
The combined effects of adjuvant therapy and screening

The combined effect of screening and adjuvant systemic treatment on breast cancer mortality was estimated to be lower than the separate interventions, because adjuvant treatment will avert the deaths of some women that would otherwise have profited from screening, and vice versa (Table 6.1). Interaction may also occur because screening changes the age at which cancer is diagnosed, which may influence the effectiveness of adjuvant therapy.[1] In a situation in which women aged 50–74 are biennially screened, the estimated reduction in the breast cancer mortality rate in the total simulated population in 2008, attributable to both adjuvant therapy and screening, was 27.4%, relative to a situation without those interventions.

Lowering the starting age of screening

Lowering the starting age of screening would increase the number of averted breast cancer deaths (Table 6.1). One extra screening round at age 48 would further reduce the breast cancer mortality rate in 2008 in the total population by 1.0%, compared to a situation with adjuvant systemic therapy and screening from age 50; two additional

a.



b.

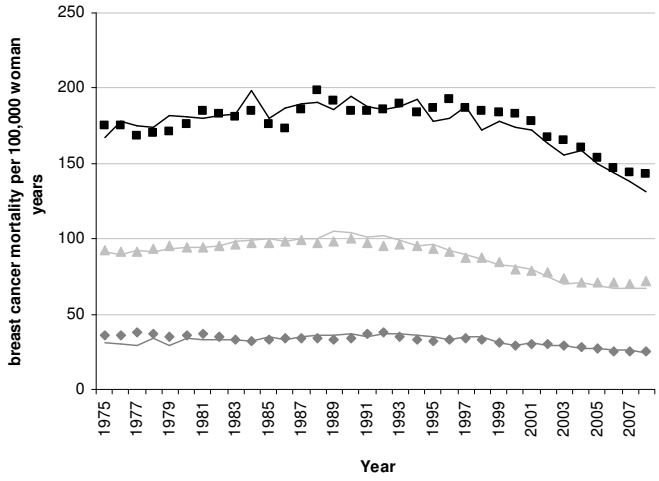


Figure 6.1a–b Observed and predicted breast cancer incidence (a) and breast cancer mortality (b) in women aged 40–49, 50–74 and 75–84 years between 1975 and 2008, per 100,000 woman-years.

Breast cancer mortality was calculated in the presence of adjuvant systemic therapy. The observed mortality was calculated as a 2-year moving average. Abbreviations**:** Eindhoven Cancer Registry (ECR)

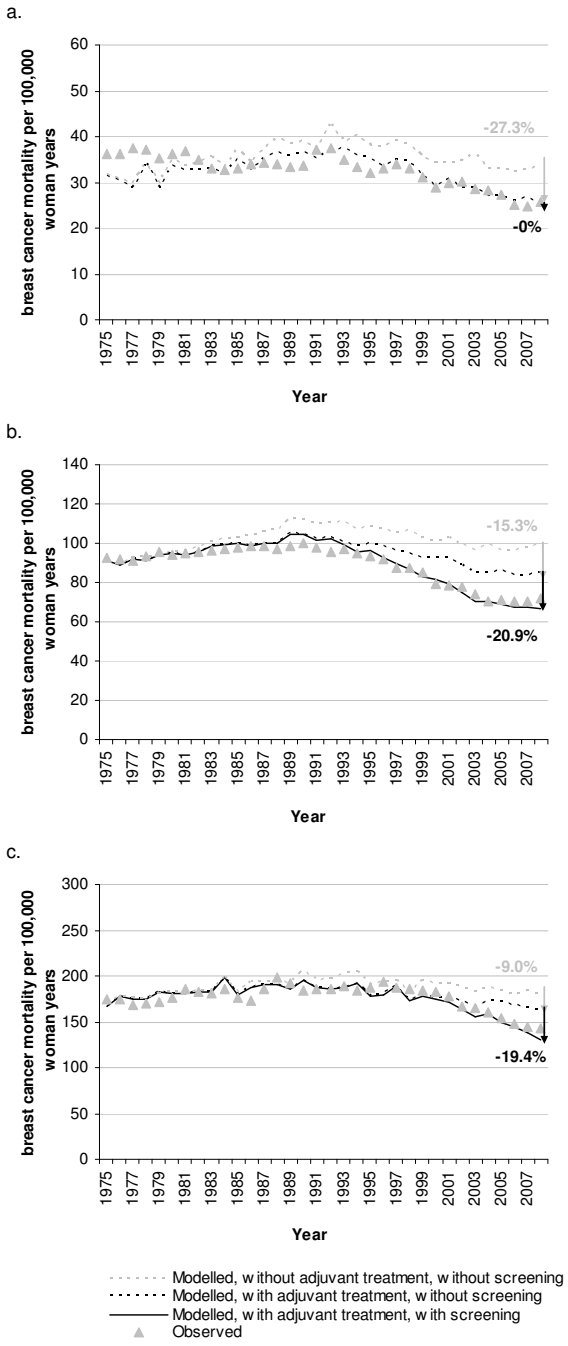


Figure 6.2a–c Predicted breast cancer mortality in the absence and presence of adjuvant treatment and screening, in a) women aged 40–49 years, b) women aged 50–74 years, and c) women aged 75–84 years

Table 6.1 Predicted breast cancer mortality per 100,000 woman-years in a situation with and without adjuvant therapy and with and without screening, in the total simulated population in 2008

	A	B		C		
	No adjuvant therapy	**Adjuvant therapy**	(B–A)/ A	**Adjuvant therapy**	(C–B)/ B	(C–A)/ A
	No screening	**No screening**		**Screening**		
	Mortality per 100,000 woman-years	**Mortality per 100,000 woman-years**	Mortality rate reduction attributable to adjuvant therapy in the absence of screening (%)	**Mortality per 100,000 woman-years**	Mortality rate reduction attributable to screening in the presence of adjuvant therapy (%)	Mortality rate reduction attributable to adjuvant therapy and screening (%)
Screening age						
50–74 (baseline)	**67.4**	**57.9**	–13.9%	**48.8**	–15.7%	–27.4%
48–74 (baseline + 1)				**48.4**	–16.5%	–28.1%
46–74 (baseline + 2)				**48.0**	–17.0%	–28.6%
44–74 (baseline + 3)				**47.6**	–17.7%	–29.2%
42–74 (baseline + 4)				**47.3**	–18.3%	–29.6%
40–74 (baseline + 5)				**47.2**	–18.6%	–29.9%
Annual 40–49; biennial 50–74 (baseline + 10)				**46.3**	–20.0%	–31.1%

screening rounds beginning at age 46 would reduce the mortality rate by 1.6%. If the current screening programme would be preceded by annual mammography between age 40 and 49, the breast cancer mortality rate in the total population could be reduced by an extra 5.1%. The overall reduction in the breast cancer mortality rate in this last scenario was estimated to be 31.1% (Table 6.1).

For women aged 40–49 years, 10 additional screening rounds starting at age 40 would reduce the breast cancer mortality rate from 25 to 22 deaths per 100,000 woman-years: a reduction of 12.9% compared to the present situation where screening starts at age 50 (which has no effect in this age group) (Figure 6.3a). In women aged 50–74 years, 10 additional screening examinations would reduce the breast cancer mortality by an extra 7.2%, to 62 breast cancer deaths per 100,000 woman-years (Figure 6.3b). Lowering the starting age of screening has no effect on breast cancer mortality in women aged 75–84 years.

Sensitivity analysis

The estimated effects of adjuvant systemic treatment and screening would drop by 8.4% if the effectiveness of adjuvant endocrine therapy was 25% lower than in the baseline

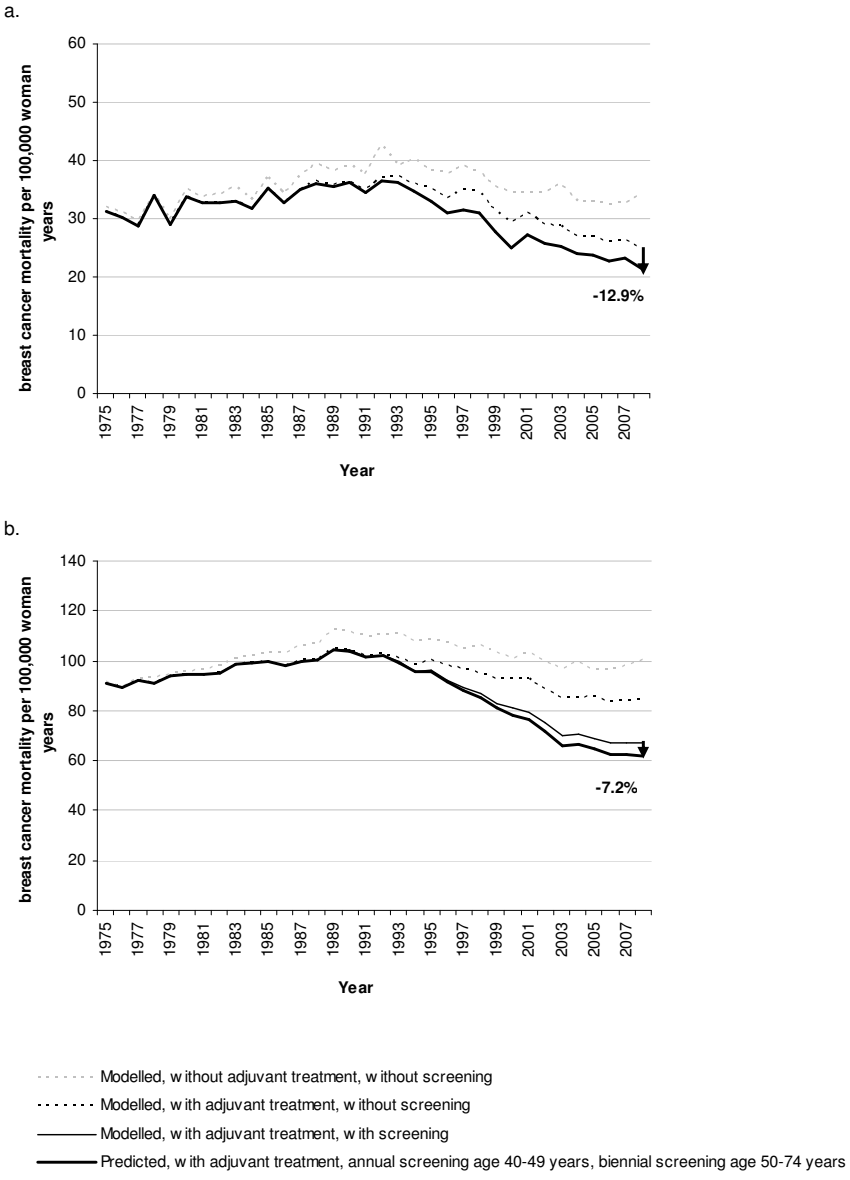


Figure 6.3a–b Predicted effects of 10 additional screening examinations on breast cancer mortality, in a) women aged 40–49 years, b) women aged 50–74 years, compared to biennial screening between age 50 and 74

model (Table 6.2). This could for instance occur when treatment adherence would be lower than in randomized controlled trials included in the EBCTCG analysis. The screening effect would remain approximately the same. The effects of adjuvant therapy and screening would maximally be 3.7% lower or higher than in the baseline model if the screening effectiveness would change (Table 6.2). Assumed trends in the survival of breast cancer only slightly affected the estimated effects of adjuvant treatment and screening.

Table 6.2 Sensitivity analysis

	Adjuvant treatment effect in the absence of screening	Screening effect in the presence of adjuvant therapy	Screening + adjuvant treatment effect
Screening scenario			
50–74, no improvement in breast cancer survival between 1990 and 2008	–13.0%	–16.5%	–27.4%
50–74, effectiveness of endocrine therapy 25% lower than baseline model	–10.9%	–15.9%	–25.1%
Annual 40–49, biennial 50–74, test sensitivity for women aged <50 same as that for women aged ≥50	–13.9%	–21.3%	–32.3%
Annual 40–49, biennial 50–74, effectiveness screening for women aged <50 is 25% lower than baseline model	–13.9%	–18.8%	–30.0%
Annual 40–49, biennial 50–74, effectiveness screening for women aged <50 is 25% higher than baseline model	–13.9%	–21.1%	–32.1%

6.4 DISCUSSION

Adjuvant systemic therapy was estimated to have reduced the breast cancer mortality rate in 2008 in the total simulated population by 13.9% compared to a situation without any intervention; biennial mammography screening between age 50 and 74 reduced the breast cancer mortality with an additional 15.7%. One additional screening round at age 48 would further reduce the breast cancer mortality rate by 1.0%; annual screening between age 40 and 49 followed by biennial screening between age 50 and 74 would reduce this rate by 5.1%.

Our results are in line with those of the Cancer Intervention and Surveillance Modeling Network (CISNET) that estimated that adjuvant therapy and screening contributed approximately equally to the observed mortality reduction in the USA.[12] In a previous analysis in the Netherlands, however, the estimated effect of screening in comparison to adjuvant therapy was higher than in the current study, with a mortality reduction of 28%–30% attributable to screening *versus* a mortality reduction of 7% attributable to adjuvant therapy.[28] This may be related to the lower estimates of the effectiveness of adjuvant therapy (based on the EBCTCG meta-analyses of 1998[29–30] instead of 2005). Also, at that time, adjuvant therapy was less frequently used, and the relatively effective[1] combination of endocrine and chemotherapy was not yet prescribed. Two recent studies reported minimal effects of screening. An analysis of mortality trends in Norway, for instance, showed that only one-third of the observed breast cancer mortality reduction of 10% could be attributed to screening.[5] Furthermore, a comparison of neighbouring countries with and without screening showed no differences in breast cancer mortality.[4] The absence of screening effects has been related to methodological issues in these studies.[31–32]

Because it is unknown how breast cancer would have progressed without any intervention, estimating the benefits of screening and adjuvant treatment in a population setting involves uncertainties. For instance, in this study we could only satisfactorily model the observed breast cancer mortality by assuming an improvement in breast cancer survival from unknown causes, in addition to the modelled screening and adjuvant systemic treatment effects. Such improvements may be attributed to further developments in adjuvant treatment or mammographic screening that were not yet accounted for in this analysis. For instance, the effectiveness of endocrine treatment may have increased, by the use of aromatase inhibitors for post-menopausal ER+ patients.[6] Screening may also have become more effective over the last decennium, for instance by the implementation of digital mammography.[8] It is, however, unlikely that screening is solely responsible for the unexplained fraction of the improvement in breast cancer survival in this study,

because breast cancer mortality in unscreened women also decreased more strongly than expected. Other factors, such as general advances in diagnostics and primary treatment for breast cancer, and increased awareness may have been important in reducing breast cancer as well.

The estimation of the effects of screening, especially those in women below age 50, entails uncertainties, because no statistically significant reduction in breast cancer mortality has been observed in trials that were specifically designed for this purpose. Based on the Swedish randomized controlled trials,[3, 23, 25–26] we predicted that one additional screening round at age 48 may lower the breast cancer mortality rate in the total population by 1.0% compared to screening from age 50. Ten extra screening rounds between age 40 and 49 would reduce the mortality rate by 5.1%. Our results were similar to those of Mandelblatt *et al.* (2009), who estimated that 5 additional screening rounds between age 50 and 40 could reduce breast cancer mortality by 3%.[13]

Besides potential benefits, screening at younger ages also involves certain risks. The number of radiation-induced breast cancers, for instance, is likely to become somewhat higher when the lower age limit for screening would be advanced.[34] The fraction of false positive test outcomes and, consequently, the number of additional imaging examinations may also increase.[35–36] If screening below age 50 is considered, these risks should be taken into account.

Conclusion

Adjuvant systemic therapy reduced breast cancer mortality by 13.9%; mammography screening additionally decreased mortality by 15.7%. Expanding the lower age limit of screening would further reduce breast cancer mortality.

REFERENCES

1. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687–717.
2. International Agency for Research on Cancer. Volume 7: Breast Cancer Screening. In: Vainio H, Bianchini F. IARC Handbooks of Cancer Preventioned. Lyon: IARCPress, 2002.
3. de Koning HJ, Boer R, Warmerdam PG, Beemsterboer PM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. Journal of the National Cancer Institute 1995;87:1217–23.
4. Autier P, Boniol M, La Vecchia C, Vatten L, Gavin A, Hery C, Heanue M. Disparities in breast cancer mortality trends between 30 European countries: retrospective trend analysis of WHO mortality database. BMJ (Clinical research ed 2010;341:c3620.
5. Kalager M, Zelen M, Langmark F, Adami HO. Effect of screening mammography on breast-cancer mortality in Norway. The New England journal of medicine 2010;363:1203–10.
6. van de Velde CJ, Verma S, van Nes JG, Masterman C, Pritchard KI. Switching from tamoxifen to aromatase inhibitors for adjuvant endocrine therapy in postmenopausal patients with early breast cancer. Cancer Treat Rev 2010;36:54–62.
7. National Evaluation Team for Breast cancer screening (NETB), National evaluation of breast cancer screening in the Netherlands 1990–2007. Twelfth evaluation report. Erasmus MC University Medical Center Rotterdam, Radboud University Nijmegen Medical Centre, 2009.
8. de Gelder R, Fracheboud J, Heijnsdijk EA, den Heeten G, Verbeek AL, Broeders MJ, Draisma G, de Koning HJ. Digital mammography screening: Weighing reduced mortality against increased overdiagnosis. Preventive medicine 2011.
9. Hery C, Ferlay J, Boniol M, Autier P. Quantification of changes in breast cancer incidence and mortality since 1990 in 35 countries with Caucasian-majority populations. Ann Oncol 2008;19:1187–94.
10. Botha JL, Bray F, Sankila R, Parkin DM. Breast cancer incidence and mortality trends in 16 European countries. Eur J Cancer 2003;39:1718–29.
11. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet 2006;368:2053–60.
12. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, Mandelblatt JS, Yakovlev AY, Habbema JD, Feuer EJ. Effect of screening and adjuvant therapy on mortality from breast cancer. The New England journal of medicine 2005;353:1784–92.
13. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, Huang H, Lee SJ, Munsell M, Plevritis SK, Ravdin P, Schechter CB, *et al.* Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Intern Med 2009;151:738–47.
14. Habbema JD, van Oortmarssen GJ, Lubbe JT, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. Computer methods and programs in biomedicine 1985;20:79–93.
15. van Oortmarssen GJ, Habbema JD, van der Maas PJ, de Koning HJ, Collette HJ, Verbeek AL, Geerts AT, Lubbe KT. A model for breast cancer screening. Cancer 1990;66:1601–12.
16. IKCnet. http://www.ikcnet.nl/.

17. Sukel MP, van de Poll-Franse LV, Nieuwenhuijzen GA, Vreugdenhil G, Herings RM, Coebergh JW, Voogd AC. Substantial increase in the use of adjuvant systemic treatment for early stage breast cancer reflects changes in guidelines in the period 1990–2006 in the southeastern Netherlands. Eur J Cancer 2008;44:1846–54.
18. Struikmans H, Nortier JW, Rutgers EJ, Zonderland HM, Bontenbal M, Elkhuizen PH, van Tienhoven G, Tjan-Heijnen VC, van Vegchel T, Tuut MK, Benraadt J. [Guideline 'Treatment of breast cancer 2008' (revision)]. Nederlands tijdschrift voor geneeskunde 2008;152:2507–11.
19. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. Cancer 1989;63:181–7.
20. Michaelson JS, Silverstein M, Sgroi D, Cheongsiatmoy JA, Taghian A, Powell S, Hughes K, Comegno A, Tanabe KK, Smith B. The effect of tumor size and lymph node status on breast carcinoma lethality. Cancer 2003;98:2133–43.
21. Michaelson JS, Silverstein M, Wyatt J, Weber G, Moore R, Halpern E, Kopans DB, Hughes K. Predicting the survival of patients with breast carcinoma using tumor size. Cancer 2002;95:713–23.
22. Sant M, Allemani C, Capocaccia R, Hakulinen T, Aareleid T, Coebergh JW, Coleman MP, Grosclaude P, Martinez C, Bell J, Youngson J, Berrino F. Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. International journal of cancer 2003;106:416–22.
23. Tabar L, Vitak B, Chen HH, Duffy SW, Yen MF, Chiang CF, Krusemo UB, Tot T, Smith RA. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. Radiol Clin North Am 2000;38:625–51.
24. Adami HO, Malker B, Holmberg L, Persson I, Stone B. The relation between survival and age at diagnosis in breast cancer. The New England journal of medicine 1986;315:559–63.
25. Nystrom 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;359:909–19.
26. Bjurstam N, Bjorneld L, Warwick J, Sala E, Duffy SW, Nystrom L, Walker N, Cahlin E, Eriksson O, Hafstrom LO, Lingaas H, Mattsson J, *et al.* The Gothenburg Breast Screening Trial. Cancer 2003;97:2387–96.
27. van Herk-Sukel MP, van de Poll-Franse LV, Voogd AC, Nieuwenhuijzen GA, Coebergh JW, Herings RM. Half of breast cancer patients discontinue tamoxifen and any endocrine treatment before the end of the recommended treatment period of 5 years: a population-based analysis. Breast cancer research and treatment 2010;122:843–51.
28. Vervoort MM, Draisma G, Fracheboud J, van de Poll-Franse LV, de Koning HJ. Trends in the usage of adjuvant systemic therapy for breast cancer in the Netherlands and its effect on mortality. British journal of cancer 2004;91:242–7.
29. Polychemotherapy for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. Lancet 1998;352:930–42.
30. Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. Lancet 1998;351:1451–67.
31. De Koning HJ. Screen evaluation without evaluating screening? Rapid response to Autier *et al.*: Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. BMJ (Clinical research ed 2011;343:d4411.

32. van Ravesteyn NT, Heijnsdijk EA, de Koning HJ. More on screening mammography. The New England journal of medicine 2011;364:282–3; author reply 5–6.
33. de Gelder R, Draisma G, Heijnsdijk EA, de Koning HJ. Population-based mammography screening below age 50: balancing radiation-induced *vs.* prevented breast cancer deaths. British journal of cancer 2011;104:1214–20.
34. Frisell J, Lidbrink E. The Stockholm Mammographic Screening Trial: Risks and benefits in age group 40–49 years. J Natl Cancer Inst Monogr 1997:49–51.
35. Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L, Force USPST. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. Ann Intern Med 2009;151:727–37, W237–42.

Appendix 6.1 Model parameters on natural history of breast cancer, and survival after adjuvant treatment and screening

DCIS	%
Fraction of tumours with a preclinical screen-detectable DCIS stage (%)	18
Fraction of DCIS that progress to preclinical invasive cancer in the absence of screening (%)	11
Fraction of DCIS that will be clinically diagnosed in the absence of screening (%)	5
Fraction of DCIS that regress in the absence of screening (%)	2

Mean duration (years) of screen-detectable preclinical phase by age and stage	Stage								
Age	DCIS	T1AN–	T1AN+	T1BN–	T1BN+	T1CN–	T1CN+	T2+N–	T2+N+
40	2.1	0.1	0.1	0.3	0.3	0.6	0.6	0.7	0.7
50	2.1	0.1	0.1	0.4	0.4	0.8	0.8	0.9	0.9
60	2.1	0.1	0.1	0.5	0.5	1.2	1.2	1.3	1.3
70	2.1	0.2	0.2	0.7	0.7	1.5	1.5	1.6	1.6

Sensitivity of mammography by age and preclinical stage	Stage								
Age	DCIS	T1AN–	T1AN+	T1BN–	T1BN+	T1CN–	T1CN+	T2+N–	T2+N+
<50	58%	39%	39%	47%	47%	68%	68%	71%	71%
≥50	77%	52%	52%	62%	62%	90%	90%	95%	95%

Long-term relative survival by clinical stage and age, without adjuvant treatment	Stage								
Age	DCIS	T1AN–	T1AN+	T1BN–	T1BN+	T1CN–	T1CN+	T2+N–	T2+N+
≤30	1.000	0.761	0.510	0.696	0.408	0.557	0.236	0.310	0.0555
40	1.000	0.798	0.575	0.741	0.481	0.618	0.310	0.386	0.102
50	1.000	0.815	0.605	0.762	0.512	0.646	0.341	0.418	0.118
60	1.000	0.796	0.568	0.738	0.472	0.612	0.298	0.375	0.0885
70	1.000	0.737	0.476	0.667	0.376	0.524	0.213	0.282	0.0524
≥80	1.000	0.678	0.383	0.597	0.279	0.435	0.128	0.189	0.0163

Long-term relative survival by clinical stage and age, with hormonal treatment	Stage								
Age	DCIS	T1AN–	T1AN+	T1BN–	T1BN+	T1CN–	T1CN+	T2+N–	T2+N+
≤30	1.000	0.854	0.701	0.814	0.639	0.730	0.534	0.579	0.424
40	1.000	0.865	0.714	0.826	0.650	0.743	0.533	0.585	0.388
50	1.000	0.860	0.699	0.819	0.629	0.731	0.499	0.558	0.330
60	1.000	0.856	0.696	0.815	0.628	0.727	0.505	0.559	0.357

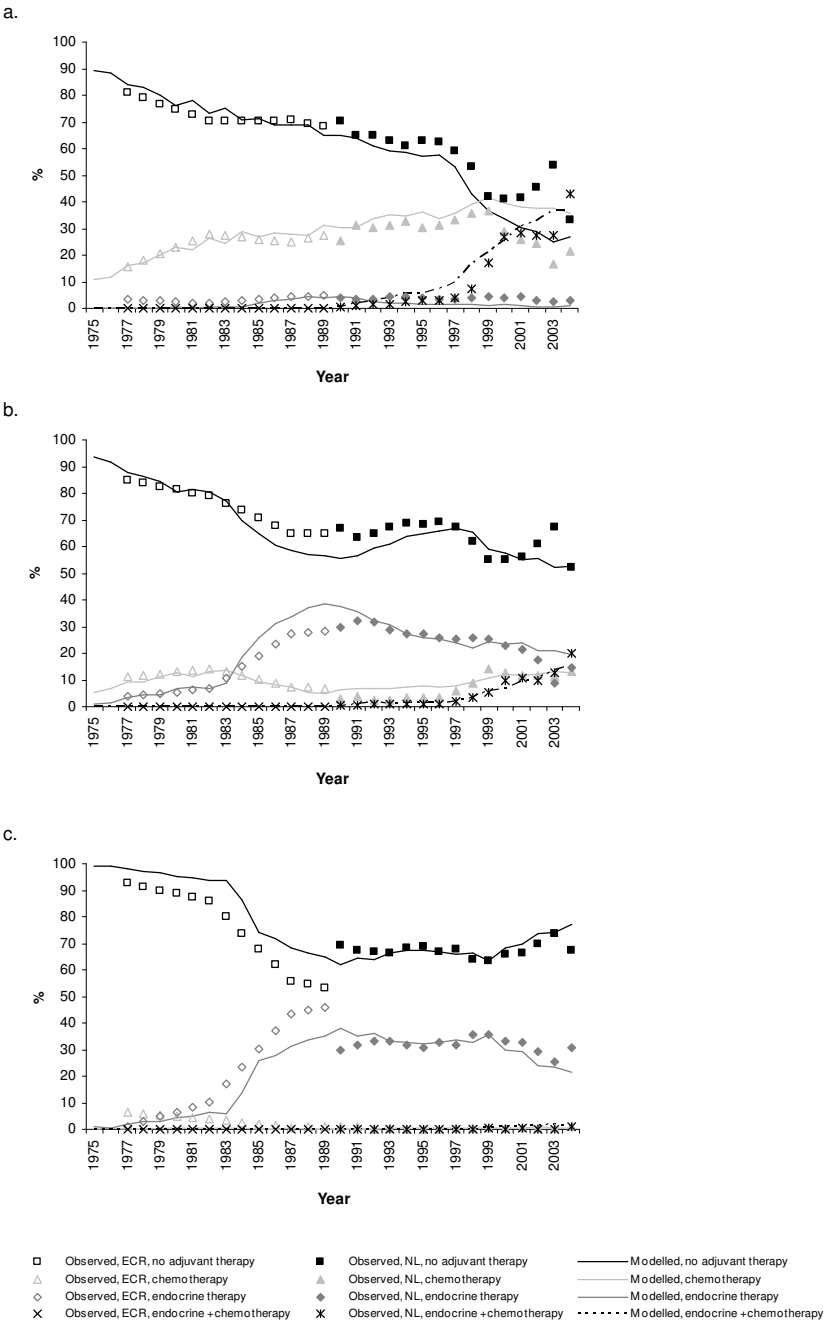
70	1.000	0.832	0.666	0.788	0.601	0.696	0.497	0.541	0.394
≥80	1.000	0.797	0.612	0.746	0.546	0.644	0.451	0.489	0.380

Long-term relative survival by clinical stage and age, with chemotherapy	Stage								
Age	DCIS	T1AN–	T1AN+	T1BN–	T1BN+	T1CN–	T1CN+	T2+N–	T2+N+
≤30	1.000	0.831	0.652	0.784	0.580	0.686	0.458	0.510	0.329
40	1.000	0.858	0.700	0.817	0.634	0.731	0.513	0.567	0.366
50	1.000	0.855	0.691	0.814	0.619	0.723	0.486	0.546	0.314
60	1.000	0.820	0.620	0.769	0.535	0.659	0.382	0.450	0.198
70	1.000	0.767	0.535	0.705	0.446	0.578	0.301	0.363	0.157
≥80	1.000	0.720	0.464	0.649	0.373	0.509	0.241	0.294	0.144

Long-term relative survival by clinical stage and age, with hormonal and chemotherapy	Stage								
Age	DCIS	T1AN–	T1AN+	T1BN–	T1BN+	T1CN–	T1CN+	T2+N–	T2+N+
≤30	1.000	0.905	0.806	0.879	0.765	0.824	0.697	0.726	0.626
40	1.000	0.912	0.814	0.887	0.772	0.833	0.697	0.730	0.602
50	1.000	0.890	0.765	0.859	0.711	0.790	0.611	0.655	0.481
60	1.000	0.872	0.729	0.835	0.669	0.757	0.559	0.608	0.428
70	1.000	0.851	0.702	0.811	0.645	0.730	0.552	0.592	0.461
≥80	1.000	0.820	0.654	0.774	0.596	0.683	0.511	0.545	0.448

Reduction in risk of dying of breast cancer by age and preclinical stage after screen-detection	Stage								
Age	DCIS	T1AN–	T1AN+	T1BN–	T1BN+	T1CN–	T1CN+	T2+N–	T2+N+
≤30	100%	85%	67%	81%	59%	71%	43%	50%	18%
40	100%	88%	72%	84%	65%	75%	50%	57%	25%
50	100%	89%	74%	85%	67%	77%	53%	60%	28%
60	100%	87%	72%	84%	64%	75%	49%	56%	24%
70	100%	88%	72%	84%	65%	75%	50%	57%	25%
≥80	100%	88%	73%	84%	65%	76%	50%	57%	25%

Abbreviations: ductal carcinoma *in situ* (DCIS); lymph node negative breast cancer with diameter of 5 mm or smaller (T1AN–); lymph node negative breast cancer with diameter of 6–10 mm (T1BN–); lymph node negative breast cancer with diameter of 11–20 mm (T1CN–); lymph node negative breast cancer with diameter of more than 20 mm (T2+N–); lymph node positive breast cancer with diameter of 5 mm or smaller (T1AN+); lymph node positive breast cancer with diameter of 6–10 mm (T1BN+); lymph node positive breast cancer with diameter of 11–20 mm (T1CN+); lymph node positive breast cancer with diameter of more than 20 mm (T2+N+)



Appendix 6.2a–c The probability of using adjuvant endocrine therapy, chemotherapy, a combination of endocrine and chemotherapy or no adjuvant therapy among women younger than 50 (a), 50–69 (b) or 70 and older (c)

Abbreviations: Eindhoven Cancer Registry (ECR); the Netherlands (NL)

Population-based mammography screening below age 50: balancing radiation-induced *vs.* prevented breast cancer deaths

Rianne de Gelder, Gerrit Draisma, Eveline A. M. Heijnsdijk, Harry J. de Koning

Br J Cancer 2011 Mar 29;104(7):1214-20. Epub 2011 Mar 1.

ABSTRACT

Introduction

Exposure to ionizing radiation at mammography screening may cause breast cancer. Because the radiation risk increases with lower exposure age, advancing the lower age limit may affect the balance between screening benefits and risks. The present study explores the benefit–risk ratio of screening before age 50.

Methods

The benefits of biennial mammography screening, starting at various ages between 40 and 50, and continuing up to age 74 were examined using micro-simulation. In contrast with previous studies that commonly used excess relative risk models, we assessed the radiation risks using the latest BEIR-VII excess absolute rate exposure-risk model.

Results

The estimated radiation risk is lower than previously assessed. At a mean glandular dose of 1.3 mGy per view that was recently measured in the Netherlands, biennial mammography screening between age 50 and 74 was predicted to induce 1.6 breast cancer deaths per 100,000 women aged 0–100 (range 1.3–6.3 extra deaths at a glandular dose of 1–5 mGy per view), against 1121 avoided deaths in this population. Advancing the lower age limit for screening to include women aged 40–74 was predicted to induce 3.7 breast cancer deaths per 100,000 women aged 0–100 (range 2.9–14.4) at biennial screening, but would also prevent 1302 deaths.

Conclusion

The benefits of mammography screening between age 40 and 74 were predicted to outweigh the radiation risks.

7.1 INTRODUCTION

Each year, millions of women in Europe have a screening mammogram. This may reduce the risk of dying from breast cancer by up to 50%[1–3]. It may, on the other hand, also cause breast cancer and breast cancer deaths due to ionizing X-ray radiation. It has been shown that the risk of tumour induction is proportional to the dose of radiation absorbed in the breast[4–9]. Although radiation doses at mammography are much lower than the doses for which cancer induction is directly observed,[10] screening a large population on a regular basis has the potential to harm.

Most nevertheless agree that the benefits of screening outweigh its radiation risks; in particular for women aged 50–69. For annual screening between age 40 and 49, the risks may also be justified, provided that a mortality reduction of at least 20% is obtained and the dose is sufficiently small.[11–12] The average denser breasts and faster growing tumours in this age group, however, may substantially reduce the screening effectiveness compared with older women. At the same time, the radiation risks increase with younger exposure age.[10] Recently, the UK Age Trial showed a non-significant 24% breast cancer mortality reduction in women aged 40–48 that were annually screened.[13] Based on these findings, we compared radiation risks to the effects of breast cancer screening, starting at various screening ages between 40 and 50 and continuing up to age 74.

For our analyses, the most recent exposure-risk model is used: the BEIR-VII model.[14] This model differs from previous models with regard to the shape of the dose–response relation: at very low (mammography) doses, the model is adjusted with a dose and dose-rate effectiveness factor (DDREF). It is an additive instead of a multiplicative model, which until recently had been the standard model for risk estimations. We will calculate radiation risks using new estimates for the average glandular dose that were measured in the Dutch nationwide screening programme,[15] and explore the threshold for a positive benefit–risk balance.

7.2 MATERIALS AND METHODS

Two models were used to estimate the ratio of screening benefits *vs.* radiation risks: the micro-simulation analysis model MISCAN[16–17] that estimates the benefits of mammography screening, and the radiation risk model of the 7th Biological Effects of Ionizing Radiation committee (BEIR-VII).[14]

MISCAN simulates the natural course of breast cancer in the absence of screening: from its early onset to, eventually, death from breast cancer or other causes. The model also assesses the impact of screening on the natural history. In short, MISCAN is a Markov-like stage transition model, in which a lesion progresses through the successive preclinical invasive TNM stages T1a, T1b, T1c and T2+. Preclinical invasive cancer may or may not be preceded by preclinical DCIS. In the absence of screening, a lesion can grow from one preclinical stage to the next or become symptomatic. With screening, a preclinical lesion can also be screen detected. The onset rate, transition probabilities between various tumour stages, the stage durations and probability of screen detection have been estimated using observations from the Dutch cancer registry and the Dutch breast cancer screening programme.[18] After a breast cancer diagnosis, some women may be cured whereas others may not. The age- and stage-specific probability of cure and the survival after diagnosis were based on several international sources.[19–24] The improvement in survival after screen detection was based on the Swedish breast cancer screening trials.[16, 24–26] Model parameters and the procedures of parameter estimation have been published elsewhere.[17] Model outcomes were predicted breast cancer incidence and mortality without screening, and predicted breast cancer mortality with screening. From this, the predicted number of prevented deaths could be derived.

The risk of developing breast cancer due to screening exposure was estimated using the excess absolute rate (EAR) model by the BEIR-VII committee,[14] who adopted and reparameterised the 'pooled analysis' EAR model by Preston *et al* (2002)[10]. The model is described as follows:

$$\lambda(t, d, E) = \lambda(t, 0) + \Sigma_i \varepsilon(t, d_i, E_i)$$

and

$$\varepsilon(t\ d, E) = d * 148 * \exp(-0.05\ (E - 30) + 3.5 \ln(t/60)) \quad \text{if } t < 50$$
$$\varepsilon(t\ d, E) = d * 94 * \exp(-0.05\ (E - 30) + \ln(t/60)) \quad \text{if } t \geq 50$$

The incidence $\lambda(t, d, E)$ per 100,000 woman-years is equal to the predicted baseline incidence without radiation $\lambda(t, 0)$ plus the sum of all induced breast cancers due to radiation $\Sigma_i \varepsilon(t, d_i, E_i)$ at each screening round. In the equation, d is the glandular dose (mGy), E is the exposure age and t is the attained age. The risk of radiation-induced breast cancer increases with younger exposure age and higher attained age. After age 50, the risk increases less steeply than before this age, which is probably related to hormonal changes around the menopause. The dose–response coefficient (148 before age 50, 94 from age 50) is slightly different from that in the BEIR-VII model, because an error was made in

the parameterisation of the original Preston model (Preston, personal communication). The lifetime risk of breast cancer in a situation with mammography was calculated by multiplying the incidence at a given age with the survival at that age and cumulating the products:

$$I_d = \Sigma_t \lambda (t, d_1, d_2, ..., E_1, E_2 ...) * S(t) \qquad t = 0 ... 100$$

The induced breast cancer mortality was calculated by multiplying the breast cancer incidence at a given age with the survival and case fatality (p(t)) at that age, and cumulating the products for all ages:

$$M_d = \Sigma_t \lambda (t, d_1, d_2, ..., E_1, E_2 ...) * p(t) * S(t) \qquad t = 0 ... 100$$

Age-specific case fatality was derived from MISCAN. Survival was calculated using a recent life table for Dutch females. We assumed that cancers that were induced by radiation could become screen detected, at a similar rate as non-induced breast cancers.

Repeated exposure to low doses such as observed at mammography screening are considered to be less harmful than the high doses that were observed in atomic bomb survivors and women that were exposed for diagnostics or therapy. Therefore, it has been suggested that the predicted number of induced breast cancers and breast cancer deaths at mammography screening should be divided by a correction factor (the DDREF).[27–31] In our paper, we used a DDREF of 1.5, as suggested by the BEIR-VII Committee.[14]

Radiation risks and screening benefits were calculated for a biennial screening program targeting women aged 50–74. The impact of extending the screening programme to women younger than 50 years of age was also assessed, by calculating risks and benefits for various starting ages between age 40 and 50. The average absorbed glandular dose at mammography screening was 1.3 mGy per view (of both breasts), as shown by Zoetelief *et al* (2006) among women from four regional breast cancer screening units in the Netherlands.[15] Two-view mammography is performed at first screening rounds, while at subsequent rounds a single view is made. Assuming that women attend all screening rounds, the total glandular dose would thus be 14 * 1.3 = 18.2 mGy at biennial screening between age 50 and 74. Because regional variations of between 1.04 and 1.63 mGy per view and a maximum dose of 5 mGy per view (in <1% of all women) were observed,[15] we also estimated the radiation risks for a glandular dose of 1, 2 and 5 mGy per view. For women between age 40 and 49, we assumed that the same number of views were made as for women aged 50–74. The test sensitivity was expected to be 25% lower than that of older women.[32] We hypothesised that breast cancer mortality could be reduced

by 24% by annual mammography in screened women of this age group, similar to that observed in the UK Age Trial.[13]

Several sensitivity analyses were performed. These include a calculation of the radiation risk:

- using the BEIR-V and BEIR-VII excess relative risk model,[14, 33]
- assuming that induced cancers could not be detected by screening,
- assuming no correction for DDREF,
- assuming a 10-year latency time,
- assuming annual screening for women aged 40–49, and biennial screening for 50- to 74-year-old women,
- assuming that the test sensitivity for women aged 40–49 is 50% lower than or similar to that for older women,
- assuming that the effectiveness of screening for women aged 40–49 is 25% lower or 25% higher than that our baseline estimate.

All effects were calculated for a cohort of 100,000 women aged 0–100, measured over their entire lifespan. To separate the consequences of radiation at younger ages from those at older ages, we also analysed the effects and radiation risks of a decade of annual screening starting at age 30 and 40. In all scenarios, a 100% participation rate was assumed. Besides screening effects, no other time-dependent changes in breast cancer incidence and mortality were assumed.

7.3 RESULTS

Without screening, a total of 12 289 breast cancers were predicted to be diagnosed in a population of 100,000 women aged 0–100 (Table 7.1). The radiation that is absorbed at biennial screening between age 50 and 74 at a glandular dose of 1.3 mGy per view would induce 7.7 extra breast cancers. The ratio of baseline incidence without screening *vs.* induced incidence would be 1596 : 1, meaning that per 1596 breast cancers, 1 would be caused by screening. At a glandular dose of 5 mGy per view, the predicted number of induced breast cancers was 29.6 per 100,000 women aged 0–100, and the ratio of baseline to induced cancers 415 : 1. Extension of the lower age limit for screening would further increase the number of induced cancers, up to 17.1 if women aged 40–74 are screened biennially (1.3 mGy per view). In this scenario, the ratio baseline : induced incidence would be 720 : 1, ranging between 936 : 1 and 187 : 1 if the glandular dose would be 1 and 5 mGy per view, respectively.

Table 7.1 Induced breast cancer incidence at various radiation doses and screening scenarios[a]

Radiation dose (mGy) per view	1	1.3	2	5
Baseline incidence per 100,000 women age 0–100, predicted without screening	12,289	12,289	12,289	12,289
Biennial screening ages 50–74				
Induced incidence per 100,000 women age 0–100	5.9	7.7	11.8	29.6
Baseline incidence: Induced incidence	2075 : 1	1596 : 1	1037 : 1	415 : 1
Biennial screening ages 48–74				
Induced incidence per 100,000 women age 0–100	7.0	9.1	14.1	35.2
Baseline incidence: Induced incidence	1747 : 1	1344 : 1	874 : 1	349 : 1
Biennial screening ages 46–74				
Induced incidence per 100,000 women age 0–100	8.3	10.8	16.6	41.5
Baseline incidence: Induced incidence	1482 : 1	1140 : 1	741 : 1	296 : 1
Biennial screening ages 44–74				
Induced incidence per 100,000 women age 0–100	9.7	12.6	19.4	48.6
Baseline incidence: Induced incidence	1264 : 1	973 : 1	632 : 1	253 : 1
Biennial screening ages 42–74				
Induced incidence per 100,000 women age 0–100	11.3	14.7	22.7	56.6
Baseline incidence: Induced incidence	1085 : 1	835 : 1	543 : 1	217 : 1
Biennial screening ages 40–74				
Induced incidence per 100,000 women age 0–100	13.1	17.1	26.3	65.6
Baseline incidence: Induced incidence	936 : 1	720 : 1	468 : 1	187 : 1

Abbreviation: BEIR-VII, 7th Biological Effects of Ionizing Radiation committee. [a]As calculated with the BEIR-VII excess absolute rate (EAR) model corrected with a 'dose and dose-rate effectiveness factor' (DDREF) of 1.5, assuming no latency time and a potential screening benefit for induced breast cancers. For women below the age of 50, the screening effectiveness was assumed to be comparable to that in the UK Age Trial[13] and the test sensitivity was estimated to be 25% lower than that for women older than 50 years of age. The grey-shaded column represents the baseline and induced breast cancer incidence at the average observed glandular dose (1.3 mGy per view) in the Netherlands

The predicted number of breast cancer deaths in a situation without screening was 4330 per 100,000 women (Table 7.2). Approximately 26% of those (1121) could be prevented by biennial screening in the age group 50–74. On the other hand, screening would cause 1.6 extra breast cancer deaths per 100,000 women, assuming a glandular dose of 1.3 mGy per view. The ratio of baseline mortality without screening *vs.* induced deaths would be 2641 : 1. Weighing the number of prevented deaths against the number of induced deaths, this ratio would be 684 : 1. A glandular dose of 5 mGy per view would increase the number of induced breast cancer deaths to 6.3, which would result in a ratio of baseline mortality *vs.* induced deaths of 687 : 1, and a ratio of prevented *vs.* induced

Table 7.2 Induced breast cancer mortality at various radiation doses and screening scenarios, calculated for 100,000 women aged 0–100[a]

Radiation dose (mGy) per view	1	1.3	2	5
Baseline mortality per 100,000 women age 0–100, predicted without screening	4330	4330	4330	4330
Biennial screening ages 50–74				
Prevented mortality (MISCAN, screening – no screening)	1121	1121	1121	1121
Induced mortality per 100,000 women age 0–100	1.3	1.6	2.5	6.3
Baseline mortality: Induced mortality	3434 : 1	2641 : 1	1717 : 1	687 : 1
Prevented mortality: Induced mortality	889 : 1	684 : 1	445 : 1	178 : 1
Biennial screening ages 48–74				
Prevented mortality (MISCAN, screening – no screening)	1172	1172	1172	1172
Induced mortality per 100,000 women age 0–100	1.5	2.0	3.0	7.5
Baseline mortality: Induced mortality	2876 : 1	2212 : 1	1438 : 1	575 : 1
Prevented mortality: Induced mortality	779 : 1	599 : 1	389 : 1	156 : 1
Biennial screening ages 46–74				
Prevented mortality (MISCAN, screening – no screening)	1216	1216	1216	1216
Induced mortality per 100,000 women age 0–100	1.8	2.3	3.6	8.9
Baseline mortality: Induced mortality	2427 : 1	1867 : 1	1213 : 1	485 : 1
Prevented mortality: Induced mortality	681 : 1	524 : 1	341 : 1	136 : 1
Biennial screening ages 44–74				
Prevented mortality (MISCAN, screening – no screening)	1247	1247	1247	1247
Induced mortality per 100,000 women age 0–100	2.1	2.7	4.2	10.5
Baseline mortality: Induced mortality	2057 : 1	1582 : 1	1028 : 1	411 : 1
Prevented mortality: Induced mortality	592 : 1	455 : 1	296 : 1	118 : 1
Biennial screening ages 42–74				
Prevented mortality (MISCAN, screening – no screening)	1275	1275	1275	1275
Induced mortality per 100,000 women age 0–100	2.5	3.2	4.9	12.3
Baseline mortality: Induced mortality	1757 : 1	1351 : 1	878 : 1	351 : 1
Prevented mortality: Induced mortality	517 : 1	398 : 1	259 : 1	103 : 1
Biennial screening ages 40–74				
Prevented mortality (MISCAN, screening – no screening)	1302	1302	1302	1302
Induced mortality per 100,000 women age 0–100	2.9	3.7	5.7	14.4
Baseline mortality: Induced mortality	1508 : 1	1160 : 1	754 : 1	302 : 1
Prevented mortality: Induced mortality	453 : 1	349 : 1	227 : 1	91 : 1

Abbreviations**:** BEIR-VII, 7th Biological Effects of Ionizing Radiation committee; MISCAN = micro-simulation screening analysis. [a]As calculated with the BEIR-VII excess absolute rate (EAR) model corrected with a dose and dose-rate effectiveness factor (DDREF) of 1.5, assuming no latency time and a potential screening benefit for induced breast cancers. For women below the age of 50, the screening effectiveness was assumed to be comparable to that in the UK Age Trial[13] and the test sensitivity was estimated to be 25% lower than that for women older than 50 years of age. The grey-shaded column represents the baseline and induced breast cancer mortality at the average observed glandular dose (1.3 mGy per view) in the Netherlands

deaths of 178 : 1. Screening from a younger age would further increase the number of induced deaths. Biennial screening between age 40 and 74, for instance, would cause 3.7 extra breast cancer deaths at 1.3 mGy per view. The benefit–risk ratio would then be 349 prevented *vs.* 1 induced death. Increasing the glandular dose to 5 mGy per view would cause 14.4 extra breast cancers deaths, resulting in the least favourable benefit–risk ratio: 91 prevented *vs.* 1 induced breast cancer death.

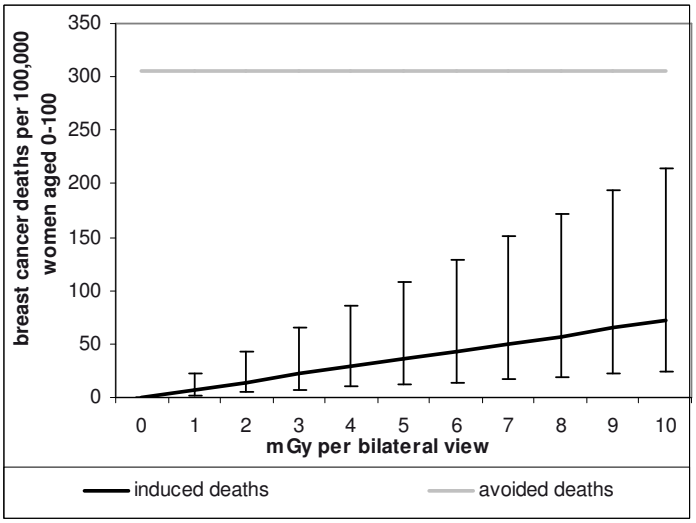
The balance between screening effects and radiation risks is also dependent on the assumed screening benefit for induced breast cancers, the assumed latency time and the shape of the dose- response; that is, whether the model is corrected for a DDREF or not (Table 7.3). Using the BEIR-V ERR model[33] instead of the BEIR-VII EAR model had no substantial effect on radiation risks, but the BEIR-VII ERR model resulted in substantially

Table 7.3 Induced and prevented breast cancer deaths and the benefit–risk ratio of breast cancer screening under various model assumptions, calculated for a glandular dose of 1.3 mGy per view, for 100 000 women aged 0–100

		Induced mortality	Prevented mortality	Ratio prevented: Induced mortality
Biennial screening ages 50–74[a]		1.6	1121	684 : 1
Biennial screening ages 50–74	Latency period of 10 years	1.4	1121	805 : 1
Biennial screening ages 50–74	BEIR V ERR	1.7	1121	658 : 1
Biennial screening ages 50–74	BEIR VII ERR	2.7	1121	419 : 1
Biennial screening ages 50–74	no screening benefit for induced cancers	2.4	1121	462 : 1
Biennial screening ages 50–74	no correction for DDREF	2.5	1121	464 : 1
Biennial screening ages 40–74[a]		3.7	1302	349 : 1
Biennial screening ages 40–74	Screening effectiveness 50- –25%	3.8	1256	334 : 1
Biennial screening ages 40–74	Screening effectiveness 50- +25%	3.7	1342	362 : 1
Biennial screening ages 40–74	Sensitivity 50- 50% lower than 50+	3.8	1225	323 : 1
Biennial screening ages 40–74	Sensitivity 50- same as 50+	3.7	1371	373 : 1
Annual screening ages 40–49 and biennial screening ages 50–74[a]		5.4	1392	259 : 1

Abbreviation: BEIR-VII, 7th Biological Effects of Ionizing Radiation committee. [a]As calculated with the BEIR-VII excess absolute rate (EAR) model corrected with a 'dose and dose-rate effectiveness factor' (DDREF) of 1.5, assuming no latency time and a potential screening benefit for induced breast cancers. For women below the age of 50, the screening effectiveness was assumed to be comparable to that in the UK Age Trial[13] and the test sensitivity was estimated to be 25% lower than that for women older than 50 years of age. The grey-shaded rows represent the baseline scenarios.

a.



b.

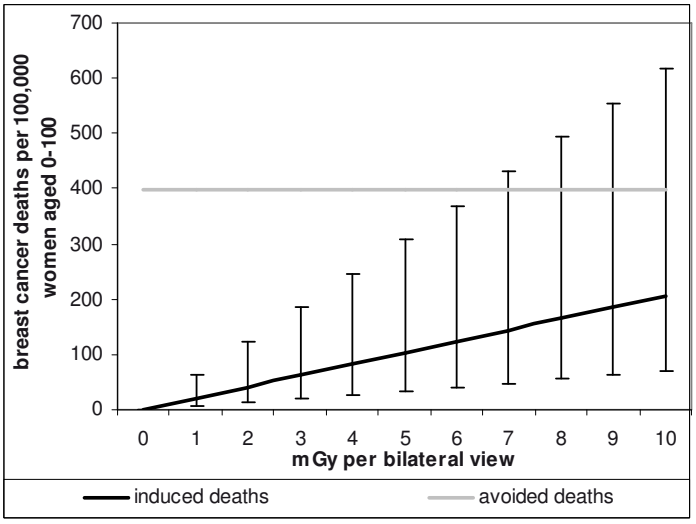


Figure 7.1a–b Prevented and induced breast cancer deaths at a decade of screening starting at age 40 (a) or age 30 (b), calculated for 100,000 women aged 0–100. Calculations were based on the BEIR-VII excess absolute rate (EAR) model, assuming no latency time. The test sensitivity for women younger than 50 was assumed to be 25% lower than that for women older than 50, and the screening effectiveness was comparable to that in the UK Age Trial[13]. For both situations, no 'dose and dose-rate effectiveness factor' (DDREF) correction was applied. Vertical lines represent an uncertainty interval around the estimated number of induced breast cancer deaths of a factor 3.

higher risk estimates. The number of induced breast cancer deaths is hardly affected by the assumed screening effectiveness or test sensitivity for women below age 50. The number of prevented breast cancer deaths, however, increases slightly, and is highest when the sensitivity is assumed to be the same for women younger and older than 50. Most breast cancer deaths can be prevented in a situation with annual screening between the ages 40 to 49, and biennial screening between age 50 and 74. However, the extra screening examinations would also result in the highest number of induced cancer deaths: 5.4 per 100,000 women. This resulted in the least favourable benefit-risk balance: per 259 prevented deaths, one breast cancer would be induced by screening.

Focussing on the first decade of screening only, annual screening from age 40 in general would avoid more deaths than it induces. Even in the unlikely situation that the average radiation dose would exceed 10 mGy per view and the model has underestimated the number of induced breast cancer deaths by a factor of 3, the radiation risks of screening would not outweigh the benefits (Figure 7.1a). Annual screening from age 30 would induce more deaths than it prevents if the average dose would be 7 mGy per view or more, and the radiation risks would be underestimated by a factor of 3 (Figure 7.1b). However, screening in this age group in the Netherlands is only recommended when women are at high risk for breast cancer, for instance because of an inherited BRCA1 or BRCA2 gene mutation.

7.4 DISCUSSION

Our study demonstrated that the risk of radiation-induced breast cancer due to mammography screening is small. Biennial screening between age 50 and 74 was predicted to cause 7.7 breast cancers and 1.6 breast cancer deaths per 100,000 women aged 0–100, but would also prevent 1121 breast cancer deaths. This indicates that the radiation risks of regular mammography are negligible.

Compared with previous estimates of the radiation risk at mammography screening of 4–23 excess breast cancers[12, 34–35] and 2–11 extra breast cancer deaths per 100,000 women,[11–12, 36] our predictions are relatively small. This may be related to the mean glandular dose in the Netherlands of 1.3 mGy per view, which is smaller than the average dose of between 1.78 and 2.35 mGy previously found in the Netherlands,[36] or the dose of between 1.5 and 2.4 mGy per view observed in other countries.[12, 34, 37–38] Furthermore, calculations were based on one-view mammography at subsequent screening rounds, but in practice, two-view mammography is increasingly performed. If we would assume a second view in ~50% of all subsequent screening rounds, as estimated by Duijm *et*

al (2009), the mean glandular dose would increase to 2 mGy per examination and the number of induced breast cancer deaths to 2.5 per 100,000 women.[39] Several screening programmes routinely use two views,[40] which would double the radiation risks compared with our estimates at 1.3 mGy per view.

Our estimates are also lower because a DDREF correction of 1.5 was used to estimate the radiation risk at the low doses absorbed at mammography screening.[14] It was further assumed that all women with radiation-induced cancers could profit from the screening programme. In reality, this will not entirely be the case, because some cancers will become clinically diagnosed in the interval between two screening rounds, or after women reach the upper age limit for screening.

Radiation risk estimates involve many uncertainties. Previous calculations of the risk of mammography screening have frequently been based on an ERR model,[11–12, 36] but in our study the EAR model was used, following the recommendations by BEIR-VII and Preston *et al* (2002).[10, 14] Our sensitivity analysis showed that the outcomes of the EAR model were comparable to those of the BEIR-V ERR model but lower than the BEIR-VII ERR model. The main conclusions did not differ. Second, different types of radiation may vary in the harm they may cause. The model that was used in this study was based on Japanese atomic bomb survivors who were exposed to γ-rays and neutrons, and medically exposed women who received high-energy X-rays or Ra-226 γ-ray radiation for diagnostics or therapy. Mammography, on the other hand, involves low-energy X-ray radiation, which may be more biologically effective, although no clear epidemiological evidence on this is currently available.[14] The estimation of the glandular dose itself is also uncertain, and may differ by screening age, breast density and breast thickness. Regional variations in glandular dose of between 1.04 and 1.63 mGy per view have been observed in the Netherlands, which were mainly related to differences in technical conditions (i.e., anode-filter combinations) of the screening units.[15] It may be difficult to distinguish the harm due to exposure to the natural background radiation from the harm due to screening. The average natural background dose is 2.4 mSv per year,[14] meaning that (assuming that 1 mSv = 1 mGy) the cumulative glandular dose during 35 years of screening (between age 40 and 75) would be 84 mGy (2.4 x 35). As a comparison, the total screening dose would be 24.7 mGy (1.3 mGy x 18 screening rounds between age 40 and 75 + 1 extra view at the initial screening round).

The overall uncertainty in the assessment of radiation risks has been estimated to be a factor 2–3.[14, 35] Nevertheless, even if we underestimated the risks by a factor of 3, the benefits of screening for women aged 50–74 would strongly outweigh the radiation risks.

Despite the observation that radiation risks increase with younger exposure age,[10] screening from age 40 would not severely jeopardise the benefit–risk ratio. Even if the screening effectiveness would be 25% lower and the radiation dose twice as high as in the current analysis, the radiation risks would be small. Our predictions for women between age 40 and 49 are more favourable than that of Mattsson *et al* (2000) and Berrington de Gonzalez and Reeves (2005), who expected a net increase in breast cancer deaths at a total glandular dose of ≥50 mGy and a mortality reduction <20%.[11–12] The difference may be related to model choice (EAR instead of ERR). Future developments in breast cancer screening, such as digital mammography, may further increase the screening benefits for women below 50,[41] and has the potential to reduce the absorbed radiation dose by 17%.[37] Our results confirm that the benefit-risk ratio of screening from age 30 on is very delicate.

Of course, the risk of radiation is just one of the possible harms of mammography screening. In a decision whether or not to screen before age 50, the risks of false-positive and false-negative mammograms, as well as the risk of overdiagnosis should be taken into account.

Conclusion

The radiation risks of mammography screening between age 40 and 74 were predicted to be negligible. From age 30, the balance between screening benefits and radiation risks would become fragile.

REFERENCES

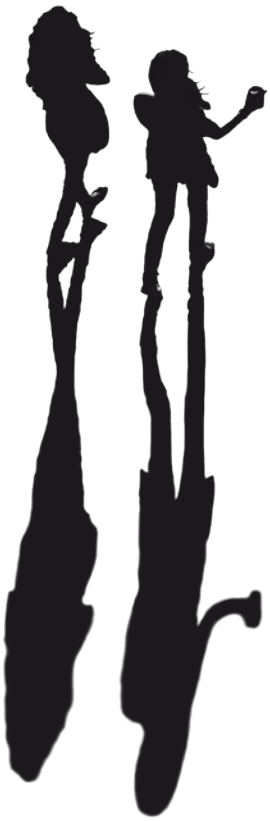
1. Allgood PC, Warwick J, Warren RM, Day NE, Duffy SW. A case-control study of the impact of the East Anglian breast screening programme on breast cancer mortality. Br J Cancer 2008;98:206–9.
2. Otto S, Boer R, Broeders MJM, Fracheboud J, Reijerink-Verheij J, Otten JDM, Verbeek ALM, de Koning HJ. Evaluation of the breast cancer screening programme in Southwest Netherlands: a case-control study. Eur J Cancer 2010;Suppl 8(3): 53 (Abstract 1N).
3. Puliti D, Miccinesi G, Collina N, De Lisi V, Federico M, Ferretti S, Finarelli AC, Foca F, Mangone L, Naldoni C, Petrella M, Ponti A, *et al.* Effectiveness of service screening: a case-control study to assess breast cancer mortality reduction. Br J Cancer 2008;99:423–7.
4. Boice JD, Jr., Land CE, Shore RE, Norman JE, Tokunaga M. Risk of breast cancer following low-dose radiation exposure. Radiology 1979;131:589–97.
5. Land CE. Low-dose radiation--a cause of breast cancer? Cancer 1980;46:868–73.
6. Land CE, Boice JD, Jr., Shore RE, Norman JE, Tokunaga M. Breast cancer risk from low-dose exposures to ionizing radiation: results of parallel analysis of three exposed populations of women. J Natl Cancer Inst 1980;65:353–76.
7. Howe GR, McLaughlin J. Breast cancer mortality between 1950 and 1987 after exposure to fractionated moderate-dose-rate ionizing radiation in the Canadian fluoroscopy cohort study and a comparison with breast cancer mortality in the atomic bomb survivors study. Radiat Res 1996;145:694–707.
8. Henderson TO, Amsterdam A, Bhatia S, Hudson MM, Meadows AT, Neglia JP, Diller LR, Constine LS, Smith RA, Mahoney MC, Morris EA, Montgomery LL, *et al.* Systematic review: surveillance for breast cancer in women treated with chest radiation for childhood, adolescent, or young adult cancer. Ann Intern Med 2010;152:444–55; W144–54.
9. Little MP, Hoel DG, Molitor J, Boice JD, Wakeford R, Muirhead CR. New models for evaluation of radiation-induced lifetime cancer risk and its uncertainty employed in the UNSCEAR 2006 report. Radiat Res 2008;169:660–76.
10. Preston DL, Mattsson A, Holmberg E, Shore R, Hildreth NG, Boice JD, Jr. Radiation effects on breast cancer risk: a pooled analysis of eight cohorts. Radiat Res 2002;158:220–35.
11. Berrington de Gonzalez A, Reeves G. Mammographic screening before age 50 years in the UK: comparison of the radiation risks with the mortality benefits. Br J Cancer 2005;93:590–6.
12. Mattsson A, Leitz W, Rutqvist LE. Radiation risk and mammographic screening of women from 40 to 49 years of age: effect on breast cancer rates and years of life. Br J Cancer 2000;82:220–6.
13. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet 2006;368:2053–60.
14. National Research Council of the National Academies. Health risks from exposure to low levels of ionizing radiation: BEIR VII Phase 2ed. Washington, D.C.: The National Academies Press, 2006.
15. Zoetelief J, Veldkamp WJ, Thijssen MA, Jansen JT. Glandularity and mean glandular dose determined for individual women at four regional breast cancer screening units in the Netherlands. Phys Med Biol 2006;51:1807–17.

16. de Koning HJ, Boer R, Warmerdam PG, Beemsterboer PM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. J Natl Cancer Inst 1995;87:1217–23.
17. Groenewoud JH, Pijnappel RM, van den Akker-Van Marle ME, Birnie E, Buijs-van der Woude T, Mali WP, 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;90:383–92.
18. National Evaluation Team for Breast cancer screening (NETB), National evaluation of breast cancer screening in the Netherlands 1990–2007. Twelfth evaluation report. Erasmus MC University Medical Center Rotterdam, Radboud University Nijmegen Medical Centre, 2009.
19. Adami HO, Malker B, Holmberg L, Persson I, Stone B. The relation between survival and age at diagnosis in breast cancer. The New England journal of medicine 1986;315:559–63.
20. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. Cancer 1989;63:181–7.
21. Michaelson JS, Silverstein M, Sgroi D, Cheongsiatmoy JA, Taghian A, Powell S, Hughes K, Comegno A, Tanabe KK, Smith B. The effect of tumor size and lymph node status on breast carcinoma lethality. Cancer 2003;98:2133–43.
22. Michaelson JS, Silverstein M, Wyatt J, Weber G, Moore R, Halpern E, Kopans DB, Hughes K. Predicting the survival of patients with breast carcinoma using tumor size. Cancer 2002;95:713–23.
23. Sant M, Allemani C, Capocaccia R, Hakulinen T, Aareleid T, Coebergh JW, Coleman MP, Grosclaude P, Martinez C, Bell J, Youngson J, Berrino F. Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. Int J Cancer 2003;106:416–22.
24. Tabar L, Vitak B, Chen HH, Duffy SW, Yen MF, Chiang CF, Krusemo UB, Tot T, Smith RA. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. Radiol Clin North Am 2000;38:625–51.
25. Bjurstam N, Bjorneld L, Warwick J, Sala E, Duffy SW, Nystrom L, Walker N, Cahlin E, Eriksson O, Hafstrom LO, Lingaas H, Mattsson J, *et al.* The Gothenburg Breast Screening Trial. Cancer 2003;97:2387–96.
26. Nystrom 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;359:909–19.
27. International Commission on Radiological Protection (ICRP). ICRP Publication 60: 1990 Recommendations of the ICRPed. Oxford, UK: Pergamon, Elsevier Science, 1990.
28. National Council on Radiation Protection and Measurements (NCRP). Limitating of Exposure to Ionizing Radiation, NCRP Report No. 116ed. Bethesda, MD, 1993.
29. Environmental Protection Agency (EPA). Estimating Radiogenic Cancer Risks. Addendum: Uncertainty Analysised. Washington DC: US Environmental Protection Agency, 1999.
30. United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR). Sources and Effects of Ionizing Radiation. UNSCEAR Report to the General Assembly, Volume II: Effectsed. New York: United Nations, 2000.
31. National Institutes of Health (NIH). Report of the NCI-CDC Working Group to Revise the 1985 NIH Radioepidemiological Tables. NIH Publication No. 03–5387ed. Bethesda, MD, 2003.
32. International Agency for Research on Cancer. Volume 7: Breast Cancer Screening. In: Vainio H, Bianchini F. IARC Handbooks of Cancer Preventioned. Lyon: IARCPress, 2002.
33. National Research Council. Health effects of exposure to low levels of ionizing radiationed. Washington, D.C.: National Academy Press, 1990.

34. Law J, Faulkner K. Cancers detected and induced, and associated risk and benefit, in a breast screening programme. Br J Radiol 2001;74:1121–7.
35. Law J, Faulkner K, Young KC. Risk factors for induction of breast cancer by X-rays and their implications for breast screening. Br J Radiol 2007;80:261–6.
36. Beemsterboer PM, Warmerdam PG, Boer R, de Koning HJ. Radiation risk of mammography related to benefit in screening programmes: a favourable balance? J Med Screen 1998;5:81–7.
37. Hendrick RE, Pisano ED, Averbukh A, Moran C, Berns EA, Yaffe MJ, Herman B, Acharyya S, Gatsonis C. Comparison of acquisition parameters and breast dose in digital mammography and screen-film mammography in the American College of Radiology Imaging Network digital mammographic imaging screening trial. AJR Am J Roentgenol 2010;194:362–9.
38. Young KC, Burch A, Oduko JM. Radiation doses received in the UK Breast Screening Programme in 2001 and 2002. Br J Radiol 2005;78:207–18.
39. Duijm LE, Groenewoud JH, de Koning HJ, Coebergh JW, van Beek M, Hooijen MJ, van de Poll-Franse LV. Delayed diagnosis of breast cancer in women recalled for suspicious screening mammography. Eur J Cancer 2009;45:774–81.
40. Broeders MJ, Scharpantgen A, Ascunce N, Gairard B, Olsen AH, Mantellini P, Mota TC, Van Limbergen E, Seradour B, Ponti A, Trejo LS, Nystrom L. Comparison of early performance indicators for screening projects within the European Breast Cancer Network: 1989–2000. Eur J Cancer Prev 2005;14:107–16.
41. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, Conant EF, Fajardo LL, Bassett L, D'Orsi C, Jong R, Rebner M, *et al.* Diagnostic performance of digital *versus* film mammography for breast-cancer screening. The New England journal of medicine 2005;353:1773–83.

Discussion

General discussion



GENERAL DISCUSSION

This thesis discussed several important benefits and harms of breast cancer screening. Despite several randomized controlled trials and extensive experience with population-based screening in several Western countries, much controversy still seemed to exist on both harms and benefits.[1–5] Using a micro-simulation model, we assessed the positive and negative screening consequences, and explored why opinions on this topic vary to a large extent. We also studied the benefits and risks of future screening scenarios, such as advancing the lower age limit for screening.

8.1 ANSWERS TO RESEARCH QUESTIONS

1. What are the benefits and harms of population-based mammography screening in the Netherlands?

1a. What are the benefits and harms of screen-film mammography screening?
In a modelled scenario with 30 years of biennial screening between 1990 and 2020, screen-film mammography was predicted to detect 26,720 breast cancers during the total lifespan of a population of 1,000,000 women aged 0–100 years in 1989 that are screened at least once (Chapter 2). For approximately 25% of all women with a screen-detected cancer early detection meant that a breast cancer death is prevented. Thus, a total of 6577 breast cancer deaths (25% of 26,720) would be averted during the lifetime of the simulated population. This equals a reduction of the breast cancer mortality of 23% compared to the predicted breast cancer mortality in a situation without screening. The modelled mortality reduction was smaller than the reductions of 20–30% observed in randomised controlled trials and nation-wide mammography screening programmes,[6–10] because in a population aged 0–100 years in 1989 some women will have had only few screening rounds. Furthermore, by calculating the effects during the remaining life span, we included years in which the modelled screening programme had already stopped.

One of the harms of breast cancer screening is overdiagnosis. We predicted that in a population of 1,000,000 women aged 0–100 years in 1989 with at least 1 screening round, 1926 tumours would be overdiagnosed during a screening period of 30 years. This is a risk of 2.1% of all diagnosed breast cancers in women with at least 1 screening examination, or 7.2% of all screen-detected cancers.

1b. How has the implementation of digital mammography affected the breast cancer mortality and risk of overdiagnosis compared to screen-film mammography?

Digital mammography led to a statistically significant 80% increase in the number of screen-detected DCIS cases compared to screen-film mammography (Chapter 2). As a consequence, digital mammography was predicted to prevent 6864 breast cancer deaths per 1,000,000 women aged 0–100 years in 1989 with at least 1 screening, while screen-film mammography would prevent 6577 deaths. This is a relative increase of 4.4%. The total predicted reduction in breast cancer mortality, relative to a situation without screening, would increase from 23% to 24%. The number of overdiagnosed tumours would also rise because of the increased DCIS detection: from 1926 per 1,000,000 screened women using screen-film mammography to 2327 cancers at digital mammography screening. The overdiagnosis rate, measured as a fraction of all diagnosed breast cancers in digitally screened women, would be 2.5% (8.2% of all screen-detected cancers). This is an increase of 21% compared to screen-film mammography.

2. Why do estimates of overdiagnosis that are reported in literature vary to a large extent?
Overdiagnosis estimates in recent literature vary between 1% of all diagnosed breast cancers in a screened population to 54% of all expected breast cancers in an unscreened population (Chapter 3). Based on the implementation of mammography screening of more than 1 million Dutch women, we illustrated that part of this difference could be ascribed to the moment at which overdiagnosis is estimated. Ideally, overdiagnosis should be assessed when screening is in a steady state phase, because the proportion of women that have a prevalent screening is then stable. Moreover, a compensatory drop in incidence will reach its maximum only if all women in the age group past the screening age had been invited to screening when they were eligible. This is not always sufficiently taken into account,[2, 5, 11–12] and could lead to an overestimation of overdiagnosis of a factor 4, in the worst-case scenario. Sufficient follow-up to allow for lead time is crucial in calculating overdiagnosis correctly. On the basis of the estimated distribution of the lead time of breast cancer in our study (the best-fitting model assumed a Weibull-distributed lead time with a median of 2 years and a mean lead time of 3.7 years), we predicted that approximately 20% of all tumours will have a lead time of more than 5 years, and 5% will have a lead time of more than 10 years. Ideally, the lead time of these tumours should be accounted for by calculating overdiagnosis several years after screening has reached the steady-state phase.

Another explanation for the observed differences in overdiagnosis estimates is the choice of denominator that defines the population at risk. Estimates that included all expected cancers in women aged 0–100 years differed by a factor 3.5 from estimates that calculated overdiagnosis as relative risk of being diagnosed with breast cancer for women in the screening age compared to a situation without screening. Finally, varying overdiagnosis estimates could also partly be ascribed to differences in background

cancer epidemiology between countries, the inclusion of DCIS in the estimate (Chapter 2) and to differences in screening characteristics and performance.

3. Do women who participate in breast cancer screening unnecessarily delay a visit to a doctor when they develop breast cancer symptoms, because of 'false reassurance'?
Breast cancer patients who had participated in the Dutch breast cancer screening program prior to noticing symptoms did not have a significantly longer period between the initial discovery of breast abnormalities and the first time presenting this to a doctor than patients who had not been screened (Chapter 4). The mean time between the first breast abnormalities and first consultation of a general practitioner ('symptom–GP period') in patients with a previous screening history was 7 days; in patients without a previous screening history 13.5 days. The proportion of patients with a symptom–GP period of ≥30 days neither differed between screened and control patients (31.2% in the symptomatic screened group, 31.0% in the control group). The median period between the first GP- and breast clinic visit was 7.0 days in screened patients and 6.0 days in control patients. These data suggest that false reassurance plays, at most, only a minor role in breast cancer screening.

4. How does the organizational form of breast cancer screening affect its benefits and costs?

4a. What is the effectiveness of opportunistic breast cancer screening compared to organized screening, in terms of breast cancer mortality and life years gained?
Organized and opportunistic breast cancer screening were predicted to be comparably effective in exemplar country Switzerland (Chapter 5). Organized screening at an 80% participation rate prevented 4921 breast cancer deaths per 1,000,000 women aged 30–70 years in 1999, which is a breast cancer mortality reduction of 13% during the lifespan of this population, and a gain of 81,000 life years. If 80% of the population would have an opportunistic mammography once every two years, a predicted 4876 breast cancer deaths per 1,000,000 women would be prevented, which equals a breast cancer mortality reduction of 13%. In this scenario, 80,400 life years would be gained.

4b. What is the cost-effectiveness of opportunistic versus organized breast cancer screening?
The costs of opportunistic screening per life year gained were twice those of organized screening: € 23,617 *versus* € 11,512 (Chapter 5). This is mainly due to the costs of follow-up diagnostics after a positive mammogram, which were more than twice as high at opportunistic screening than at organized screening. The elevated costs were associated with the frequency with which additional imaging examinations were being performed in combination with an opportunistic mammogram. Even if the rate of false-negative mammograms at opportunistic screening would be 25% lower than at organized screening, cost-effectiveness would not change markedly. For opportunistic screening to become

equally cost-effective as organized screening, costs and use of additional diagnostics should therefore be reduced.

5. What are the effects of adjuvant systemic breast cancer treatment and mammography screening on breast cancer mortality in women younger and older than 50?

5a. What are the predicted effects of adjuvant systemic therapy on breast cancer mortality in women aged 50 or older?
Adjuvant systemic therapy was estimated to have reduced the breast cancer mortality rate in 2008 in the total simulated population aged 19–100 years by 14%, compared to a situation without such treatment (Chapter 6). In women aged 40–49 years, the estimated reduction in the 2008 breast cancer mortality rate was 27%, in women aged 50–74 years 15%, and in women aged 75–84 years 9%, compared to a situation without the intervention.

5b. In the presence of adjuvant systemic therapy, what are the predicted effects of mammography screening of women aged 50–74 years on breast cancer mortality?
In the presence of adjuvant systemic therapy, mammography screening of women aged 50–74 years was estimated to reduce the 2008 breast cancer mortality rate in the total simulated population aged 19–100 years by 16%, compared to a situation without screening (Chapter 6). The screening-related reduction in breast cancer mortality in 2008 would be 21% in women aged 50–74 years and 19% in women aged 75–84 years. The combined effect of screening and adjuvant treatment was somewhat lower than the separate effects, because screening will avert some of the deaths that would have been prevented by adjuvant treatment, and vice versa. Also, the age distribution of women with screen-detected cancers may be different than that of women with symptomatic cancers, which could affect the age-dependent effects of adjuvant treatment. We estimate that the reduction in breast cancer mortality, attributable to the combined use of adjuvant systemic treatment and screening of women aged 50–74 years, would be 27% compared to a situation without any intervention, in the total simulated population aged 19–100 years in 2008.

5c. What are the predicted effects of adjuvant treatment and screening starting between age 40 and 50 on breast cancer mortality?
Expanding the lower age limit of screening has the potential to further reduce the mortality rate (Chapter 6). One additional screening round at age 48 would decrease the breast cancer mortality rate in the total simulated population from 48.8 per 100.000 woman-years to 48.4 per 100.000 woman-years: a reduction of 1.0% compared to a situation with adjuvant therapy and screening from age 50. Two additional screening

rounds starting at age 46 would lower this rate by 1.6%, to a breast cancer mortality rate of 48.0 per 100.000 woman-years. If the current biennial screening programme is to be preceded by annual mammography between age 40 and 49, the breast cancer mortality would be reduced by an extra 5.1%, to 46.3 per 100.000 woman-years. In this last scenario, the overall reduction in the breast cancer mortality rate attributable to adjuvant systemic therapy and screening was estimated to be 31% in the total simulated population in 2008.

6. How do the benefits of mammography screening in women younger and older than 50 relate to the risk of radiation-induced breast cancer and breast cancer death?

6a. What are the radiation-induced risks of breast cancer screening in women aged 50–74 years, as compared to the benefits of screening?
Our study (Chapter 7) demonstrated that the risk of radiation-induced breast cancer due to mammography screening is very small. The average glandular screening dose in the Netherlands is 1.3 mGy per view. At this dose, biennial screening between age 50 and 74 was estimated to cause 7.7 breast cancers and 1.6 breast cancer deaths per 100,000 women aged 0–100 years, but would also prevent 1121 breast cancer deaths. The benefits of mammography screening in women aged 50–74 years thus strongly outweigh the radiation risks. The outcomes of the study were unlikely to be strongly affected by model assumptions or variations in the absorbed radiation dose. If the risks of mammography screening would be calculated at average glandular doses of 1 to 5 mGy/ view – a range which has been observed between screening units in the Netherlands – the number of induced breast cancers would range between 5.9 and 26.6 per 100,000 women. The number of induced breast cancer deaths would range between 1.3 and 6.3 per 100,000 women.

6b. How would an extension of the lower age limit for screening to ages below 50 affect radiation risks, as compared to the benefits?
Biennial screening of women aged 40–74 years was predicted to induce 3.7 breast cancer deaths (range: 2.9–14.4 at doses between 1 and 5 mGy per view), but would also prevent 1302 deaths. Even in the unlikely situation that the average radiation dose would exceed 10 mGy per view and the model has underestimated the number of induced breast cancer deaths by a factor of 3, the radiation risks of screening would not outweigh the benefits. Annual screening from age 30 would induce more deaths than it prevents if the average dose would be 7 mGy per view or more, and the radiation risks would be overestimated by a factor of 3. However, screening in this age group in the Netherlands is only recommended when women are at high risk for breast cancer, for instance because of an inherited BRCA1 or BRCA2 gene mutation.

8.2 INTERPRETATION OF OUR FINDINGS

The predicted effects of screening

Our study showed that breast cancer screening is effective in reducing breast cancer mortality. In Chapter 2, the predicted reduction in breast cancer mortality rate that was attributable to screening was estimated to be 23%–24%. In the other chapters of this thesis, however, mortality reductions of between 13% and 26% were predicted (Table 8.1). Similar inconsistency has been observed in the predicted number of annually prevented breast cancer deaths in the Netherlands. Estimates ranged between 680 and 1160 averted deaths per year (Table 8.1). In Dutch screening literature, estimates of between 500 to 800 prevented deaths[13–15] seem to contradict the estimate of 1300 prevented breast cancer deaths that would be obtained using the average prognosis of a screen-detected cancer (for approximately 25% of the 5500 women with a screen-detected cancer each year, early detection means that breast cancer death can be prevented.[16–17]

This broad range could partly be related to differences in the models that were used, which depended on the research question under study. For instance, the demography, breast cancer epidemiology and screening characteristics in Switzerland (Chapter 5) differed from the Netherlands (Chapter 2, 6 and 7). By modelling a rising breast cancer trend in Chapter 6, the absolute screening benefits may have been higher than calculations based on a constant incidence. On the other hand, screening benefits in this chapter may be smaller because we included adjuvant therapy in the model, which can reduce screening benefits.[18] The number of averted breast cancer deaths may also depend on the length of the screening period, the screening ages and attendance rates, which varied between all chapters, again depending on the study question (Table 8.1).

The estimator that was used to calculate the mortality reduction (Table 8.1) has likely influenced the magnitude of the screening benefits. For instance, by measuring screening effects during the entire lifespan of a population (such as in Chapters 2, 5 and 7), all potential screening benefits could be taken into account. In 2008, on the contrary, the mortality reduction had not yet reached its maximum, because not all women older than 75 had received a mammogram and nobody in the targeted age group had had all 13 screening rounds. For this reason, estimates in Chapter 6 were smaller than the other estimates. Using an infinite screening period (Chapter 7) increased the average number of screening mammograms per woman, and thereby the mortality reduction.

Based on the outcomes of our analyses, we estimated that the total number of prevented breast cancer deaths in the Netherlands in 2011 ranged between the estimates of Chapter 5 and 6: between 683 and 858 breast cancer deaths were avoided by screening. If

Table 8.1 Estimates of the screening-related reduction in breast cancer mortality according to study design and estimator

Chapter	Research question	Screening period, interval, age group, attendance	Simulated population	Period of measure-ment	Predicted number of prevented breast cancer deaths compared to no screening (breast cancer mortality rate reduction)	Observed or estimated Dutch population	Estimated number of prevented breast cancer deaths per year of screening
Chapter 2	Effects and risks of digital versus screen-film mammo-graphy	1990–2020; biennial screening of women aged 50–69 years between 1990 and 1998, and of women aged 50–74 years between 1998 and 2020; 82% participation	Women aged 19–69years in 1989, with at least 1 mammogram, dynamic population model	Deaths during entire lifespan (up to 2070)	6577–6864 per 1,000,000 women at screen-film and digital mammography, respectively (23%–24%)	4,875,625 women * 0.85 (= simulated fraction of population with at least 1 mammogram) = 4,144,281 women aged 19–69 years in 1989	909–948
Chapter 5	Effects, costs and cost-effectiveness of opportunistic and organized screening	1999–2020; biennial screening of women aged 50–69 years; 80% participation/ screening uptake	Women aged 30–70 years in 1999, dynamic population model	Deaths during entire lifespan (up to 2099)	4876–4921 per 1,000,000 women (13%)	4,170,423 women aged 30–70 years in 1989	968–977
		1999–2020; biennial screening of women aged 50–69 years; 80% participation/ screening uptake	Women aged 30–70 years in 1999, dynamic population model	Mortality rate ratio in 2018	918–980 per 4,110,450 women, at opportunistic and organized screening, respectively (23%–25%)	An estimated 3,600,000 women aged 49–89 years in 2018	804–858

Table 8.1 (continued)

Chapter 6	Effects of screening as compared to adjuvant therapy	1990–2027; biennial screening of women aged 50–69 years between 1990 and 1998, and of women aged 50–74 years between 1998 and 2020; 82% participation	Women aged 0–100 years in 1989, dynamic population model	Mortality rate ratio in 2008	9.1 per 100,000 women (16%)	7,511,371 women aged 0–100 years in 1989	683
Chapter 7	Prevented versus radiation-induced breast cancer deaths	Biennial screening of women aged 50–74 years, 100% participation	Women born in year 0, cohort model	Deaths during entire lifespan (up to year 100)	1121 per 100,000 women (= 8,044,350 woman-years according to MISCAN life tables) (26%)	8,311,420 woman-years in population aged 0–100 in 2008	1160 (928 if an 80% participation rate is assumed)

future screening effects are taken into account, this number could further increase to approximately 928 prevented breast cancer deaths per year (Chapter 7). Such estimates, however, should be used with caution. On the one hand, population growth and increasing breast cancer incidence, which were not included in the present study (except for Chapter 6), may raise this number, but on the other hand, improving breast cancer treatment could also lower screening effects.[19] Careful evaluation of breast cancer mortality trends therefore remains crucial.

Comparison with other breast cancer screening studies

A vast amount of literature supports our findings that mammography screening effectively reduces breast cancer mortality.[20–22] Randomized controlled trials showed reductions of 25%–30% in women in the screening arm of the trial; case-control studies estimated the mortality effects of screening to range between 38% and 70%.[22] Several recent publications in high impact journals, however, suggest that the effects of breast cancer screening are much smaller than previously assumed, or even absent. Kalager *et al.* (2010), for instance, showed that mammography screening in Norway reduced breast cancer mortality by only 3%.[3] They compared incidence-based breast cancer mortality between women living in regions that had and regions had not started screening, and between the current situation and a historical situation without any screening. The mortality rate reduction in screened women, as compared to their historical counterparts, was 10% larger than

the reduction in non-screened women. However, because of the decreasing mortality in non-screened women, only a third of the 10% mortality reduction could actually be attributed to screening.

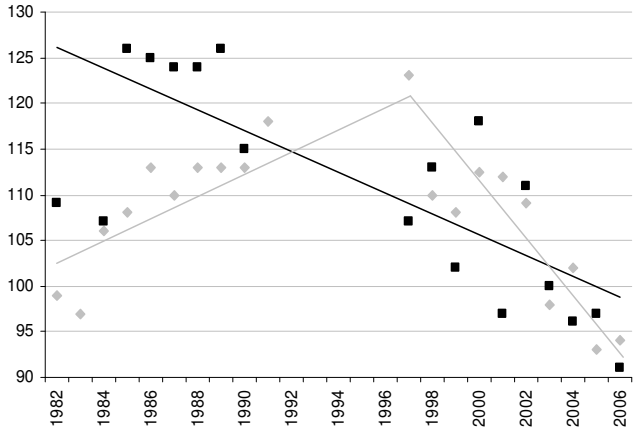
A Danish study by Jorgensen *et al.* (2010) compared trends in breast cancer mortality between a 10 years period before screening started, and a 10 years period of 'steady state' screening, starting at least 5 years after the introduction of screening. Women living in regions with a screening program ('screened group') were compared to women of the same age living in regions without a screening program ('control group').[4] The authors reported that breast cancer mortality in the screened group dropped by 1% per year, whereas breast cancer mortality in control group decreased by 2% per year. They therefore concluded that screening has no effect.

A pair-wise comparison of matched neighbouring countries with and without screening programmes showed similar breast cancer mortality trends over time, suggesting that mammography screening did not play a role in reducing mortality.[1]

Why are these recent findings so different from former studies? In the study by Kalager *et al.*, the absence of a screening effect in the study is likely related to the short period of follow-up: the average follow-up after diagnosis was 2.2 years. In a response letter to the NEJM, van Ravesteyn *et al.* (2011) showed that it is unreasonable to expect a screening effect so soon.[23] Using MISCAN, they indeed predicted a 10% breast cancer mortality reduction at a 2.2 years follow-up, but the screening effect would quickly increase to a reduction of 16% if the follow-up period is prolonged by 5 years.

The mortality trend analysis by Jorgensen was based on two periods: a pre-screening period between 1982 and 1991, and a post-screening period between 1997 and 2006. Available observations for the period 1971–1981 and 1991–1996 were ignored (Figure 8.1a). Because the authors gave no convincing reasons why these data points were disregarded, we estimated a trend line between all available observations (1971–2006) and estimated the annual percent change in the control group to be 0.8% between 1971 and 1997, and –2.6% between 1997 and 2006. This would mean that the breast cancer mortality rate in 2006, relative to expected mortality rate based on the period 1971–1997, had decreased by 25%. In the screened group, the annual percent change was 1.2% between 1971 and 1993, and –2.8% between 1993 and 2006. The estimated mortality rate reduction in 2006, compared to the expected pre-screening trend, was 39% (Figure 8.1b). Thus, in contrast with Jorgensen *et al.*, our analysis that included all data observations showed that the breast cancer mortality reduction in screened women is larger than in non-screened women, indicating a screening effect. In a web appendix

a.



b.

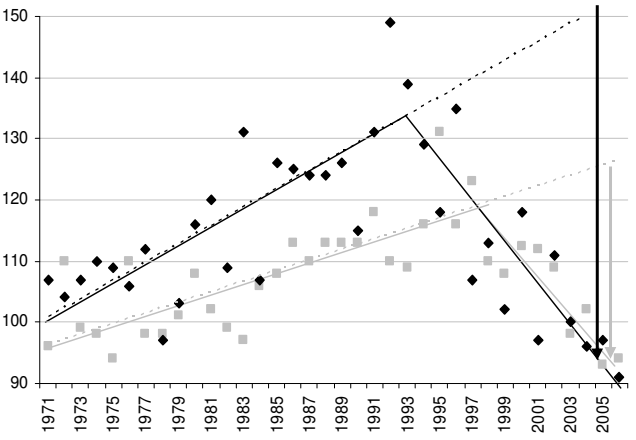


Figure 8.1a–b Reduction in breast cancer mortality in Denmark, calculated by: a) comparing data observations between the periods 1982–1991 and 1997–2006, similar to Jorgensen *et al* (2010)[4], and b) including all data observations between 1971 and 2006. Black diamonds: observed breast cancer mortality in screening group; black solid lines: estimated breast cancer mortality trend in screening group; grey squares: observed breast cancer mortality in control group; grey solid lines: estimated breast cancer mortality in control group; black dotted lines: expected breast cancer mortality in screening group without trend break; grey dotted lines: expected breast cancer mortality in control group without trend break

that was additionally published in the BMJ, Jorgensen *et al.* performed an extra analysis based on all available data. Although they now reported a stronger mortality decline in screened than in non-screened regions, they held on to their previous conclusions, because mortality reductions were also observed in women aged 35–54 who, according to them, could never have profited from screening. However, in Denmark, women are invited to have a screening mammogram from age 50. As a consequence, some women aged 50–54 could theoretically have benefitted from their mammogram. The conclusions from Jorgensen *et al.* therefore appear to be incorrect.

The pair-wise comparison of breast cancer mortality trends in countries with and without screening of Autier *et al.* (2011) suggested that no relevant differences existed between countries that could have affected mortality.[1] However, the study did not account for breast cancer incidence, tumour type and stage distribution, and only superficially described country differences in breast cancer treatment or the use of opportunistic mammography screening.[24] A simulation model, such as MISCAN, has the potential to specifically include confounders.

The predicted risk of overdiagnosis

In chapters 2 and 3 we assessed the risk of overdiagnosis. In Chapter 2, this risk was estimated to range between 2.1% of all diagnosed breast cancers at screen-film mammography and 2.5% of all diagnosed breast cancers at digital mammography. In Chapter 3, the overdiagnosis ranged was estimated to range between 2.8% and 9.7%, depending on the denominator that is used to define the population at risk. The choice of denominator will be determined by the research question, but should ideally be the same as the denominator that is used to estimate the screening benefits, in order to be able compare those two. We estimated the ratio between benefits and overdiagnosis risk to be 3 : 1, or, with other words, for 3 prevented breast cancer deaths, 1 breast cancer will be overdiagnosed and over-treated. This highly contradicts ratios of 10 overdiagnosed cancers *versus* 1 averted breast cancer death that are suggested in literature.[25]

8.3 THE BALANCE BETWEEN BENEFITS AND HARMS OF SCREENING

Benefits

For screening to be justified, its benefits should outweigh the risks. In this thesis we estimated that, currently, between 683 and 858 breast cancer deaths are annually prevented by screening in the Netherlands. For each prevented breast cancer death, approximately 16.5 life years would be gained (Chapter 5). The total number of life years gained would thus range between 11,270 and 14,157. Besides the prevention of mortality and gain of

life years, screening could also improve the stage at which cancer is diagnosed, which could lower the need of radical treatment, morbidity and the number of outpatient clinic visits, thereby increasing quality of life.[26]

Harms

Having a mammogram can be perceived as painful and inconvenient, and although the quality of life is only impaired slightly and for a relatively short period,[26] the number of women that undergo screening each year is large. In the Netherlands, almost 900,000 women have a mammogram each year. To prevent 683–858 breast cancer deaths, we therefore estimated that 1049–1317 mammograms need to be performed (= number needed to screen, 'NNS'). This means that for 1048–1316 women with mammogram, screening will have no effect. A NNS of 1049–1317 is higher than predicted in Chapters 2 and 5, where we estimated the number needed to screen to range between 798 and 856. This discrepancy may be related to the fact that these chapters accounted for future screening effects, by calculating breast cancer mortality over an entire lifespan or in 2018.

Of all women who have a mammogram, 1.8% is referred for further diagnostic assessment because of an abnormal mammogram. In 30% of these cases, a breast cancer is detected; in the other 70%, the mammogram was ultimately found to be normal, or with other words, was 'false-positive'.[17] False-positive mammograms require additional diagnostic testing and invasive biopsies, which can cause distress. Moreover, false-positive mammograms are associated with anxiety and an increased number of health care visits.[19, 27–28] However, for most women this anxiety is short-lived,[27] and viewed as an acceptable consequence of screening.[19] The cumulative risk of a false-positive mammogram in the Netherlands, assuming that a woman participates in all 13 screening rounds, is 9%.[17] Because of low referral rates, rates of false-positives in the Netherlands are much lower than in other countries.[17] In the United States, for instance, the cumulative risk of a false-positive mammogram calculated over 10 screening rounds was 50%; in Norway this risk was 21%.[19]

Of those women with a true-positive mammogram, i.e. a screen-detected cancer, MISCAN predicts that only 26% was estimated to actually benefit from screening. In 56% of the cases, a woman would also have survived breast cancer without screening, in 13% the woman would die despite screening, and in 6% of the cases, the tumour would have never been diagnosed without screening ('overdiagnosis'). Therefore, 74% of all women with a screen-detected cancer live longer with the diagnosis breast cancer while they would not have profited. They are however treated and suffer from the anxiety and

distress that comes with a diagnosis. Some women may suffer from morbidity because of their treatment.

A negative mammogram may also involve certain risks. 'False reassurance' because of a prior negative screening mammogram may cause some women to delay the presentation of breast cancer symptoms. Although Chapter 4 showed that screened women did not significantly differ from non-screened women in the mean period between noticing a first sign of breast cancer and presenting the symptoms to a health care professional for the first time, 4% of women with screen-detected cancers reported to have had symptoms prior to the mammogram, and had delayed symptom presentation for a substantial time. These women were diagnosed in stage T1c (N unknown) and T1cN+. Because the percentage of women that delay to present symptoms is low and statistically non-significant, we assume that the harms are negligible. False-reassurance may be a more pronounced screening risk in other countries or screening programmes, and could for instance depend on the information that is provided to screening attendees with regard to symptoms that could arise between two screening mammograms.

Mammography itself may cause breast cancer due to ionizing X-ray radiation (Chapter 7). We estimated that for every 684 breast cancer deaths that are averted, screening induces 1 breast cancer death (thus, 0.0015 induced breast cancer deaths for each averted death). Because this number almost equals the estimated number of annually prevented breast cancer deaths in the Netherlands (n=683, Chapter 6), we predict that screening would cause 1 breast cancer death per year.

Quality of life

As discussed above, mammography screening may at the same time positively and negatively affect quality of life. In Chapter 5, we presented estimates of breast cancer utilities, based on a study by de Haes *et al.* (1991). These utilities describe how various stages of breast cancer and treatment can affect quality of life. The study showed that while having a mammogram and undergoing breast cancer treatment can lower one's quality of life, preventing terminal illness could also counterbalance such harms. Considering that screening on the one hand increases breast cancer incidence, but on the other hand also lowers mortality, we estimated the loss in quality of life to be 3824 life years per 1,000,000 women. This equals 4.7% of the number of life years gained by screening in this population (81,000). As we estimated that for each averted breast cancer deaths 16.5 life years would be gained, we can assume that 0.8 (4.7% out of 16.5) quality-adjusted life years may be lost for each prevented breast cancer death. One should note, however, that utility estimates are rather arbitrary (the utility value for metastatic breast cancer

Table 8.2 Benefits and harms of breast cancer screening

Benefits	Harms
1 breast cancer death prevented[a]	1183 mammograms performed[b]; 1182 without benefit
16.5 life years gained	0.8 quality-adjusted life year lost
	23 referrals
	16 false-positive mammograms
	7 true positive mammograms; 5 of which without benefit
	3.7 would have survived anyhow
	0.9 would have died anyhow
	0.5 would never have been diagnosed ('overdiagnosis')
	Extra deaths because of false reassurance: negligible
	Extra breast cancers because of radiation: 0.0069
	Extra breast cancer deaths because of radiation: 0.0015

[a] To obtain the actual number of benefits and harms in the Netherlands per year, all numbers in this table should be multiplied by a factor 771 (the average between 683 and 858 avoided breast cancer deaths). [b] Calculated as the average between 1049 and 1317 women needed to screen.

ranged from –0.52 to 0.882 in literature) and strongly depended on study methodology.[29] Quality of life assessments should therefore be interpreted with caution.

Balancing benefits and harms

Table 8.2 compares the benefits and harms of screening. For each breast cancer death that is prevented, 16.5 life years would be gained. On the downside, 0.8 quality-adjusted life year may be lost because of screening and the consequential increase in breast cancer. Moreover, 1300 mammograms would be needed to avoid 1 breast cancer death, at which 23 women would be referred for follow-up diagnostics, and 16 would receive a false-positive result. Seven women would receive a true positive mammogram, but for 5 of them screening would have no benefit, because they would have died of breast cancer anyhow (3.7 women), they would have survived anyhow (0.9 woman), or they would never have been diagnosed with breast cancer if they were not screened (0.5 woman).

For women that consider having a mammogram, it is therefore important to realize that although screening inevitably involves certain risks, the benefits of mammography screening outweigh the harms.

8.4 LIMITATIONS OF OUR STUDY

The main limitation of our study is, of course, the need to make some assumptions for our model. Because, in principle, all diagnosed breast cancers are treated, it is not exactly known how the disease would have developed without intervention. Hypotheses have to be made on the onset rate of preclinical cancer and the average sojourn time (the period in which a pre-clinical tumour can become detected by screening). The fraction of invasive breast cancers that is preceded by DCIS is uncertain, and it is unknown whether all preclinical DCIS cases progress, or if some are dormant or regress. Furthermore, assumptions need to be made on the age- and stage specific fatality of cancer and the impact of a screening program. The role of underlying time trends, such as an increasing prevalence of certain risk factors for breast cancer and developments in breast cancer treatment and screening further complicates our analyses.

This does not imply that the natural history of breast cancer is completely unknown. Some indirect inferences can be made from randomized controlled trials and screening data. The average duration of pre-clinical cancer, for instance, is proportional to the ratio between the detection rate in the initial screening round and the clinical incidence rate without screening. Based on stage-specific incidence rates of clinically diagnosed and screen-detected breast cancer and rates of interval cancer, progression and regression rates might be assessed. Using detection rates per screening round and interval cancer rates, the sensitivity of mammography could be estimated. The survival after a breast cancer diagnosis and the influence of a screening program on these survival rates can be deducted from the randomized trials. In MISCAN, the effectiveness of screening was based on the Swedish Trials.[9, 14, 30–31] Despite the fact that these trials have been performed in the seventies and eighties, the observed breast cancer mortality in the Netherlands (and other countries) could be reasonably modelled using these data.

MISCAN, however, cannot provide exact estimates of the natural course of breast cancer. Chapter 2 showed that observed cancer rates could be simulated with strongly divergent model assumptions: models with extreme assumptions on the progression and regression rate of DCIS fitted equally well as our baseline model. As a consequence, the predicted mortality reduction of screening may range between 20.8% and 22.8% under extreme assumptions on the progression and regression rate of DCIS. The risk of overdiagnosis in this situation would range between 1.2% and 5.5%, compared to the breast cancer incidence in women aged 0–100 years in a situation without screening. If the consequences of digital screening are calculated, taking into account the increased detection of DCIS, the predicted breast cancer mortality reduction would range between 20.8% and 24.8%, and the overdiagnosis risk would range between 1.4% and 7.7%. In the best

case scenario, digital screening would thus imply that the risk of overdiagnosis increases by 14% compared to screen-film mammography, but in the worst case, this increase would be 43%. It is therefore important to assess the true progression and regression rate of DCIS, and embed model assumptions in available evidence.

Except for Chapter 6, our analyses were based on a constant background incidence of breast cancer over time (that is, the incidence that would be expected without screening). However, several studies showed that the background incidence and mortality of breast cancer in non-screened women in the Netherlands increased.[32–33] Although we simulated average incidence between 1990 and 2004 reasonably, the modelled incidence was higher than observed before 1990 and lower than observed since 2000.[17]

We also assumed a constant background breast cancer mortality trend in all chapters (except Chapter 6). These models simulated the observed breast cancer mortality fairly well. However, by including the increasing incidence trend in chapter 6, the modelled mortality also increased by time, to a level considerably higher than the observed trend from 1990 onwards. Even the modelled effects of adjuvant systemic therapy could not sufficiently explain the difference. We therefore assumed an additional survival improvement related to unknown causes, which may include improvements in diagnostics, primary and (neo-)adjuvant treatment, screening and increasing breast cancer awareness. Future modelling research should explore these incidence and survival trends and how they interact with screening.

8.5 THE FUTURE OF BREAST CANCER SCREENING

The effects of lowering the age limit of screening

Our study demonstrated that the current Dutch breast cancer screening program, targeting 50–74 years old women, effectively reduces breast cancer mortality at reasonably low risks. With steadily breast increasing incidence rates in the age groups that are not yet invited, the question has been raised whether to start screening from an earlier age. So far, studies on the effectiveness of screening 40–49 years old women are inconsistent. For instance, meta-analyses of randomized controlled trials that included women under age 50 estimated that annual screening between age 40 and 49 could significantly reduce breast cancer mortality by 15% to 19%.[20, 27] These trials, however, were not explicitly designed to study screening effects in this age group. The only randomized controlled trial that specifically aimed to measure the effects of annual screening below age 50 found a non-significant 17% reduction of breast cancer mortality.[34] In contrast, a statistically

significant breast cancer mortality reduction of 29% was observed in Swedish women aged 40–49 years who attended screening.[35]

Modelling the effects of mammography screening under age 50

International recommendations on the proper screening age are highly divergent: the 2009 US Preventive Services Task Force (USPSTF) recently recommended against screening service below age 50,[36] but other authoritative institutions advise women aged 40–49 years to have an annual or biennial mammogram.[37] Whatever guideline is correct, it is important to bear in mind that under certain circumstances the effects of breast cancer screening below age 50 may differ from those of the trials. Countries with higher breast cancer incidence, a less favourable tumour stage distribution, worse survival or better screening performance may observe significant mortality reductions by screening younger women. The decision whether or not to screen before age 50 should therefore be supported by modelling studies that can account for such factors.

In this thesis, we predicted that in the presence of adjuvant systemic therapy for breast cancer, screening between age 50 and 74 could reduce breast cancer mortality by 15.7%, and that 1 additional screening round at age 48 could lower this rate by an extra 1.0%. Ten annual screening rounds between age 40 and 49 would reduce the breast cancer mortality rate by an extra 5.1%, compared to biennial screening between age 50 and 74 (Figure 8.2). Although this affect appears small in relative terms, it would equal a reduction of 190 breast cancer deaths in addition to the estimated 683 breast cancer deaths that were estimated to be avoided by the screening programme in 2008. The additional mortality reduction to be obtained by screening under age 50 becomes relatively smaller the younger screening is started: 37 extra breast cancer are prevented by adding 1 screening round at age 48, 66 by two extra screening rounds from age 46, 91 by 3 screening rounds, 113 by 4 screening rounds, and 127 by 5 extra screening rounds starting at age 40. The relative mortality gain of (biennial) screening from age 40 is comparable to that estimated for the USA: a 3% mortality rate reduction would be predicted compared to biennial screening from age 50.[38]

The risks of mammography screening under age 50

The ultimate decision whether to initiate breast cancer screening before age 50 should be based on a balance between benefits and harms. The potential harms of commencing screening at earlier age include an increased number of false-positive and false-negative mammograms, additional diagnostics, overdiagnosis, and radiation-induced breast cancer. Elmore *et al.* (1998) observed a 56.2% cumulative risk of a false-positive mammogram after 10 screening rounds starting at age 40, and a 47.3% cumulative risk of a false-positive mammogram after 10 screening rounds starting at age 50.[39] More recently,

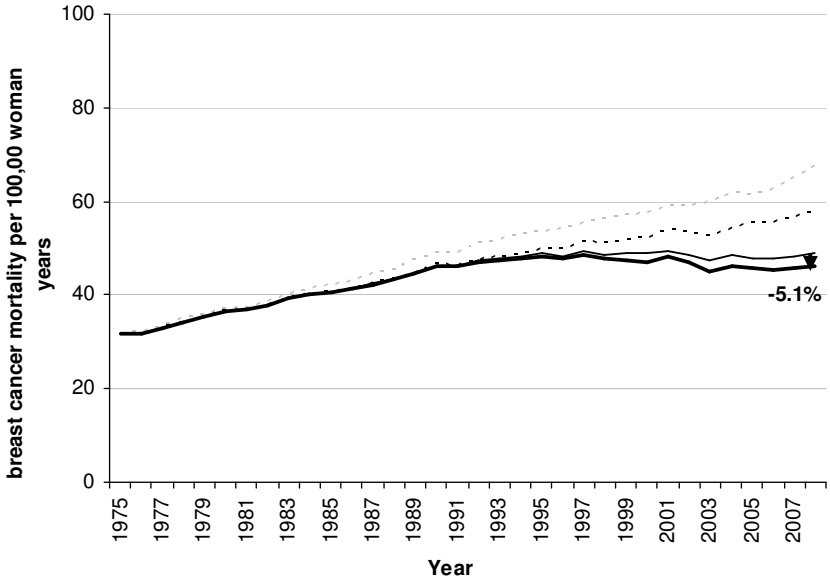


Figure 8.2 The effects of adjuvant systemic therapy, biennial mammography screening between age 50–74, and 10 additional screening examinations starting at age 40 on breast cancer mortality in 19–100 years old women. Grey dotted line: no adjuvant therapy, no screening; black dotted line: adjuvant therapy, no screening; black thin solid line: adjuvant therapy, biennial screening starting at age 50; black thick solid line: adjuvant therapy, annual screening between age 40 and 49 followed by biennial screening between age 50 and 74.

Salas *et al.* (2011) showed that the odds of having a false-positive mammogram increased by 43% if screening started at age 45 instead of age 50. The risk of a false-positive mammogram followed by an invasive diagnostic procedure also increased significantly.[40] Because referral and false-positive rates in the Netherlands are relatively low, screening below age 50 will be less consequential as, for instance, the USA in this regard.

The lifetime risk of being overdiagnosed will increase when more mammograms are performed, although the rate below age 50 will not be as high as that in older ages, because the risk of competing causes of mortality is smaller.[38] The rate of false-negatives and biopsies will be smaller at younger than at older screening ages, whereas the rates of additional imaging will be higher.[20]

Cost-effectiveness of screening below age 50

Because of these increased risks, a cost-effectiveness analysis is needed before any decisions on screening below age 50 can be made. Schousboe *et al.* (2011), who studied the cost-effectiveness of screening below age 50, showed that mammography screening

of younger women in the USA would only be cost-effective in those groups that are at higher-than-average risk of developing breast cancer.[41] The cost-effectiveness ratio in the Netherlands, however, would probably be much more favourable, as was screening in women older than 50.[42–43] Future research should determine the exact cost-effectiveness ratio of screening women younger than 50.

Digital mammography screening

Since 2010, conventional screen-film mammography has been replaced by digital mammography in all Dutch screening regions. The benefits in terms of breast cancer mortality prevention and harms in terms of overdiagnosis were discussed in Chapter 2. The estimated effects of digital screening in this thesis were based on an observed increase in the detection of DCIS in the first 3 years of digital screening in the Netherlands, but the detection of small invasive breast cancers is also likely to rise. Between 2004 and 2007, the detection rate of small invasive cancers was significantly higher at digital screening: 4.7/ 1,000 screening examinations (*versus* 4.0/ 1,000 screen-film mammograms).[17] Because of the relatively favourable prognosis of these tumours, the predicted mortality reduction is likely to be higher than the 23.7% as predicted in Chapter 2. Additional risks of implementing digital screening include higher detection rates and a lower positive predictive value or referral recommendations.[17] Because the increase in harms is proportionally larger than the increase in benefits, the cost-effectiveness ratio of mammography screening is likely to be affected, but will probably not surpass the WHO threshold for cost-effectiveness, as cost-effectiveness of screen-film mammography screening the Netherlands used to be highly favourable.

Digital mammography screening for women younger than 50

Digital mammography screening may be particularly beneficial to women younger than 50, who, on average, have denser breasts than older women. Because digital screening has the potential to increase contrast resolution in mammograms of dense breasts, abnormalities may become more visible. It should be noted that women below 50 may have a higher risk of false-positive mammograms than women over 50, because of the higher breast density. Biopsy rates, on the contrary, were found to be lower in younger women than in older women. The number of women needed to screen to prevent 1 breast cancer death is likely to be higher than in a screening program that targets women from age 50. Future research should investigate the potential benefits, harms and cost-effectiveness of digital screening for women younger than 50.

Unravelling the natural course of DCIS

Estimating the harms of (digital) screening is limited by the largely unknown and heterogeneous natural history of DCIS. For example, if DCIS would be a mainly progressive

disease, the detection of such lesion would prevent the development of invasive cancers, and possibly, breast cancer death. If DCIS would be dormant or regress, the detection of DCIS would mean that a woman would be unnecessarily treated. Unravelling the course of the disease would therefore mean a large step forward in the reduction of the harms of screening. A relevant aspect would be to develop a set of prognostic markers that could identify those DCIS lesions that are at high risk to recur later in life, and those that are not. If the latter group could be treated more conservatively, much of the harm involved with detecting DCIS would be averted.

Developing prognostic markers

Various attempts have been made to identify a set of prognostic markers for DCIS. The presence of comedonecrosis, focality, surgical margin width, method of detection, nuclear grade, lesion size,[46] growth pattern,[47–48] younger age, and not being treated with radiation therapy[49] were all found to be significantly associated with breast cancer recurrence in a meta-analysis of studies on DCIS. Estrogen (ER) and progesterone (PR) receptor negativity and Human epidermal growth factor-2 oncoprotein (HER2/neu) overexpression may also be also be related to recurrence, but the evidence is limited and inconsistent.[46]

Attempts to combine predictors into one prognostic index have not yet resulted in an instrument with sufficient discriminatory power. For instance, the Van Nuys Prognostic Index,[49] which considers lesion size, grade, margin width and age, may insufficiently distinguish recurring from non-recurring lesions.[50–51] Future research should therefore aim to further develop risk-stratification models to identify subsets of women who are at low or high risk for breast cancer recurrence or progression. Such research should also take into account that factors associated with the recurrence of an initial DCIS as a subsequent DCIS may be different from the factors associated with recurrence as an invasive breast cancer.[52] Once such models have been developed, an important research aim would be to study the consequences of modifying breast cancer treatment according to risk level in a randomized controlled trial, for instance by offering active surveillance only to low-risk women, and surgical intervention to those at high risk.

8.6 MAIN CONCLUSIONS

- High quality mammography screening in women aged 50–74 years is predicted to reduce breast cancer mortality by 23% in a population aged 19–100 years with at least 1 screening examination, compared to a situation without screening. An estimated 683 to 858 breast cancer deaths are annually prevented in the current screening situation.
- The risk of overdiagnosis in the Dutch breast cancer screening programme is 2.1% of all diagnosed breast cancers in women with at least 1 screening examination, or 7.2% of all screen-detected breast cancers.
- The implementation of digital mammography screening resulted in an 80% increase in the detection rate of DCIS. As a consequence, the number of breast cancer deaths that is prevented by screening was predicted to increase by 4.4%, in addition to the breast cancer deaths that were already prevented by screen-film mammography screening. This means that the reduction in breast cancer mortality would increase from 23% to 24%. The risk of overdiagnosis would increase by 21%, to 2.5% of all diagnosed breast cancers in a population aged 0–100 years with at least 1 screening examination.
- Overdiagnosis estimates in literature vary between 1% and 54%. Such variations can be ascribed to insufficient follow-up to allow for lead time, and differences in the choice of the estimator that is used to define the population at risk for overdiagnosis.
- Breast cancer patients who had participated in the Dutch breast cancer screening programme prior to noticing symptoms did not have a significantly longer period between the initial discovery of breast abnormalities and the first time presenting these to a doctor than women who had not been screened. False reassurance therefore plays, at most, only a minor role in breast cancer screening.
- In a specific country with decentralized screening, opportunistic mammography was predicted to be comparably effective as organized mammography screening, although the costs would be twice as high, related to the frequency with which additional diagnostic examinations were performed.
- In a population aged 19–100 years, adjuvant systemic therapy was estimated to reduce breast cancer mortality by 14%, compared to a situation without such treatment or screening. Biennial mammography screening between age 50 and 74 would further reduce breast cancer mortality by 16%. If screening would then be extended with one additional screening examination at age 48, the breast cancer mortality rate would further decrease by 1.0%. Ten additional mammograms between age 40 and 49 could reduce breast cancer mortality by an extra 5.1%, compared to biennial screening between age 50 and 74.

- The benefits of mammography screening such as currently practiced are not lower than had been observed in randomized controlled trials.
- Biennial screening between age 50 and 74 was estimated to cause 7.7 breast cancers and 1.6 breast cancer deaths per 100,000 women aged 0–100 years, but would also prevent 1121 breast cancer deaths. If biennial screening would start at age 40, 3.7 breast cancer deaths would be induced, but 1302 breast cancer deaths would be averted. The benefits of mammography screening in women aged 50–74 years thus strongly outweigh the radiation risks, also when the lower age limit for screening would be lowered to age 40.

8.7 RECOMMENDATIONS FOR FUTURE RESEARCH AND PRACTICE

- Since 2010, all screening units in the Netherlands have replaced screen-film with digital mammography. As a consequence, the risk of overdiagnosis may increase. The effects of digital mammography screening therefore remain to be carefully monitored.
- The long-term consequences of the increased detection of DCIS are uncertain, because the natural history of DCIS is largely unknown. Some lesions may never progress to invasive breast cancer if they are left untreated, and some may not recur when they are treated with local excision only. If we would know in advance which cases are relatively harmless, the harms involved with treating these cases could be avoided. Future research should therefore aim to develop stratification models that could allocate DCIS into subgroups at low and high risk for progression/ recurrence. As a second step, research should determine which subgroups could be treated with interventions that are less invasive than those that are currently used, and those who would need more intensive treatment.
- Strong variations in overdiagnosis rates have been observed, which could be related to methodological differences between studies. To increase comparability, researchers should agree on a uniform method to calculate overdiagnosis. The observed differences in overdiagnosis estimates were also related to the moment at which the risk had been calculated. Future studies that assess the risk of overdiagnosis in screening programmes should therefore allow for sufficient follow-up to account for the effect of lead time.
- No evidence for false reassurance was observed in the Dutch screening programme. This does not necessarily mean that the same is true for other (breast cancer) screening programmes. The risk of false reassurance may, for instance, depend on pre-existing knowledge about the disease or the information that is provided to potential screening participants, which could be inferior to that in the Netherlands. In the evaluation of screening programmes, it is therefore recommended to assess the

role of false reassurance. Furthermore, providing women with valid information about the risk of false-negative screening outcomes and the possibility that symptoms may develop in the period between 2 screening examinations, is relevant to minimize the risk of delay and advanced disease.

- Although organized and opportunistic mammography screening were predicted to be equally effective, the cost of opportunistic screening would be twice as high. Strong improvements can be obtained if the use of imaging diagnostics would be diminished and costs were decreased. Centralizing screening would enhance the possibilities of multiple readings, discussion and feedback, which should increase the screening volume and performance. It would also enable a more effective use of equipment and a faster acquisition of work-up diagnostics. Furthermore, continuous quality control and evaluation of screening could ensure that maximum benefits are obtained at reasonable costs.
- Both adjuvant systemic treatment and mammography screening played a significant role in reducing breast cancer mortality. The observed mortality reduction in the Netherlands, however, could not be fully explained by these factors alone. A possible explanation may be that the effects of screening or adjuvant therapy have been under-estimated, because recent developments such as the use of aromatase inhibitors or digital screening were not taken into account. It may also be the case that general advances in diagnostics and primary treatment for breast cancer and increased awareness have played an additional role. Future research should further elicit the role of screening and adjuvant systemic therapy in the observed reduction of breast cancer mortality.
- In the Netherlands, a cost-effectiveness analysis that includes the increased costs for false-positive and false-negative mammograms, additional diagnostics, overdiagnosis, and radiation-induced breast cancer is needed before decisions on the feasibility of screening below age 50 can be made.
- Digital mammography may be particularly beneficial for (subgroups of) women younger than 50 years. A study on the long-term effects of digital mammography screening in this age group has not yet been performed, and is therefore recommended.
- Women who are invited to participate in screening need to be fully informed on the benefits and harms of having a mammogram. Future research should elicit how women balance benefits and risks, and what their specific preferences and needs are with regard to informed decision making.

REFERENCES

1. Autier P, Boniol M, Gavin A, Vatten LJ. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. BMJ (Clinical research ed 2011;343:d4411.
2. Jorgensen KJ, Gotzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. BMJ (Clinical research ed 2009;339:b2587.
3. Kalager M, Zelen M, Langmark F, Adami HO. Effect of screening mammography on breast-cancer mortality in Norway. N Engl J Med 2010;363:1203–10.
4. Jorgensen KJ, Zahl PH, Gotzsche PC. Breast cancer mortality in organised mammography screening in Denmark: comparative study. BMJ (Clinical research ed 2010;340:c1241.
5. Zahl PH, Strand BH, Maehlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study. BMJ (Clinical research ed 2004;328:921–4.
6. Blanks RG, Moss SM, McGahan CE, Quinn MJ, Babb PJ. Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990–8: comparison of observed with predicted mortality. BMJ (Clinical research ed 2000;321:665–9.
7. Tabar L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Grontoft O, Ljungquist 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 and Welfare. Lancet 1985;1:829–32.
8. Jonsson H, Nystrom L, Tornberg S, Lenner P. Service screening with mammography of women aged 50–69 years in Sweden: effects on mortality from breast cancer. J Med Screen 2001;8:152–60.
9. Nystrom 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;359:909–19.
10. 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. BMJ (Clinical research ed 2005;330:220.
11. Jorgensen KJ, Zahl PH, Gotzsche PC. Overdiagnosis in organised mammography screening in Denmark. A comparative study. BMC women's health 2009;9:36.
12. Morrell S, Barratt A, Irwig L, Howard K, Biesheuvel C, Armstrong B. Estimates of overdiagnosis of invasive breast cancer associated with screening mammography. Cancer Causes Control 2010;21:275–82.
13. van der Maas PJ, de Koning HJ, van Ineveld BM, van Oortmarssen GJ, Habbema JD, Lubbe KT, Geerts AT, Collette HJ, Verbeek AL, Hendriks JH, *et al.* The cost-effectiveness of breast cancer screening. Int J Cancer 1989;43:1055–60.
14. de Koning HJ, Boer R, Warmerdam PG, Beemsterboer PM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. J Natl Cancer Inst 1995;87:1217–23.
15. van der Maas PJ. [Population screening for breast cancer: an interim assessment]. Nederlands tijdschrift voor geneeskunde 2000;144:1096–9.
16. van der Maas PJ. Breast cancer screening programme in The Netherlands: an interim review. Breast 2001;10:12–4.

17. National Evaluation Team for Breast cancer screening (NETB), National evaluation of breast cancer screening in the Netherlands 1990–2007. Twelfth evaluation report. Erasmus MC University Medical Center Rotterdam, Radboud University Nijmegen Medical Centre, 2009.
18. Cronin KA, Feuer EJ, Clarke LD, Plevritis SK. Impact of adjuvant therapy and mammography on U.S. mortality from 1975 to 2000: comparison of mortality results from the cisnet breast cancer base case analysis. J Natl Cancer Inst Monogr 2006:112–21.
19. Fletcher SW. Breast cancer screening: a 35-year perspective. Epidemiol Rev 2011;33:165–75.
20. Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L, Force USPST. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. Ann Intern Med 2009;151:727–37, W237–42.
21. Schopper D, de Wolf C. How effective are breast cancer screening programmes by mammography? Review of the current evidence. Eur J Cancer 2009;45:1916–23.
22. Paap E, Verbeek AL, Puliti D, Paci E, Broeders MJ. Breast cancer screening case-control study design: impact on breast cancer mortality. Ann Oncol 2011;22:863–9.
23. van Ravesteyn NT, Heijnsdijk EA, de Koning HJ. More on screening mammography. N Engl J Med 2011;364:282–3; author reply 5–6.
24. De Koning HJ. Screen evaluation without evaluating screening? Rapid response to Autier *et al.*: Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO mortality database. BMJ (Clinical research ed 2011;343:d4411.
25. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane database of systematic reviews (Online) 2009:CD001877.
26. 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;49:538–44.
27. International Agency for Research on Cancer. Volume 7: Breast Cancer Screening. In: Vainio H, Bianchini F. IARC Handbooks of Cancer Preventioned. Lyon: IARCPress, 2002.
28. van der Steeg AF, Keyzer-Dekker CM, De Vries J, Roukema JA. Effect of abnormal screening mammogram on quality of life. The British journal of surgery 2011;98:537–42.
29. Peasgood T, Ward SE, Brazier J. Health-state utility values in breast cancer. Expert review of pharmacoeconomics & outcomes research 2010;10:553–66.
30. Tabar L, Vitak B, Chen HH, Duffy SW, Yen MF, Chiang CF, Krusemo UB, Tot T, Smith RA. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. Radiol Clin North Am 2000;38:625–51.
31. Bjurstam N, Bjorneld L, Warwick J, Sala E, Duffy SW, Nystrom L, Walker N, Cahlin E, Eriksson O, Hafstrom LO, Lingaas H, Mattsson J, *et al.* The Gothenburg Breast Screening Trial. Cancer 2003;97:2387–96.
32. Louwman WJ, Voogd AC, van Dijck JA, Nieuwenhuijzen GA, Ribot J, Pruijt JF, Coebergh JW. On the rising trends of incidence and prognosis for breast cancer patients diagnosed 1975–2004: a long-term population-based study in southeastern Netherlands. Cancer Causes Control 2008;19:97–106.
33. Otten JD, Broeders MJ, Fracheboud J, Otto SJ, de Koning HJ, Verbeek AL. Impressive time-related influence of the Dutch screening programme on breast cancer incidence and mortality, 1975–2006. Int J Cancer 2008;123:1929–34.

34. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet 2006;368:2053–60.
35. Hellquist BN, Duffy SW, Abdsaleh S, Bjorneld L, Bordas P, Tabar L, Vitak B, Zackrisson S, Nystrom L, Jonsson H. Effectiveness of population-based service screening with mammography for women ages 40 to 49 years: evaluation of the Swedish Mammography Screening in Young Women (SCRY) cohort. Cancer 2011;117:714–22.
36. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2009;151:716–26, W–236.
37. Qaseem A, Snow V, Sherif K, Aronson M, Weiss KB, Owens DK. Screening mammography for women 40 to 49 years of age: a clinical practice guideline from the American College of Physicians. Ann Intern Med 2007;146:511–5.
38. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, Huang H, Lee SJ, Munsell M, Plevritis SK, Ravdin P, Schechter CB, *et al.* Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Intern Med 2009;151:738–47.
39. Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. N Engl J Med 1998;338:1089–96.
40. Salas D, Ibanez J, Roman R, Cuevas D, Sala M, Ascunce N, Zubizarreta R, Castells X. Effect of start age of breast cancer screening mammography on the risk of false-positive results. Preventive medicine 2011;53:76–81.
41. Schousboe JT, Kerlikowske K, Loh A, Cummings SR. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. Ann Intern Med 2011;155:10–20.
42. Groenewoud JH, Otten JD, Fracheboud J, Draisma G, van Ineveld BM, Holland R, Verbeek AL, de Koning HJ. Cost-effectiveness of different reading and referral strategies in mammography screening in the Netherlands. Breast cancer research and treatment 2007;102:211–8.
43. Tosteson AN, Stout NK, Fryback DG, Acharyya S, Herman BA, Hannah LG, Pisano ED, Investigators D. Cost-effectiveness of digital mammography breast cancer screening. Ann Intern Med 2008;148:1–10.
44. World Health Organization. Macroeconomics and health: investing in health for economic development: report of the Commission on Macroeconomics and Healthed. Geneva, 2001.
45. International Monetary Fund. World Economic and Financial Surveys, World Economic Outlook Database, http://www.imf.org 2010.
46. Wang SY, Shamliyan T, Virnig BA, Kane R. Tumor characteristics as predictors of local recurrence after treatment of ductal carcinoma *in situ*: a meta-analysis. Breast cancer research and treatment 2011;127:1–14.
47. Pinder SE, Duggan C, Ellis IO, Cuzick J, Forbes JF, Bishop H, Fentiman IS, George WD. A new pathological system for grading DCIS with improved prediction of local recurrence: results from the UKCCCR/ANZ DCIS trial. Br J Cancer 2010;103:94–100.
48. Bijker N, Peterse JL, Duchateau L, Julien JP, Fentiman IS, Duval C, Di Palma S, Simony-Lafontaine J, de Mascarel I, van de Vijver MJ. Risk factors for recurrence and metastasis after breast-conserving therapy for ductal carcinoma-in-situ: analysis of European Organization for Research and Treatment of Cancer Trial 10853. J Clin Oncol 2001;19:2263–71.
49. Silverstein MJ, Lagios MD, Craig PH, Waisman JR, Lewinsky BS, Colburn WJ, Poller DN. A prognostic index for ductal carcinoma *in situ* of the breast. Cancer 1996;77:2267–74.

50. Boland GP, Chan KC, Knox WF, Roberts SA, Bundred NJ. Value of the Van Nuys Prognostic Index in prediction of recurrence of ductal carcinoma *in situ* after breast-conserving surgery. The British journal of surgery 2003;90:426–32.
51. MacAusland SG, Hepel JT, Chong FK, Galper SL, Gass JS, Ruthazer R, Wazer DE. An attempt to independently verify the utility of the Van Nuys Prognostic Index for ductal carcinoma *in situ*. Cancer 2007;110:2648–53.
52. Kerlikowske K, Molinaro AM, Gauthier ML, Berman HK, Waldman F, Bennington J, Sanchez H, Jimenez C, Stewart K, Chew K, Ljung BM, Tlsty TD. Biomarker expression and risk of subsequent tumors after initial ductal carcinoma *in situ* diagnosis. J Natl Cancer Inst 2010;102:627–37.

Summary

SUMMARY

Introduction

Breast cancer is the most common cancer among women in Western countries. Presently, women in the Netherlands have a 1 : 7 chance of developing breast cancer during their lifetime. Although the probability of cure has improved over the last decennia, for a third of all women with breast cancer, the disease will be fatal. This makes breast cancer the most common cause of cancer death in women in Europe.

Randomized controlled trials have shown that early detection of breast cancer by mammography screening can reduce breast cancer mortality. In a population setting, however, the benefits of mammography may be different. Several recent studies suggest that breast cancer screening has no, or only small effects on breast cancer mortality. Moreover, an increasing number of studies argue that the harms of mammography may be larger than previously assumed.

This thesis therefore assessed the benefits of breast cancer screening, in particular the benefits in terms of breast cancer mortality reduction and gain of life years. The harms of screening, with regard to overdiagnosis and false-reassurance, were also estimated, and the costs and cost-effectiveness of two types of mammography screening was calculated (**Part 1**). We focussed on possible future directions of screening, such as lowering the age limit of screening to include women younger than 50 years (**Part 2**). **Part 3** of this thesis (General discussion) balanced benefits and harms, and focuses on current breast cancer screening debates.

To assess the effects of screening, the micro-simulation analysis model MISCAN was used. The model can predict breast cancer incidence, tumour stage distribution, breast cancer mortality and life expectancy in the absence of screening, and compare these measures to a modelled situation with screening.

Effects and risks of the current breast cancer screening programme

The first study objective was to assess the mortality benefits and overdiagnosis risk of population-based mammography screening in the Netherlands (**Chapter 2**). The effects and risks of screen-film mammography screening were then compared to those of digital mammography screening, which has replaced conventional screen-film mammography since 2010. In women with at least 1 screening examination, screen-film mammography screening was predicted to reduce breast cancer mortality in by 23%. The fraction over-diagnosed tumours of all diagnosed breast cancers was 2.1%. Based on the observed increase in the detection of ductal carcinoma *in situ* (DCIS), we estimated that digital

mammography screening would further reduce breast cancer mortality by 4.4%, at a 21% increased overdiagnosis rate. The consequences of digital screening, however, were strongly dependent on underlying model assumptions on the natural history of DCIS.

The estimated risk of overdiagnosis strongly varies in literature, with estimates ranging from 1% to 54%. **Chapter 3** aimed to explains such variations. Using the gradual implementation of mammography screening in the Netherlands as an example, we estimated that such differences could largely be explained by the use of different estimators to calculate overdiagnosis, such as for instance women of all ages, women in the screening age only, or screen-detected cancers. Differences can also be explained by the fact that several studies based their analyses on early screening phases, with which they insufficiently accounted for lead time. In 2006, when screening in the Netherlands was in a steady-state, the overdiagnosis rate was 2.8%, as compared to all predicted cancers in women aged 0–100 years in the absence of screening.

The risk of false reassurance was assessed in **Chapter 4**. False reassurance was defined as diagnostic delay due to having participated in screening. It could occur after a negative screen result, when a patient or doctor, perceiving the risk of developing cancer to be small, is consequently less alert to symptoms when they occur. Possible consequences are increased tumour sizes, more positive lymph nodes and decreased long-term survival. In the present study, breast cancer patients who had participated in the Dutch screening programme did not have a significantly longer period between the initial discovery of breast abnormalities and first consultation of a general practitioner than patients who had not been screened. Only two women with a previous screening history had relatively long periods between the initial symptoms and first GP visit: 2.5 and 4 years, respectively. This study therefore showed that false reassurance played, at most, only a minor role in breast cancer screening.

The aim of **Chapter 5** was to assess how the organizational form of breast cancer screening affects its benefits and costs. Various centralized screening programmes have been shown to reduce breast cancer mortality at reasonable costs, but opportunistic screening, such as for instance performed in Switzerland, may under certain circumstances be a cost-effective alternative. Using MISCAN, we predicted that organized and opportunistic screening were comparably effective, and could reduce breast cancer mortality by 13% in the total simulated population (or by 20% in women aged 55–74 years). For each prevented breast cancer death, 16.5 life years were gained. The cost-effectiveness ratio of opportunistic screening, on the other hand, was twice less favourable than that of organized screening. This was mainly related to high costs of opportunistic mammograms and the frequency with which additional imaging diagnostics were used.

Future directions of mammography screening

In **Chapter 6**, the effects of mammography screening were compared to the effects of adjuvant systemic therapy. Randomized controlled trials have shown that both interventions reduce breast cancer mortality, but in a population setting the effects of screening may interact with those of adjuvant treatment, and vice versa. Particular attention was paid to the effects of screening below age 50, thereby taking into account possible interaction with adjuvant therapy. By using micro-simulation, we estimated the effects of adjuvant therapy, the additional effect of biennial screening between age 50 and 74 (current screening programme), and the effect of extending the current screening programme with 1–10 extra examinations between age 40 and 50. Adjuvant treatment was estimated to have reduced the breast cancer mortality rate in the simulated population in 2008 by 13.9%, compared to a situation without treatment. Biennial screening between age 50 and 74 further reduced the mortality rate by 15.7%. Advancing the lower age limit of the screening programme to age 48 would reduce the mortality rate by 1.0% compared to screening from age 50; 10 additional screening rounds between age 40 and 49 would lower this rate by 5.1%.

Chapter 7 aimed to estimate the radiation risks of breast cancer screening. Because the radiation risk increases with lower exposure age, particular attention was paid to radiation risks involved with advancing the lower age limit of screening to ages younger than 50. Using MISCAN, we compared the risks with the potential benefits of screening starting between age 40 and 50. Our calculations were based on the latest estimates of glandular dose and the most recent excess absolute rate exposure-risk model. The study showed that the estimated radiation risk is lower than previously assessed. At a mean glandular dose of 1.3 mGy per view, biennial mammography screening between age 50 and 74 was predicted to induce 1.6 breast cancer deaths per 100,000 women aged 0–100, against 1121 avoided deaths in this population. Advancing the lower screening age limit to 40 (5 additional screening examinations) was predicted to induce 3.7 breast cancer deaths per 100,000 women aged 0–100, but would also prevent 1302 deaths. The benefits of mammography screening between age 40 and 74 were therefore predicted to outweigh the radiation risks.

General discussion

The answers to the research questions were discussed in **Chapter 8**. Furthermore, the various model predictions of all chapters were combined and interpreted. We compared our findings to other screening studies, including those who reported small effects and large risks of mammography screening. Attention was paid to several methodological issues that may be involved with using simulation models.

Our main conclusion was that high quality mammography screening in women aged 50–74 years can reduce breast cancer mortality by 23%, measured in a population aged 19–100 years with at least 1 screening examination, compared to a situation without screening. In the current screening situation, between 683 and 858 breast cancer deaths are annually prevented. Even in a situation in which many breast cancer patients are treated with adjuvant systemic therapy, screening effects remain substantial. Advancing the lower age limit for screening could further reduce breast cancer mortality, without substantially increasing the radiation risk.

The risk of overdiagnosis in the Dutch breast cancer screening programme is small, and was estimated to be 2.1% of all diagnosed breast cancers in women with at least 1 screening examination, or 7.2% of all screen-detected breast cancers. The overdiagnosis risk was predicted to increase as a consequence of implementing digital screening, because of increased detection of DCIS. The effects of digital screening should therefore remain to be carefully monitored. Future research should aim to develop stratification models that could allocate DCIS into subgroups at low and high risk for tumour progression or recurrence, identifying those that could be treated with interventions that are less aggressive than those that are currently used, and those who would need more intensive treatment.

Samenvatting

SAMENVATTING

Inleiding

Borstkanker is de meest voorkomende kanker bij vrouwen in westerse landen. Tegenwoordig hebben vrouwen in Nederland een kans van 1 : 7 om tijdens hun leven borstkanker te ontwikkelen. Hoewel de kans op genezing de afgelopen decennia is toegenomen, is voor een derde van alle vrouwen met borstkanker de ziekte fataal. Hierdoor is borstkanker de meest voorkomende oorzaak van kankersterfte bij vrouwen in Europa.

Gerandomiseerde trials hebben aangetoond dat vroege ontdekking van borstkanker door screening borstkankersterfte kan voorkomen. In de werkelijke populatie kunnen de effecten van borstkankerscreening echter anders zijn. Verschillende recente studies lijken aan te tonen dat screening geen, of slechts een klein effect heeft op borstkankersterfte. Bovendien suggereert een toenemend aantal onderzoeken dat de nadelen van borstkankerscreening groter zijn dan voorheen aangenomen werd.

In dit proefschrift werden daarom de effecten van borstkankerscreening onderzocht, met name de effecten op borstkankersterfte en gewonnen levensjaren. De risico's van screening met betrekking tot overdiagnose en 'onterechte geruststelling' werden ook geschat, en de kosten en kosten-effectiviteit van twee types borstkankerscreening werden berekend **(Deel 1)**. Nadruk werd gelegd op mogelijke toekomstige ontwikkelingen op het gebied van screening, zoals het verlagen van de onderste leeftijdsgrens om vrouwen van onder de 50 te includeren (**Deel 2**). In **Deel 3** van dit proefschrift (Algemene discussie) worden effecten en risico's tegen elkaar afgewogen, en worden huidige screeningsdebatten belicht.

Om de effecten van screening te bepalen werd het microsimulatie model 'MISCAN' gebruikt. Het model kan borstkankerincidentie, stadiumverdeling, borstkankersterfte en levensverwachting in de afwezigheid van screening voorspellen, en kan deze maten vergelijken met een situatie met screening.

Effecten en risico's van het huidige bevolkingsonderzoek

De eerste onderzoeksdoelstelling was het bepalen van de sterftereductie en het risico op overdiagnose in het Nederlands bevolkingsonderzoek naar borstkanker (**Hoofdstuk 2**). De effecten en risico's van analoge mammografie werden vervolgens vergeleken met die van digitale screening, dat sinds 2010 de conventionele analoge screening vervangen heeft. Bij vrouwen met tenminste 1 screeningsonderzoek werd voorspeld dat analoge mammografie borstkankersterfte met 23% heeft teruggebracht. De fractie overgediagnosticeerde tumoren bedroeg 2,1% van alle gediagnosticeerde borstkankers.

Op basis van de waargenomen toename in de detectie van ductaal carcinoma *in situ* (DCIS), schatten we dat digitale mammografie de borstkankersterfte verder zou kunnen terugbrengen met 4,4%, terwijl het overdiagnosecijfer met 21%zou toenemen. De consequenties van digitale screening zijn echter sterk afhankelijk van onderliggende modelaannames over het natuurlijk verloop van DCIS.

Het geschatte risico op overdiagnose varieert sterk in de literatuur, met schattingen die variëren tussen 1% en 54%. In **Hoofdstuk 3** worden deze verschillen verklaard. Op basis van de geleidelijke invoering van borstkankerscreening in Nederland schatten we dat zulke verschillen grotendeels verklaard kunnen worden door het gebruik van verschillende schatters om overdiagnose te berekenen, zoals bijvoorbeeld vrouwen van alle leeftijden, vrouwen in alleen de screeningsleeftijd, of screeningscarcinomen. Verschillen kunnen ook verklaard worden door het feit dat verschillende studies hun uitkomsten baseerden op vroege screeningsfasen, waardoor zij onvoldoende rekening met 'lead time' hielden. In 2006, toen screening in Nederland in een stabiele fase verkeerde, was het overdiagnosecijfer in Nederland 2,8%, ten opzichte van alle voorspelde kankers bij vrouwen in de leeftijd 0–100 jaar in de afwezigheid van screening.

Het risico op onterechte geruststelling werd bepaald in **Hoofdstuk 4**. Onterechte geruststelling werd gedefinieerd als uitstel van diagnose door deelname aan het bevolkingsonderzoek. Het kan optreden na een negatieve uitslag van een screeningsonderzoek, wanneer een arts of patiënt het risico op het ontwikkelen van kanker als klein ervaart, en daardoor minder alert is op symptomen als deze optreden. Mogelijke gevolgen zijn toegenomen tumorgrootte, meer positieve lymfeklieren en afgenomen lange termijn overleving. In de huidige studie hadden borstkankerpatiënten die met het Nederlands bevolkingsonderzoek hadden meegedaan geen significant langere periode tussen de eerste ontdekking van borstafwijkingen en het eerste consult bij een huisarts dan patiënten die niet waren gescreend. Slechts twee vrouwen die eerder gescreend waren hadden een relatief lange periode tussen de eerste klachten en het eerste huisartsconsult: respectievelijk 2,5 en 4 jaar. Deze studie toonde dus aan dat onterechte geruststelling op zijn hoogst slechts een kleine rol speelt in borstkankerscreening.

Het doel van **Hoofdstuk 5** was te bepalen hoe de organisatorische vorm van borstkankerscreening de voor- en nadelen beïnvloedt. Van verschillende centraal georganiseerde screeningsprogramma's is aangetoond dat zij de borstkankersterfte tegen redelijke kosten kunnen terugbrengen, maar opportunistische screening, zoals dat bijvoorbeeld wordt uitgevoerd in Zwitserland, kan onder bepaalde omstandigheden een kosteneffectief alternatief zijn. Met MISCAN voorspelden we dat georganiseerde en opportunistische screening even effectief zijn, en borstkankersterfte in de totale gesimuleerde populatie

met 13% kunnen reduceren (of met 20% bij vrouwen in de leeftijd 55–74 jaar). Voor ieder voorkomen sterfgeval aan borstkanker kunnen 16,5 levensjaren gewonnen worden. De kosteneffectiviteitsratio van opportunistische screening was daarentegen twee keer minder gunstig dan die van georganiseerde screening. Dit was voornamelijk gerelateerd aan de hoge kosten van een opportunistisch mammogram en de frequentie waarmee extra beeldvormende diagnostiek werd gebruikt.

Toekomstige ontwikkelingen in borstkankerscreening

In **Hoofdstuk 6** werden de effecten van screening vergeleken met de effecten van adjuvante systemische behandeling van borstkanker. Gerandomiseerde trials hebben aangetoond dat beide interventies borstkankersterfte verlagen, maar in een daadwerkelijke populatie kunnen de effecten van screening die van adjuvante therapie beïnvloeden, en andersom. Bijzondere aandacht werd besteed aan de effecten van screening bij vrouwen die jonger zijn dan 50 jaar, waarbij mogelijke interactie met adjuvante therapie in beschouwing werd genomen. Door middel van microsimulatie konden we de effecten van adjuvante therapie, het toegevoegde effect van tweejaarlijkse screening bij 50 tot 74 jarige vrouwen (het huidige screeningsprogramma), en het effect van het uitbreiden van het huidige screeningsprogramma met 1–10 extra screeningsonderzoeken tussen leeftijd 40 en 50 schatten. We schatten dat adjuvante therapie het sterftecijfer aan borstkanker in de totale populatie in 2008 met 13,9% verlaagd heeft. Tweejaarlijkse screening tussen leeftijd 50 en 74 bracht het sterftecijfer verder terug met 15,7%. Het verlagen van de onderste leeftijdsgrens van het screeningsprogramma tot 48 jaar zou het sterftecijfer aan borstkanker met 1.0% reduceren ten opzichte van screening vanaf 50 jaar; 10 extra screeningsronden tussen leeftijd 40 en 49 zouden dit cijfer met 5.1% terugbrengen.

Hoofdstuk 7 had als doel de stralingsrisico's van borstkankerscreening te schatten. Omdat het risico van straling toeneemt met een jongere leeftijd van blootstelling, werd nadruk gelegd op de risico's die gepaard gaan met het verlagen van de onderste leeftijdsgrens voor screening met leeftijden onder de 50. Met MISCAN vergeleken we de risico's met mogelijke gunstige effecten van het starten van screening tussen 40 en 50-jarige leeftijd. Onze berekeningen waren gebaseerd op de laatste schattingen van geabsorbeerde dosis en het meest recente 'extra absoluut risico' blootstelling-risicomodel. De studie toonde aan dat geschatte stralingsrisico lager is dan voorheen werd aangenomen. Bij een gemiddelde glandulaire dosis van 1,3 mGy per opname voorspelden we dat per 100.000 vrouwen in de leeftijd 0 tot 100 jaar tweejaarlijkse screening tussen leeftijd 50 en 74 1,6 sterfgevallen aan borstkanker zou veroorzaken, tegen 1121 voorkomen sterfgevallen in dezelfde populatie. Het verlagen van de onderste leeftijdsgrens tot 40 (5 extra screeningsonderzoeken) zou per 100.000 vrouwen in de leeftijd 0 tot 100 jaar 3,7

sterfgevallen aan borstkanker veroorzaken, maar zou ook 1302 sterfgevallen voorkomen. We voorspelden daarom dat de voordelen van borstkankerscreening tussen leeftijd 40 en 74 jaar zwaarder wegen dan de stralingsrisco's.

Algemene discussie

De antwoorden op de onderzoeksvragen werden bediscussieerd in **Hoofdstuk 8**. Daarnaast werden de verschillende modelvoorspellingen uit alle hoofdstukken gecombineerd en geïnterpreteerd. We vergeleken onze bevindingen met andere screeningsstudies, onder andere met studies die kleine effecten en grote risico's van borstkankerscreening rapporteerden. Aandacht werd besteed aan methodologische problemen die gepaard kunnen gaan met het gebruik van simulatiemodellen.

Onze voornaamste conclusie was dat, wanneer een hoge kwaliteit screening bereikt wordt, het bevolkingsonderzoek bij vrouwen van 50 tot 74 jaar borstkankersterfte met 23% kan terugbrengen, gemeten in een populatie van 19–100 jaar met tenminste 1 screeningsonderzoek, ten opzichte van een situatie zonder screening. In de huidige screeningssituatie worden jaarlijks 683 tot 858 sterfgevallen aan borstkanker voorkomen. Zelfs in een situatie waarbij veel patiënten behandeld worden met adjuvante systemische therapie blijven screeningseffecten substantieel. Het vervroegen van de onderste leeftijdsgrens van het bevolkingsonderzoek kan borstkankersterfte verder verminderen, zonder dat daarbij het stralingsrisico sterk toeneemt.

Het risico op overdiagnose in het Nederlands bevolkingsonderzoek is klein, en werd geschat op 2,1% ten opzichte van alle gediagnosticeerde kankers bij vrouwen met tenminste 1 screeningsonderzoek, of 7,2% van alle screeningscarcinomen. We voorspelden dat het overdiagnoserisico zal toenemen als gevolg van het implementeren van digitale screening, vanwege toegenomen detectie van DCIS. De effecten van digitale screening moeten daarom zorgvuldig in de gaten worden gehouden. Toekomstig onderzoek zou zich moeten richten op het ontwikkelen van stratificatiemodellen die DCIS onderverdelen in subgroepen met lagere en hogere kans op progressie en recidivering, waarmee patiënten geïdentificeerd kunnen worden die behandeld kunnen worden met therapieën die minder agressief zijn dan de huidige, en patiënten die intensievere behandelingen nodig hebben.

Annex

ANNEX

The consequences of screening on population health can be assessed using micro-simulation models, such as the MIcrosimulation SCreening ANalysis model 'MISCAN' that was developed at the Erasmus MC.[1–2] With MISCAN, the results of randomized trials can be extrapolated to different screening ages, intervals and (improved) tests. The model consists of a part that simulates the demography of the population under study, a part that simulates the natural history of breast cancer, and a part that models the influence of screening on this natural history.

The demography part of the model first simulates individual life histories without breast cancer. For each person, a date of birth and a date of death from other causes than breast cancer are simulated. The distribution of births and deaths over calendar time can be adjusted to represent the population under study.

The natural history part of MISCAN simulates the development of breast cancer in the population. A graphical representation of the natural history in the model is given in Figure A.1. A certain percentage of women develop a preclinical screen-detectable breast cancer, which can become symptomatic. MISCAN is a semi-Markov model that simulates the growth of breast cancer as a progression through discrete disease stages. Without screening, preclinical breast cancer may develop through the invasive preclinical stages T1a (≤5 mm), T1b (6–10 mm), T1c (11–20 mm) and T2+ (>20 mm), with or without invasion to regional or distant lymph nodes. A certain fraction of the pre-clinical invasive cancers is preceded by pre-clinical screen-detectable DCIS. Preclinical DCIS may also regress. In every stage, the cancer may be diagnosed because of symptoms. The durations of each preclinical stage are assumed to be exponentially distributed. The natural history parameters were estimated using data from the Dutch Cancer Registry and screening results.[3–4] These data include age- and calendar year specific incidence rates of DCIS and invasive breast cancer; the age-, stage- and calendar year specific incidence of clinically diagnosed and screen-detected breast cancer; age-, stage, calendar year and detection-round specific detection rounds of breast cancer and age-, stage-, calendar year and detection-round specific rates of interval cancers. Survival after clinical diagnosis depends on the stage of the cancer and was derived from several international sources.[5–10]

Screening interrupts the development of breast cancer. With screening, breast cancers may be detected at an earlier stage than in a situation without screening. In this way, the probability of survival may increase. In MISCAN, screening characteristics of the programme under study, such as the test that is used, screening ages, intervals and attendance rates are specified. The sensitivity of the test, which determines how many

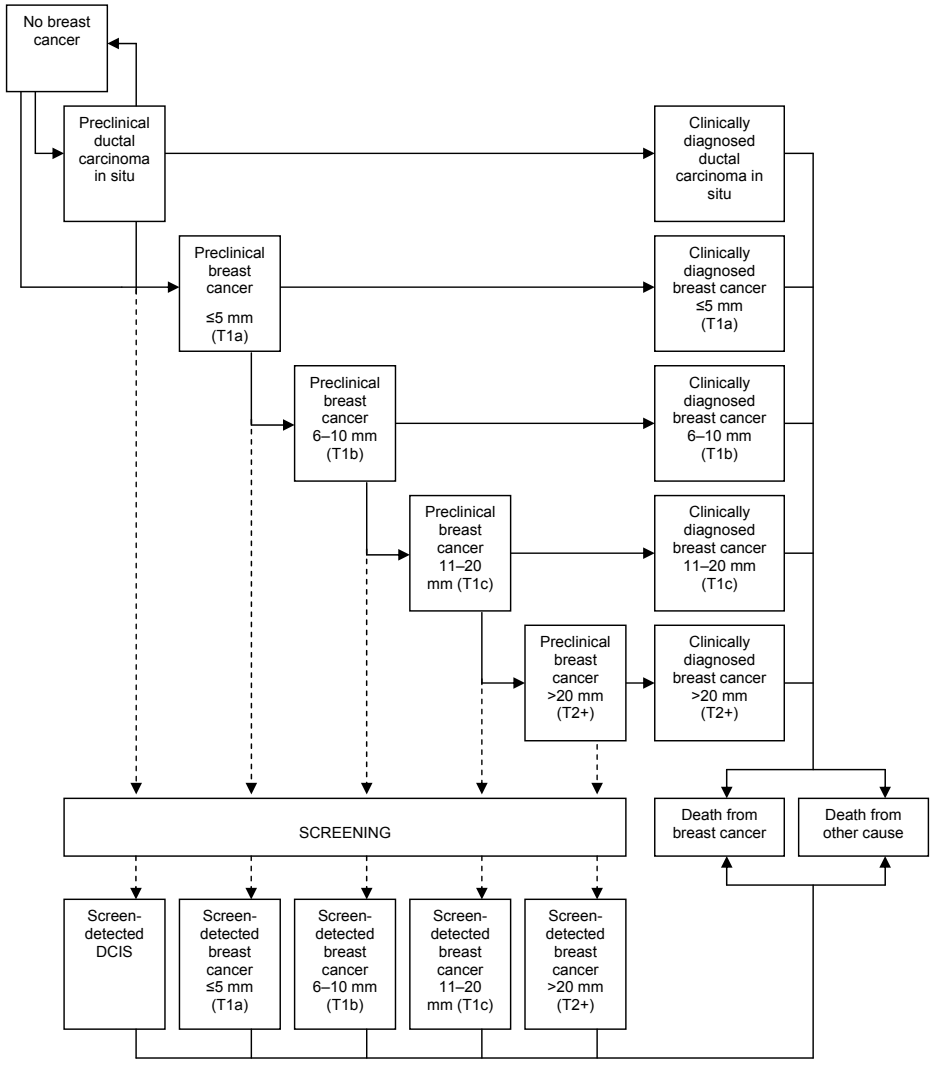


Figure A.1. Transitions in the MISCAN model

preclinical breast cancers are detected, is estimated in a similar manner as the natural history parameters. The improvement in prognosis after screen-detection was based on the Swedish randomized controlled trials.[7, 11–13] The consequences of screening are then calculated by comparing breast cancer incidence and mortality in a population with and without screening. The benefits and harms of different screening policies can be compared by applying them to identical natural histories.

Table A.1 MISCAN parameters and range of parameter estimates within this thesis

Model parameter		Value
Fraction of invasive cancers preceded by preclinical screen-detectable DCIS at age 50 (%)		18%
	Fraction of pre-clinical screen-detectable DCIS that becomes clinically diagnosed	5%
	Fraction of pre-clinical screen-detectable DCIS that progresses to invasive cancer	11%
	Fraction of pre-clinical screen-detectable DCIS that regresses	2%
Duration of pre-clinical breast cancer per stage at age 50 (year)	DCIS	2.1–5.2 years
	T1a	0.1–0.3 year
	T1b	0.4–0.6 year
	T1c	0.8–1.4 years
	T2+	0.8–1.0 year
Test sensitivity of mammography at age 50 (%)	DCIS	48%–80%
	T1a	47%–70%
	T1b	62%–75%
	T1c	80%–90%
	T2+	95%–100%

Although the results described in this thesis (except for Chapter 4) were all calculated with MISCAN, the model parameters may differ somewhat between the chapters, depending on the research question under question. Table A.1 describes the model parameters used in the various chapters of this thesis.

REFERENCES

1. Habbema JD, van Oortmarssen GJ, Lubbe JT, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. Computer methods and programs in biomedicine 1985;20:79–93.
2. van Oortmarssen GJ, Habbema JD, van der Maas PJ, de Koning HJ, Collette HJ, Verbeek AL, Geerts AT, Lubbe KT. A model for breast cancer screening. Cancer 1990;66:1601–12.
3. IKCnet. http://www.ikcnet.nl/. http://www.ikcnet.nl/. April 6, 2010
4. National Evaluation Team for Breast cancer screening (NETB), National evaluation of breast cancer screening in the Netherlands 1990–2007. Twelfth evaluation report. Erasmus MC University Medical Center Rotterdam, Radboud University Nijmegen Medical Centre, 2009.
5. Adami HO, Malker B, Holmberg L, Persson I, Stone B. The relation between survival and age at diagnosis in breast cancer. N Engl J Med 1986;315:559–63.
6. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. Cancer 1989;63:181–7.
7. Tabar L, Vitak B, Chen HH, Duffy SW, Yen MF, Chiang CF, Krusemo UB, Tot T, Smith RA. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. Radiol Clin North Am 2000;38:625–51.
8. Michaelson JS, Silverstein M, Sgroi D, Cheongsiatmoy JA, Taghian A, Powell S, Hughes K, Comegno A, Tanabe KK, Smith B. The effect of tumor size and lymph node status on breast carcinoma lethality. Cancer 2003;98:2133–43.
9. Michaelson JS, Silverstein M, Wyatt J, Weber G, Moore R, Halpern E, Kopans DB, Hughes K. Predicting the survival of patients with breast carcinoma using tumor size. Cancer 2002;95:713–23.
10. Sant M, Allemani C, Capocaccia R, Hakulinen T, Aareleid T, Coebergh JW, Coleman MP, Grosclaude P, Martinez C, Bell J, Youngson J, Berrino F. Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. Int J Cancer 2003;106:416–22.
11. de Koning HJ, Boer R, Warmerdam PG, Beemsterboer PM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. J Natl Cancer Inst 1995;87:1217–23.
12. Nystrom 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;359:909–19.
13. Bjurstam N, Bjorneld L, Warwick J, Sala E, Duffy SW, Nystrom L, Walker N, Cahlin E, Eriksson O, Hafstrom LO, Lingaas H, Mattsson J, *et al.* The Gothenburg Breast Screening Trial. Cancer 2003;97:2387–96.

Dankwoord

DANKWOORD

'Nothing is really work unless you'd rather be doing something else'.
James Barrie

Na ruim 6 jaar, 5 verhuisdozen papierwerk, 54 reviewcommentaren, 2762 koppen dubbele cappuccino, 35 liter bloed, zweet en tranen (yek) en 1 grijze haar kan ik nog altijd zeggen dat werk (meestal) niet als werk voelt. En daar ben ik veel mensen dankbaar voor. Een aantal wil ik hier persoonlijk noemen.

Harry, dank je voor het vertrouwen dat je in me hebt gehad. Je hebt me gestimuleerd om als onderzoeker altijd het beste uit mezelf te halen. En ook al zuchtte ik wel eens na weer een ronde cryptische commentaren, door jouw opmerkingen ligt hier een proefschrift waar ik stiekem best trots op ben.

Gerrit, jij hebt me wegwijs gemaakt in het doen van onderzoek. Met veel enthousiasme legde je me transitiekansen, survivalfuncties, hazardratios and detectiecijfers uit. Heel wat A4-tjes vol formules moeten er de afgelopen jaren door heen gegaan zijn, tijdens jouw geduldige pogingen om me de fijne MISCAN kneepjes bij te brengen. Dankjewel daarvoor!

Eveline, bedankt voor je betrokken begeleiding tijdens de laatste jaren van dit hele gebeuren. Op de jou kenmerkende efficiënte manier van doen stond je altijd voor me klaar, of het nu ging om onderdak toen ik kamerloos werd of het meedenken met artikelen, analyses, inleidingen en discussies. Ik vind het dan ook erg jammer dat je er op 4 april niet bij kunt zijn, maar hé, belangrijker zaken eerst!

Ik wil mijn promotiecommissie, Prof. Hunink, Prof. van Leeuwen en Prof. Rutten, bedanken voor het beoordelen van mijn manuscript. De coauteurs van de verschillende hoofdstukken van dit proefschrift: dank jullie wel voor het samen tot stand brengen van mooie artikelen.

Mijn borstkankerscreencollega's wil ik bedanken voor de fijne samenwerking de afgelopen jaren. Jacques, dank je dat ik altijd bij je terecht kon voor cijfertjes om mijn 'black box' mee te vullen. Je was altijd bereid om uitgebreid voor mijn vragen te gaan zitten, wat dan regelmatig uitmondde in een uur kletsen over reizen, muziek, ambities of wat dan ook. Ik ben trouwens nog steeds heel blij met de CD's!

Nicolien, wat was het fijn om niet meer de enige junior in het team te zijn! Thanks dat ik m'n MISCAN frustraties, succesjes en congresstress bij je kon spuien.

De overige leden van het Landelijk Evaluatie Team Borstkankerscreening: Mireille, Ard, Roland, Martin, Hans en Suzie: dank jullie wel voor de waardevolle suggesties en adviezen op verschillende hoofdstukken van dit proefschrift. Niet de minste in dit rijtje: Arry, dank je dat ik altijd bij je kon binnenstormen: voor de agenda van Harry, adreslijsten, mailtjes, Hora Est formulieren... Vooral toen ik twee banen moest gaan combineren was het heel fijn dat je me met dit soort dingen wilde helpen!

Niets is echt werk wanneer je fijne collega's hebt. Bedankt, MGZ'ers, dat jullie me onder mijn koptelefoon vandaan plukten voor een kop thee, lunchwandeling of babbeltje bij het koffieautomaat. Ik vind het heel bijzonder dat ik verschillende collega's ook buiten de Erasmus muren mocht leren kennen. De hardlooptrainingen in Het Park en door de Koopgoot, de filmfestivalfilms met nachtmerriedetectives, zwevende geesten en Sloveense supersterren, het 'schaatsen' en ijsbreken op Zweedse meren, een picknick in het park, samen het WK voetbal kijken tussen de Oxforder hooligans, fietsen in Zeeland, Maggies housewarmingbrunch, Lifangs Chinese BBQ inclusief Elisabeths taartkunsten, en nog vele andere mooie momenten zullen me altijd bijblijven...

Onder het kopje 'collega's' wil ik in het bijzonder mijn kamergenootjes bedanken. Mateja en Isabelle: jullie introduceerden me – toen ik als groentje op MGZ begon – in de wondere wereld van het promoveren. Jullie relativerende opmerkingen hebben me geholpen om niet al te lang stil te stil te staan bij dingen die soms anders gaan dan je verwacht. Ik wil jullie ook bedanken voor de niet typische Nederlandse 'gezelligheid'! Janneke: soms hoorde je dagenlang niet veel anders van me als 'grom' en 'zucht', als ik weer diep in concentratie over een analyse gebogen zat. Toch hebben we, tijdens al die jaren in kamer AE-137, ook veel lief en leed gedeeld, en gelukkig niet alleen over promotiedingen. Dankjewel voor de fijne tijd! Ik hoop dat we nog eens afspreken als we allebei onze boekjes af hebben.

Ik heb MGZ als een heel prettige en stimulerende werkplek ervaren. Het is een luxe wanneer je statistiek advies, administratieve ondersteuning, wetenschappelijke artikelen en ieder benodigd computerprogramma tot je beschikking hebt staan. Iedereen die hier verantwoordelijk voor was: dankjewel!

Ik wil ook mijn nieuwe collega's van de GG&GD bedanken. Ik ben blij dat ik zo'n warm nest terecht ben gekomen: dat maakte de overgang van Rotterdam naar Utrecht een stuk

makkelijker. Het is fijn dat jullie zo meeleven met deadlines, covers, offertes en de zoektocht naar De Ware Stelling, ook wanneer ik eigenlijk aan URHIS zou moeten werken ☺.

Mijn lieve vrienden: ja hij is echt af! Jullie mogen me weer bellen om af te spreken: ik heb eindelijk weer tijd! Bedankt dat jullie er voor me waren, met een mail of sms'je om te vragen hoe het gaat, met gezellige etentjes, filmavondjes, wandelingen, onbeperkte sushi en weekendjes weg, en vooral met lieve woorden en goede adviezen. Ook al lijkt het soms misschien niet zo, maar jullie steun heeft zeker bijgedragen aan het afronden van dit boekje.

Saskia en Debora: ik ben heel blij dat jullie, als vriendinnen en als paranimfen, naast me willen staan op dit spannende moment! Het is heel fijn dat we samen alle promotietoestand kunnen delen. Nog even, en dan vieren we ook jullie feestjes: zet 'm op!

Christine, wat een geweldige cover heb je ontworpen, en wat ziet de binnenkant er mooi uit! Het is fantastisch dat je dat kan, en dat je me zo wilde helpen met m'n 'levenswerk'.

Na 14 woorden (of beter gezegd: alinea's) van waardering kom ik bij de 'last but not least'. Lieve Ro, Martien, Cas, Chris, Faab, Lies en Dan: bedankt! Dank jullie wel voor jullie altijd oprechte belangstelling, ook al is dat hele promoveren soms maar een vreemde bedoeling. Ik weet dat jullie supertrots op me zijn, en daar ben ik heel blij mee. Pappa en mamma: dank jullie wel dat ik mag weten dat jullie achter me staan, promotie of geen promotie. Jullie hebben altijd in me geloofd. Bedankt daarvoor. Ik hou van jullie!!

About the author

CURRICULUM VITAE

Rianne de Gelder was born on February 29, 1980 in Rotterdam, the Netherlands. After completing her secondary education at the Gereformeerde Scholengemeenschap Randstad in 1998, she studied Biomedical Sciences at Leiden University. During her studies, she did several research internships which resulted in international publications. Rianne obtained her Master's degree in 2005. Between 2006 and 2011, she was appointed as junior researcher at the department of Public Health at the Erasmus MC in Rotterdam. During this period, she studied the benefits and risks of breast cancer screening, the results of which are described in this thesis. Since January 2011, Rianne is employed as health researcher at the Municipal Health Service Utrecht (GG&GD).

Rianne de Gelder werd geboren op 29 februari 1980 te Rotterdam, Nederland. Nadat zij in 1998 haar VWO diploma aan de Gereformeerde Scholengemeenschap Randstad behaalde, studeerde zij Biomedische Wetenschappen aan de Universiteit Leiden. Tijdens haar studie liep zij verschillende onderzoeksstages die resulteerden in internationale publicaties. Rianne behaalde in 2005 haar Master diploma.
Tussen 2006 en 2011 was zij als junior onderzoeker verbonden aan de afdeling Maatschappelijke Gezondheidszorg van het Erasmus MC in Rotterdam. Tijdens deze periode deed zij onderzoek naar de gunstige en ongunstige effecten van borstkankerscreening, waarvan de resultaten in dit proefschrift beschreven worden. Sinds Januari 2011 is Rianne als gezondheidsonderzoeker werkzaam bij de GG&GD Utrecht.

LIST OF PUBLICATIONS

de Gelder R, van As E, Tilanus-Linthorst MM, Bartels CC, Boer R, Draisma G, de Koning HJ. *Breast cancer screening: evidence for false reassurance?* Int J Cancer. 2008 Aug 1;123(3):680-6.

de Gelder R, Bulliard JL, de Wolf C, Fracheboud J, Draisma G, Schopper D, de Koning HJ. *Cost-effectiveness of opportunistic versus organised mammography screening in Switzerland.* Eur J Cancer. 2009 Jan;45(1):127-38. Epub 2008 Nov 27.

de Gelder R, Fracheboud J, Heijsdijk EAM, Broeders M, Verbeek A, den Heeten G, Draisma G, de Koning HJ. *Digital mammography screening: Weighing reduced mortality against increased overdiagnosis.* Prev Med. 2011 Sep 1;53(3):134–40. Epub 2011 Jun 21.

de Gelder R, Heijnsdijk EAM, Fracheboud J, van Ravesteyn N, de Koning HJ. *Interpreting overdiagnosis estimates in population-based breast cancer screening.* Epidemiol Rev. 2011;33(1):111–21. Epub 2011 Jun 27.

de Gelder R, Heijnsdijk EAM, van Ravesteyn N, Fracheboud J, Draisma G, de Koning HJ. *Interpretatie van schattingen van overdiagnose bij borstkankerscreening.* Gamma Professional 2011 Okt;61(4):15-22

de Gelder R, Draisma G, Heijnsdijk EAM, de Koning HJ. *Population-based breast cancer screening for women below age 50: balancing radiation-induced versus prevented breast cancer deaths.* Br J Cancer 2011 Mar 29;104(7):1214-20. Epub 2011 Mar 1.

de Gelder R, Heijnsdijk EAM, Fracheboud J, Draisma G, de Koning HJ. *The effects of population-based mammography screening starting between age 40 and 50 compared to the effects of adjuvant systemic therapy. Submitted.*

PhD PORTFOLIO

Summary of PhD training

Name PhD student:	Rianne de Gelder
Erasmus MC department:	Public Health
PhD period:	2006–2012
Promotor:	prof.dr. H.J. de Koning
Supervision:	dr. G. Draisma dr. E.A.M. Heijnsdijk

PhD training	Year	Workload
General courses		
Department of Public Health, Erasmus MC, Rotterdam:		
- Presentation course	2008	8 hours
- Computer/ statistics courses	2006–2008	8 hours
Specific courses		
Nihes, Erasmus MC, Rotterdam:		
- Planning and Evaluation of Screening	2006	40 hours
Erasmus Winter/ Summer Programme, Rotterdam:		
- Biostatistics for clinicians	2008	40 hours
- Survival analysis for clinicians	2008	40 hours
- Regression analysis	2009	40 hours
Karolinska Institute Stockholm:		
- Essentials of descriptive cancer epidemiology	2009	40 hours
Presentations		
- 'Cost-effectiveness of organized and opportunistic breast cancer screening in Switzerland' *Meeting Swiss Cancer League* *Bern, Switzerland*	2007	40 hours
- 'Evidence for false reassurance in breast cancer screening?' *European Breast Cancer Conference Berlin, Germany*	2008	40 hours

-	'Overdiagnosis in the Dutch breast cancer screening programme' *Research meeting Erasmus MC, Department of Public Health, Rotterdam*	2009	40 hours
-	Overdiagnosis in nation-wide breast cancer screening and the role of detecting DCIS' *Symposium Cancer screening: trials and modeling to guide public health policies* *Erasmus MC, Rotterdam*	2009	20 hours
-	'Estimating overdiagnosis in the Dutch breast cancer screening programme' *Meeting Cancer Intervention and Surveillance Modelling Network (CISNET), Rotterdam*	2009	20 hours
-	'Overdiagnosis in the Dutch Breast Cancer Screening Program' *Meeting International Cancer Screening Network (ICSN), Oxford, UK*	2010	20 hours
-	'Harms and benefits of breast cancer screening' *IX Convegno Osservatorio Nazionale Screening* *Verona, Italy*	2010	40 hours
-	'The Effects of Population-based Mammography Screening Starting Between Age 40 and 50 Compared to the Effects of Adjuvant Systemic Therapy' *European Breast Cancer Conference* *Vienna, Austria*	2012	40 hours

(Inter)national conferences

-	European Breast Cancer Conference *Berlin, Germany*	2008	40 hours
-	Symposium Cancer screening trials and modeling to guide public health policies *Erasmus MC, Rotterdam*	2009	8 hours
-	NIH State-of-the-Science Conference: Diagnosis and Management of Ductal Carcinoma in Situ (DCIS) *Bethesda, USA*	2009	24 hours
-	Meeting International Cancer Screening Network (ICSN) *Oxford, UK*	2010	24 hours
-	IX Convegno Osservatorio Nazionale Screening *Verona, Italy*	2010	16 hours
-	European Breast Cancer Conference *Vienna, Austria*	2012	40 hours

Seminars, meetings and workshops

-	Seminars and research meetings of the Department of Public Health *Erasmus MC, Rotterdam*	2006–2010	140 hours
-	Seminars Erasmus Winter/ Summer Programme *Erasmus MC, Rotterdam*	2006–2007	16 hours