Quantifying and Improving Outcomes of Breast Cancer Screening

Evaluation and long-term model predictions

Valérie Devi Varisha Sankatsing

ISBN: 978-94-6419-095-3

Cover illustration: Jeffrey Pisa
Layout: Rowen Aker, persoonlijkproefschrift.nl
Printing: Gildeprint Enschede, gildeprint.nl

This thesis was financially supported by the Department of Public Health, Erasmus MC, Rotterdam.

Quantifying and Improving Outcomes of Breast Cancer Screening

Evaluation and long-term model predictions

Het kwantificeren en verbeteren van de uitkomsten van borstkanker screening
Evaluatie en lange termijn voorspellingen

Proefschrift

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

Prof.dr. F.A. van der Duijn Schouten

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op

woensdag 3 februari 2021 om 10:30 uur

door

Valérie Devi Varisha Sankatsing

geboren te Nijmegen

Erasmus University Rotterdam

Promotiecommissie

Promotor

Prof.dr. H.J. de Koning

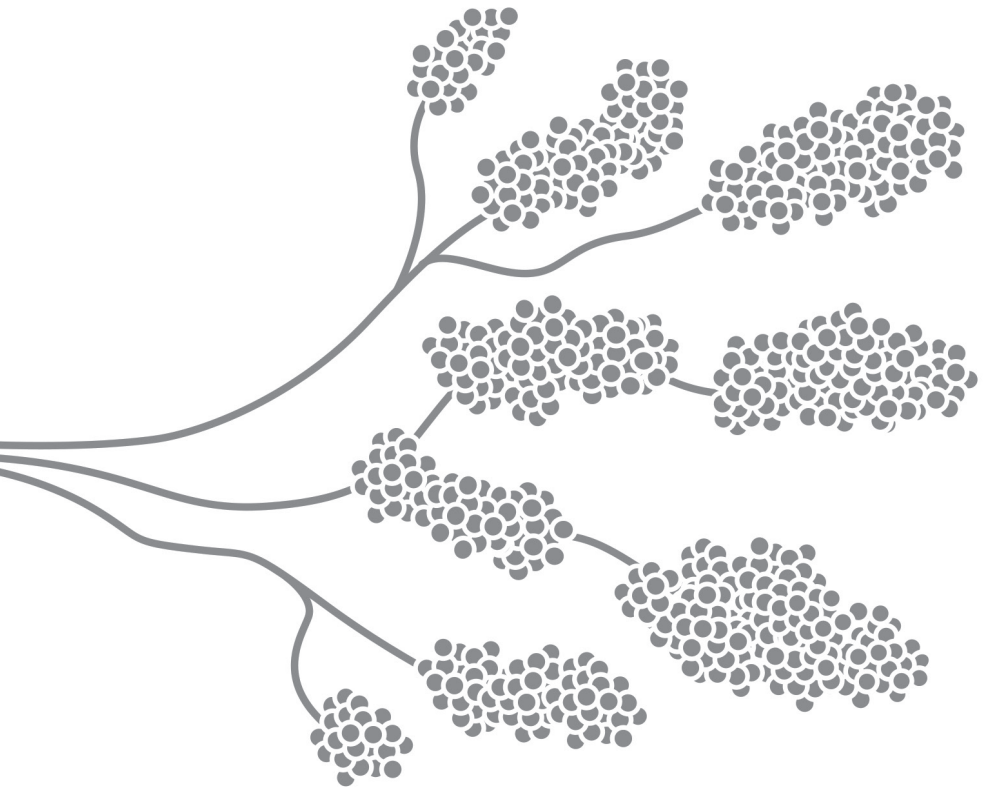
Overige leden

Prof.dr. S. Siesling
Prof.dr. C.H. van Gils
Prof.dr. C. Verhoef

Copromotor

Dr. N.T. van Ravesteyn

CONTENTS



Chapter 1

Introduction

BREAST CANCER

As all cancers, breast cancer starts with uncontrolled cell division as a result of (often multiple) mutations, due to damaged DNA, which are not picked up by the DNA repair system. Cells that originate from uncontrolled cell division can develop into a tumour, which is either benign or malignant. With respect to breast cancer, uncontrolled cell growth starts in the milk ducts, lobules (milk producing glands) or the connective tissue in between the ducts and lobules. If the tumour remains in its original place it is called an in situ carcinoma, which is usually a ductal carcinoma in situ (DCIS) in case of breast tumours. Lobular carcinoma in situ (LCIS) is uncommon, without symptoms and generally not visible on a mammogram.[1]

A tumour becomes invasive when it breaks through the basement membrane, after which it can invade nearby tissues and enter the bloodstream.[1] Once in the blood vessels, malignant cells can spread (metastasize) to various sites. This is only possible when certain criteria are met, e.g. the growth of new blood vessels from the pre-existing vasculature (angiogenesis). There is no consensus on whether or not an invasive breast cancer is always preceded by an in situ carcinoma.

The stage of breast cancer is often determined using the TNM-classification system, which defines the size of the tumour (T), possible regional lymph node involvement (N) and possible distant metastases (M). Invasive disease is categorized in stages T1 – T4, which represent different tumour sizes and possible tumour extension to chest wall and/or skin (T4).[2] Stage T1 is subdivided in T1a, T1b and T1c, which differ in size (Table 1). In case of regional lymph node involvement, i.e. regional lymph node metastases, the disease is classified as 'node positive'. Other, non-regional, lymph node metastases and metastases which spread through the bloodstream are coded as 'distant metastases' (M1). Distant metastases often occur in the lungs, pleura, bones, liver, brain, lymph nodes, peritoneum, skin or adrenals.[3]

Table 1. TNM classification for breast cancer

Stage		
Tis (DCIS)		Ductal carcinoma in situ
T1		Invasive tumor; greatest dimension ≤ 2 cm
	T1a	Invasive tumor; greatest dimension ≤ 0.5 cm
	T1b	Invasive tumor; greatest dimension > 0.5 cm and ≤ 1 cm
	T1c	Invasive tumor; greatest dimension > 1 cm and ≤ 2 cm
T2		Invasive tumor; greatest dimension > 2 cm and ≤ 5 cm
T3		Invasive tumor; greatest dimension > 5 cm
T4		Invasive tumor; any size with direct extension to chest wall and/or skin
N0		No regional lymph node metastases
N1		Movable ipsilateral level I, II axillary lymph node(s) metastases
N2		Metastases ipsilateral level I, II axillary lymph node(s) that are clinically fixed or in clinically detected ipsilateral internal mammary nodes in the absence of axillary lymph node metastases
N3		Metastases ipsilateral infraclavicular (level III axillary) lymph node(s) with or without level I, II axillary lymph node or internal mammary lymph node involvement
M0		No distant metastases
M1		Distant metastases

Adapted from: Edge S BD, Comptom CC,Frits AG, Greene FL, Trotti A. AJCC Cancer Staging Manual. 7 ed: Springer-Verlag New York; 2010. XV, 648 p.

Risk factors

Although many breast cancers are caused by random errors in DNA replication, there are known factors that could increase breast cancer risk. Besides gender and age, risk factors that are associated with a high relative risk are dense breasts, a previous biopsy and a family history of breast cancer. The latter could indicate inherited cases of breast cancer, often associated with mutations in BRCA1 and BRCA2 genes, which are known to increase the risk of breast cancer substantially.

Other known mutations, which moderately elevate breast cancer risk, include CHECK2, ATM and NF1.[4]

Several reproductive factors are also associated with elevated risk of breast cancer, for example advanced age at first birth, nulliparity, and low age at menarche increase breast cancer risk.[5, 6] Lifestyle factors that affect the risk of breast cancer include a lack of physical activity, smoking, being overweight or excessive alcohol consumption.[7] Reproductive factors can also lead to a lower than average risk of breast cancer. Parity, ever breastfeeding and young age at menopause have been shown to be associated with a moderately decreased breast cancer risk.[5, 6]

BREAST CANCER TREATMENT

Almost all women with breast cancer are treated. Most tumours, DCIS and invasive cancers, are primarily treated through breast-conserving surgery (lumpectomy) or mastectomy, with or without radiation. The majority of women with (invasive) breast cancer is additionally treated with adjuvant treatment, including hormone therapy, chemotherapy, radiotherapy and targeted therapy (or combinations of these therapies). Adjuvant treatment considerably improved over the last decades. Adjuvant systemic therapy, including chemotherapy, hormonal therapy and immunotherapy, can be administered to minimize the recurrence risk and the risk of metastasizing or to control metastatic breast cancer. A patient can also be treated with neo-adjuvant treatment to decrease the tumour size before surgery.[8]

BREAST CANCER SCREENING

Population-based cancer screening is performed in an asymptomatic, healthy population. Screening enhances diagnosis of breast cancer in an early (localized) stage, improving tumour stage distribution compared to clinically diagnosed cancers. Breast cancers diagnosed at an early stage have a higher chance of being treated successfully, which may lead to prolonged survival. The stage shift caused by screening results in less advanced treatment among women with screen-

detected cancers compared to non-screened women. Although screening can result in improved survival and may avert breast cancer deaths, it can also have adverse effects and cause harm. As the population invited to screening is considered to be healthy and screening is thus only beneficial for a rather small number of women, disadvantages associated with screening should be limited. The balance between benefits and harms of screening has often been a topic of debate.[9-11]

Because screening brings the date of diagnosis forward, the period between diagnosis and death (lead time) is generally longer with screening, even if breast cancer death is not postponed or prevented (Figure 1). Prolonged breast cancer survival due to screening may thus be misleading and breast cancer mortality (reduction) is therefore a better measure to assess the effect of screening.

Population based breast cancer screening programmes make use of mammography screening. A mammogram is a x-ray image of the breast with moderately high test sensitivity and high specificity.[12]

Benefits of mammography screening

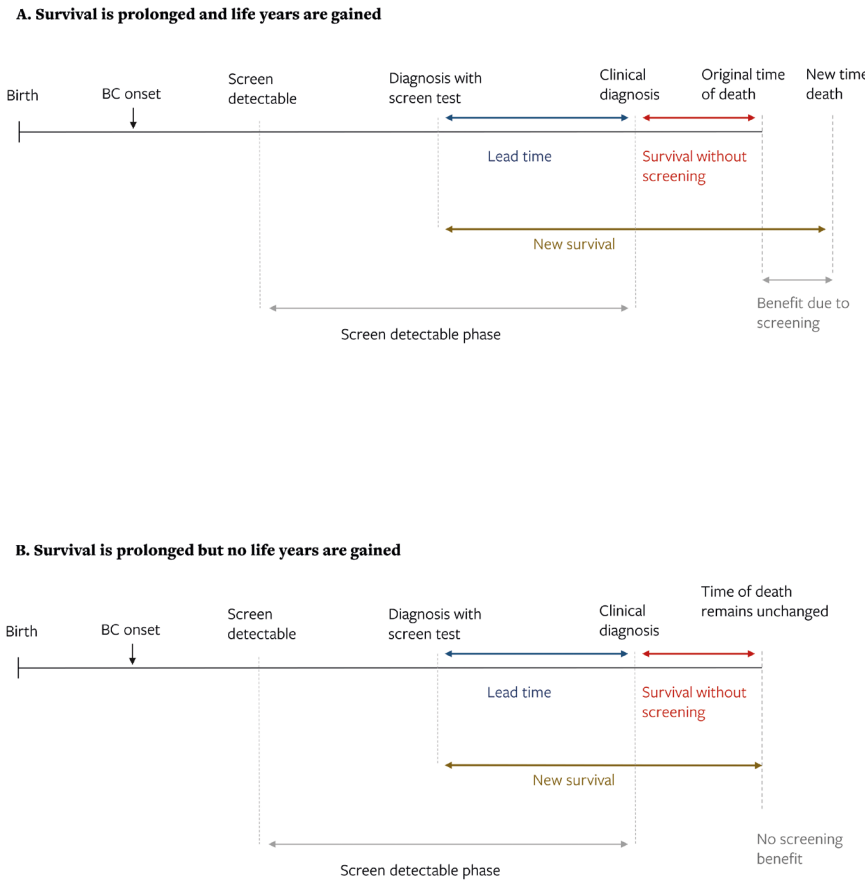
Ten randomised trials of mammography screening were conducted in the 1970s and 1980s.[13-17] Meta-analyses of these trials, showed a reduction in breast cancer mortality of around 20%.[18, 19]

Because of the extensive debate about the balance between the benefit and harms of mammography screening, an independent panel of experts on mammography screening was formed in the UK.[11] After conducting a large review including a meta-analysis, with 13 years of follow-up, the panel estimated that the breast cancer mortality reduction due to mammography screening was 20% for women invited to screening.

In addition to the trials, numerous observational studies have been conducted to estimate the effect of screening on breast cancer mortality. Using the evidence from observational studies, the International Agency for Research on Cancer recently estimated the reduction in breast cancer mortality as a result of mammography

screening to be 40% in women aged 50 to 69 years who attended screening.[20] The reduction in breast cancer mortality was 23% for women in the same age range who were invited to screening. The observational evidence for a reduction in breast

Figure 1. In a situation with screening, breast cancers can be detected earlier during the screen detectable phase. The time that the diagnosis is brought forward is referred to as the 'lead time'. Screening is beneficial if life years are gained and the time of death is thus postponed (A). However, it is important to note that survival is always prolonged with screening, as the diagnosis is brought forward in time, even if no life years are gained and the time of death is not postponed (B). BC: breast cancer



cancer mortality from mammography screening for women aged 70 to 74 years was also considered to be sufficient.[20]

Next to breast cancer mortality reduction in a screened versus non-screened population, beneficial outcome measures of mammography screening are the number of life years gained, less advanced breast cancer stage and less advanced treatment.

Harms of mammography screening

One of the most important potential adverse outcomes of screening is overdiagnosis of breast cancer. Overdiagnosis is defined as screen-detection of tumours that would never have presented clinically during an individual's lifetime in the absence of screening. Overdiagnosis can occur because some cases of screen-detected DCIS or indolent invasive breast cancer may never present clinically during a woman's lifetime, due to slow growth, a complete lack of growth or regression of the lesion[21, 22]. In this case, an individual will die of another cause than breast cancer. However, overdiagnosis is also possible with respect to lesions with average or high growth rates if women die of other causes. Overdiagnosis results in more individuals being diagnosed in the presence of screening and may lead to overtreatment in the screening setting. Complications or side effects as a consequence of overtreatment are undesired, since treatment of overdiagnosed cancers will not improve survival.

Other harms associated with screening are false-positive findings and false reassurance. It has been suggested that false-positive mammograms increase short-term, but not long-term anxiety.[23] False reassurance might occur when, in case of a false-negative screen result, an individual is less perceptive to symptoms of breast cancer because of the reassuring negative screen.

Evaluation of a screening programme

The performance of a breast cancer screening programme is measured using certain performance indicators, including the breast cancer detection rate, the

interval cancer rate, the referral rate and the false-positive rate. In addition, the stage distributions of screen- and clinically detected cancers are also monitored.

For screening to be effective, the interval cancer rate should be rather low and the detection rate relatively high. Interval cancers are usually defined as breast cancers diagnosed after a negative screening examination (i.e. not resulting in a referral), within the first two years after screening.

Performance indicators are used to calculate programme sensitivity and specificity. Programme sensitivity is defined as the proportion of true-positives among all breast cancers (true-positives and false-negatives), diagnosed within two years after the screening examination.[24] To calculate the programme sensitivity, interval cancers are often used to approximate the number of false-negative findings. However, the number of interval cancers includes not only cancers missed at screening, but also fast growing cancers that were not detectable during the screening examination. The programme sensitivity is therefore lower than the test sensitivity of mammography. The programme specificity is calculated as the proportion of true-negative findings among all negative screening examinations (true negative and false-positive findings), within the first two years after screening. Another measure is the positive predictive value: the chance of having breast cancer after a positive screening examination. Also important for the effectiveness of a screening programme are a high attendance rate and, as mentioned earlier, a more favourable stage distribution for screen-detected cancers.

Breast cancer screening in the Netherlands

Biennial breast cancer screening was gradually implemented in the Netherlands between 1989 and 1997. Initially only women aged 50-69 years were invited, until 1998, when the upper age limit of screening was extended to 74 years between 1998 and 2001. In the period 2004-2010, screen-film mammography was replaced by digital mammography, reaching full transition in June 2010. Before 2004, 2-view mammography (cranial-caudal and mediolateral-oblique) was performed at first screens whereas at subsequent screens only 1-view examinations (mediolateral-oblique) were performed. The number of 2-view mammograms at subsequent

screens increased steadily and after 2010, 2-view mammography was performed at all subsequent screening rounds. The reading policy in the Netherlands is double reading with, in case of disagreement, consensus or arbitration.[25]

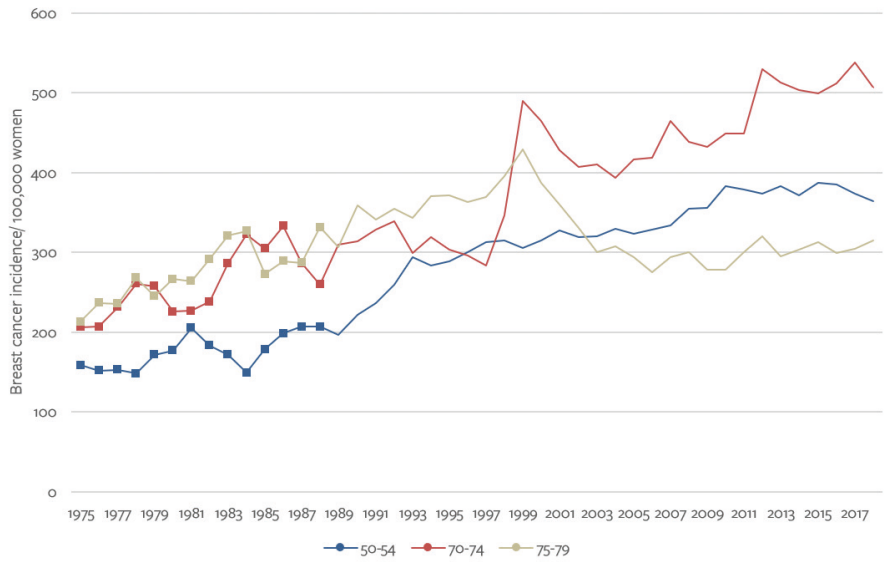
The attendance rate over the last years has been around 80% in the Netherlands.[24] Analyses of Dutch screening data have shown that the percentages of advanced stage breast cancer (stages III and IV) among women who were screened and women who were not screened or irregularly screened were 10% and 23% respectively.[26]

In the Netherlands, breast-conserving therapy is more common among screen-detected cancers than among cancers detected outside of screening, namely 71% versus 38%.[24] In addition, the percentage of adjuvant therapy after surgery is significantly lower for screen-detected cancers than for breast cancers in women who were not screened (50% versus 68%).[24]

BREAST CANCER INCIDENCE AND MORTALITY IN THE NETHERLANDS

The breast cancer incidence in the Netherlands is one of the highest in Europe.[27] Breast cancer incidence in women aged 50-54 years increased substantially around the implementation of mammography screening in the Netherlands (1989-1994), compared to the years before implementation (Figure 2). This steep rise in incidence, caused by the detection of prevalent cases of breast cancer during the first screening rounds, attenuated around 1994. There is a peak in the incidence trend of women aged 70-74 in 1999, during the extension of the screening programme to age 70-74 years. Around the same time, the incidence in older women, aged 75-79, decreases significantly because part of the breast cancers in this age group are detected earlier due to screening in the age group 70-74 years (Figure 2).

Figure 2. Breast cancer incidence in the Netherlands for different age groups between 1975 and 2018.



The squares represent observations based on regional data as national data was lacking during this period.

Breast cancer mortality has decreased significantly over the last 25 years in the Netherlands. This decline is present in all five-year age groups between 40 and 79 years (Figure 3).[24, 28] The reduction in breast cancer mortality over the years was probably caused by the implementation of the breast cancer screening programme around 1990 and the introduction of adjuvant chemo- and hormonal therapy.[29, 30] Adjuvant treatment improved substantially over the last 25 years, which additionally contributed to breast cancer mortality reduction.[31] In the highest age group, 75-79 years, there is a steep decline in breast cancer mortality shortly after the extension of the screening programme to 70-74 years between 1998 and 2001 (Figure 3).[24]

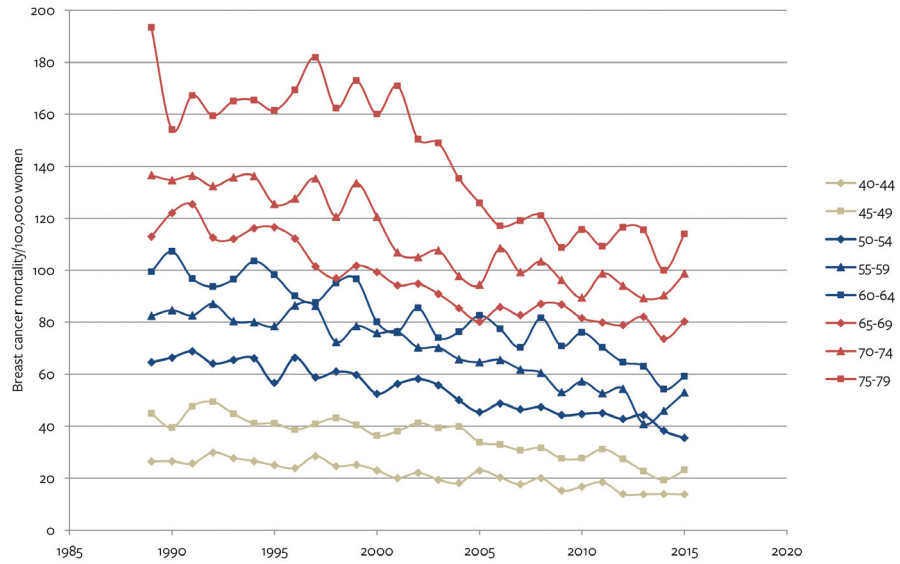
Survival

As concluded above, breast cancer mortality decreased over the last decades, despite the rise in breast cancer incidence. Survival after diagnosis thus increased

over this period. In the Netherlands, 5-year survival after diagnosis of invasive breast cancer changed from 78% with a diagnosis between 1991-1995 to 88% with a diagnosis between 2011-2015.[32]

Improved survival is, as well as breast cancer mortality reduction, probably the result of both early detection due to screening and improvements in (adjuvant) therapy. As discussed earlier, prolonged survival does not always result in delayed or averted breast cancer death.

Figure 3. Breast cancer mortality in the Netherlands for different age groups between 1985 and 2015.



POSSIBLE NEW DEVELOPMENTS IN BREAST CANCER SCREENING

Screening before age 50 years

Breast cancer incidence among women aged 40-49 years has been increasing over the last decades.[33, 34] The impact of screening on breast cancer mortality may be different for women aged 40-49 because of several factors associated with younger

age, including a lower breast cancer incidence, lower sensitivity of mammography due to greater breast density and, possibly, more aggressive tumour growth. Meta-analyses of randomised controlled trials in which women aged 40 to 49 at entry were included, report a breast cancer mortality reduction of 8-15%.[19, 35] However, methodologically these analyses may be flawed as the benefit of screening observed in women aged 40 to 49 at randomization, may be partially attributable to early detection of breast cancer by screening these women at age 50 and older[36]. Findings of the United Kingdom (UK) age trial suggested a breast cancer mortality reduction of 12% from annual mammography starting at 39-41 years, with 17 years of follow-up.[37] The reduction was however only statistically significant after 10 years of follow-up (25% reduction) and not at 17 years of follow-up.

The largest cohort study that compared breast cancer mortality in Swedish women aged 40 to 49 years, between women invited and not invited to screening, showed a breast cancer mortality reduction of 26% for women invited to screening and 29% for women attending screening.[38] The effect of lowering the starting age of breast cancer screening is therefore interesting to explore.

Risk-based screening

In most countries, organized screening programmes invite all women in a specific age group (often 50-69 or 50-74 years), regardless of their risk of breast cancer. However, breast cancer risk varies for different ages as a result of changes in risk factors with increasing age. A woman's individual risk of breast cancer, dependent on her age and the presence or absence of risk factors (e.g. breast density, family history of breast cancer), may affect the balance between benefits and harms associated with screening. Better individual harm-benefit ratios may also lead to an improved balance on the population level. An alternative to offering a uniform screening strategy is a risk-based approach, in which the target age range, the screening interval and possibly the screening method can be adjusted to different risk groups. There are several European studies in which risk factors for breast cancers are being identified.[39-42]

New screening technnologies

There are multiple studies on alternatives for mammography, for the whole screening population or a specific – often high-risk - subgroup. Frequently studied technologies are magnetic resonance imaging (MRI), digital breast tomosynthesis and ultrasound.

Because MRI has a relatively high sensitivity compared to other imaging modalities for breast cancer detection, it is used to screen women with a high risk of breast cancer, due to familial or genetic predispositions.[43-45] The use of MRI to screen women with dense breasts is currently investigated.[46] Compared to mammography, MRI has prolonged acquisition time and higher false-positive rates[47] and costs.

Digital breast tomosynthesis is of particular interest as it has been suggested as a replacement for digital mammography, for the whole eligible screening population, in the long term. By generating a 3D-like image of the breast, tomosynthesis has the potential to overcome the issue of overlapping breast tissue on a 2D mammogram, which may result in improved diagnostic accuracy compared to digital mammography.[48, 49] Tomosynthesis screening leads however to an increase in reading time, compared to digital mammography and higher costs.[50, 51] Estimates for the false-positive rate with tomosynthesis vary considerably between studies from substantial decreases to significant increases compared to digital mammography.[52-55]

MICROSIMULATION MODELLING OF BREAST CANCER SCREENING

Microsimulation models are used in an increasing number of studies to evaluate the effect of disease interventions, including cancer screening. By extrapolating the results of randomised controlled trials, models are able to estimate the impact of screening under many different circumstances, including those that are not feasible to test in trials due to ethical, time- or cost-related issues.[56, 57] Conditions that can easily be tested or varied using modelling are lifelong follow-up and the effect of screening for different sub-groups of the population. In addition, models

can simulate multiple screening scenarios, whereas in trials usually only one or a few strategies are tested. Optimum screening policies are therefore often determined using models.[58] Furthermore, models have the ability to adjust the estimated effect of screening to changes in screening policy (e.g. transition to digital mammography), while maintaining other conditions.[59] Also, models can make predictions for the future and are able to quantify unobservable factors associated with screening including overdiagnosis[59] and the relative contribution of screening and treatment to cancer mortality reduction.[29, 30]

MISCAN

The MISCAN (MIcrosimulation SCreening ANalysis) model for breast cancer screening was developed in the 1980's and has frequently been used for the (economic) evaluation of mammography screening and for recommendations for screening policy.[59-61] MISCAN has been well reported and validated in the past and has been frequently recalibrated and updated.[61] Important components of the model are: the population demographics, the natural history of breast cancer, the screening component and the treatment component. The model simulates a population consisting of individual life histories, based on life-tables of Dutch women. Subsequently, the natural history of breast cancer (without screening) is simulated resulting in the onset of breast cancer in a subset of women in the population at a certain point in time, which may eventually lead to breast cancer death. Model outcomes are then estimated for a situation without screening. Hereafter, mammography screening and improvements in prognosis of survival after screen-detection are modelled. In the presence of screening, tumours can become screen-detected during the preclinical detectable phase, before clinical symptoms are present. Screening can therefore lead to earlier detection and treatment of breast cancer and may, thereby, improve survival and may prevent breast cancer death. In the model, screen-detected and clinically detected tumours are primarily treated with surgery and may be treated with adjuvant therapy, based on Dutch treatment probabilities.

In MISCAN, breast cancer is modelled through tumour progression and starts with the development of preclinical ductal carcinoma in situ (DCIS), which may

progress through the invasive successive stages T1a, T1b, T1c and T2+ (modelled as a semi-Markov process). DCIS progression into T1a varies from immediate transition to slow progression. A small fraction of DCIS is assumed to regress. At each stage, a tumour can become screen-detected in the presence of screening, clinically diagnosed if symptoms are present or progress to the next stage (Figure 4).

The monitored performance indicators from the evaluation of the national screening programme are used as input for MISCAN, either as direct input or as data used for model calibration.

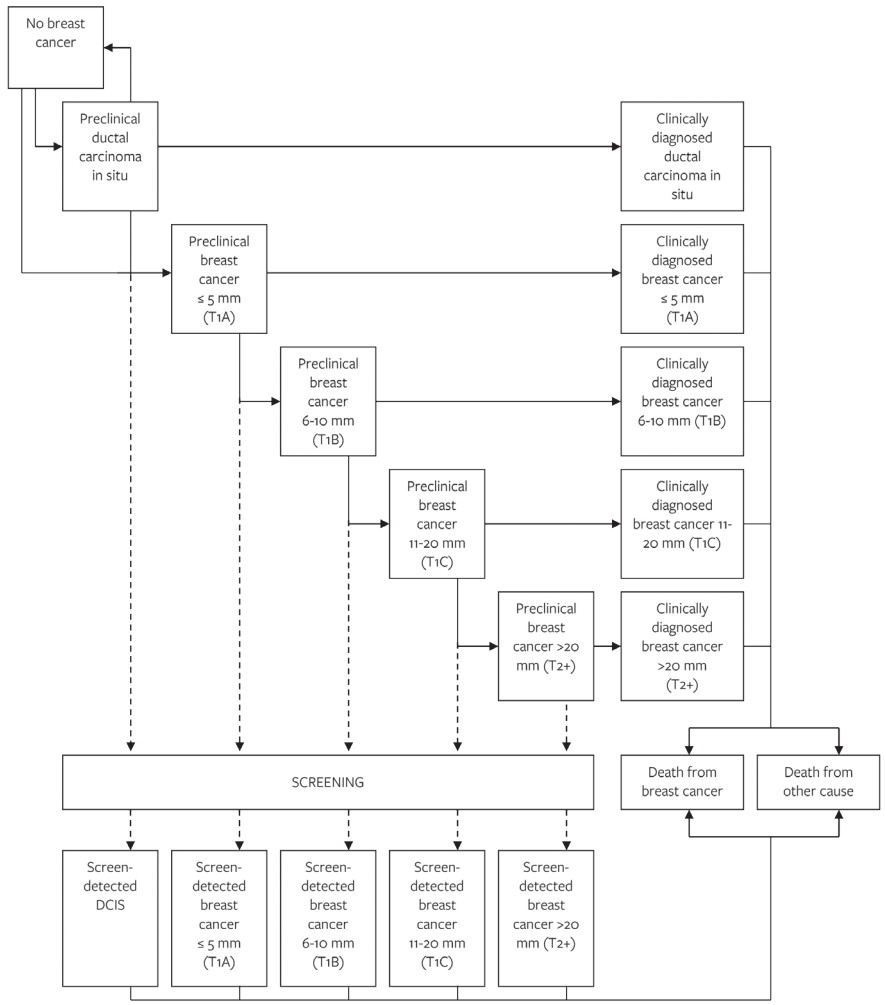
Cost-effectiveness analysis

In a cost-effectiveness analysis the benefits of screening are compared to the costs of screening. Benefits of breast cancer screening are often defined as life years gained, which result from averted breast cancer deaths. A cost-effectiveness threshold, often referred to as a willingness-to-pay-threshold, is used to determine whether or not an intervention is cost-effective.

In European countries, organized breast cancer screening has been demonstrated to be cost-effective.[25, 62] Most European breast cancer screening programmes are targeted to women aged 50-69 years, with a screening interval of 2 years.[63] This age range has been extended to 40 years, 74 years, or both in some European countries. However, even if there is general consistency among European countries with respect to their screening policies, the cost-effectiveness of screening may differ between countries because it depends on country-specific characteristics such as the breast cancer incidence; tumour stage distribution and breast cancer mortality before the start of screening; the target age range and screening interval; the structure and organization of the health care system; the coverage of the population by invitation[64]; the costs of screening and the costs of diagnostics and treatment. Another important factor that affects the cost-effectiveness is the attendance rate of the programme. Attendance rates differ substantially between European countries, ranging from 19% to 89%.[63]

It is also important to assess the cost-effectiveness of on-going screening programmes, as the ratio of effects and costs may change over time. Assessing the cost-effectiveness of current screening programmes is particularly relevant when changes in screening policies are considered - for example extension of the age range - or when a new screening technology is available.

Figure 4. Possible transitions in the MISCAN model.



RESEARCH QUESTIONS AND OUTLINE OF THESIS

The aim of this thesis is to quantify the effects of breast cancer screening. This thesis consists of two parts. The first part evaluates the performance indicators over the period with digital mammography screening in the Netherlands and compares these to the performance indicators of former screen-film mammography. The effect of both screen-film and digital mammography screening on breast cancer mortality rates in the Netherlands is also assessed in this part. In the second part of this thesis, the benefits, harms and cost-effectiveness of different alterations to the current screening strategy are described.

Part One: Evaluation of current breast cancer screening in the Netherlands

Research question 1: How do the performance indicators of current breast cancer screening in the Netherlands change over time and what is the effect of screening on breast cancer mortality rates?

In chapter 2, the detection rates, interval cancer rates and other important performance indicators after the transition to digital mammography are evaluated and compared to indicators with screen-film mammography. In chapter 3, the reduction in breast cancer mortality trends in different age groups over the last 20 years, after the introduction of the Dutch screening programme, is quantified.

Part Two: Quantifying the cost-effectiveness, harms and benefits of different screening strategies and screening modalities in the Netherlands using microsimulation modelling

Research question 2: To what extent do alterations to current screening change the harm-benefit ratios and cost-effectiveness estimates of current screening?

Chapter 4 presents the cost-effectiveness of different screening strategies starting before age 50 years in the Netherlands and discusses the corresponding harms.

In Chapter 5, we identify optimal screening strategies for women with a low and high relative risk of breast cancer, under the condition that the cost-effectiveness of screening in the Netherlands is not negatively affected. In Chapter 6, the incremental cost-effectiveness of screening with digital breast tomosynthesis in the Netherlands, compared to digital mammography screening, is assessed by conducting a probabilistic sensitivity analysis.

REFERENCES

1. American Cancer Society. Available from: https://www.cancer.org/cancer/breast-cancer/understanding-a-breast-cancer-diagnosis.html
2. Edge S BD CC, Frits AG, Greene FL, Trotti A. AJCC Cancer Staging Manual. 7 ed: Springer-Verlag New York; 2010. XV, 648 p.
3. Sobin LH GM, Wittekind C. TNM Classification of Malignant Tumours, Seventh Edition. New York: Wiley-Blackwell 2009:310-p.
4. Easton DF, Pharoah PD, Antoniou AC, Tischkowitz M, Tavtigian SV, Nathanson KL, Devilee P, Meindl A, Couch FJ, Southey M, Goldgar DE, Evans DG, et al. Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med* 2015;372: 2243-57.
5. Lambertini M, Santoro L, Del Mastro L, Nguyen B, Livraghi L, Ugolini D, Peccatori FA, Azim HA, Jr. Reproductive behaviors and risk of developing breast cancer according to tumor subtype: A systematic review and meta-analysis of epidemiological studies. *Cancer Treat Rev* 2016;49: 65-76.
6. Collaborative Group on Hormonal Factors in Breast C. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol* 2012;13: 1141-51.
7. Rookus MA MT, van Leeuwen FE. Risicofactoren van mammacarcinoom. In: Wobbes T, Nortier JWR, Koning CCE (eds.). Handboek Mammacarcinoom. Utrecht: De Tijdstroom, 2007.
8. Steenbruggen TG, van Ramshorst MS, Kok M, Linn SC, Smorenburg CH, Sonke GS. Neoadjuvant Therapy for Breast Cancer: Established Concepts and Emerging Strategies. *Drugs* 2017;77: 1313-36.
9. Autier P, Boniol M. Mammography screening: A major issue in medicine. *Eur J Cancer* 2018;90: 34-62.
10. Gotzsche PC. Mammography screening is harmful and should be abandoned. *J R Soc Med* 2015;108: 341-5.
11. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. *Br J Cancer* 2013;108: 2205-40.
12. 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.
13. Shapiro S. Periodic screening for breast cancer: the HIP Randomized Controlled Trial. Health Insurance Plan. *J Natl Cancer Inst Monogr* 1997: 27-30.

14. Alexander FE, Anderson TJ, Brown HK, Forrest AP, Hepburn W, Kirkpatrick AE, Muir BB, Prescott RJ, Smith A. 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. *Lancet* 1999;353: 1903-8.

15. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. *Ann Intern Med* 2002;137: 305-12.

16. Miller AB, To T, Baines CJ, Wall C. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50-59 years. *J Natl Cancer Inst* 2000;92: 1490-9.

17. 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.

18. Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2013: CD001877.

19. Nelson HD, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L. Effectiveness of Breast Cancer Screening: Systematic Review and Meta-analysis to Update the 2009 U.S. Preventive Services Task Force Recommendation. *Ann Intern Med* 2016;164: 244-55.

20. Lauby-Secretan B, Scoccianti C, Loomis D, Benbrahim-Tallaa L, Bouvard V, Bianchini F, Straif K, International Agency for Research on Cancer Handbook Working G. Breast-cancer screening--viewpoint of the IARC Working Group. *N Engl J Med* 2015;372: 2353-8.

21. Yen MF, Tabar L, Vitak B, Smith RA, Chen HH, Duffy SW. Quantifying the potential problem of overdiagnosis of ductal carcinoma in situ in breast cancer screening. *Eur J Cancer* 2003;39: 1746-54.

22. Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. *Lancet Oncol* 2007;8: 1129-38.

23. Tosteson AN, Fryback DG, Hammond CS, Hanna LG, Grove MR, Brown M, Wang Q, Lindfors K, Pisano ED. Consequences of false-positive screening mammograms. *JAMA Intern Med* 2014;174: 954-61.

24. National Evaluation Team for Breast cancer screening (NETB). National evaluation of breast cancer screening in the Netherlands 2004-2014. Fourteenth evaluation report. Erasmus MC University Medical Center Rotterdam, Radboud University Nijmegen Medical Centre, 2019.

25. Groenewoud JH, Otten JD, Fracheboud J, Draisma G, van Ineveld BM, Holland R, Verbeek AL, de Koning HJ, Netb. Cost-effectiveness of different reading and referral strategies in mammography screening in the Netherlands. *Breast Cancer Res Treat* 2007;102: 211-8.

26. de Munck L, Fracheboud J, de Bock GH, den Heeten GJ, Siesling S, Broeders MJM. Is the incidence of advanced-stage breast cancer affected by whether women attend a steady-state screening program? *Int J Cancer* 2018;143: 842-50.

27. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, Forman D, Bray F. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer* 2013;49: 1374-403.

28. 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.

29. de Gelder R, Heijnsdijk EA, Fracheboud J, Draisma G, de Koning HJ. The effects of population-based mammography screening starting between age 40 and 50 in the presence of adjuvant systemic therapy. *Int J Cancer* 2015;137: 165-72.

30. Plevritis SK, Munoz D, Kurian AW, Stout NK, Alagoz O, Near AM, Lee SJ, van den Broek JJ, Huang X, Schechter CB, Sprague BL, Song J, et al. Association of Screening and Treatment With Breast Cancer Mortality by Molecular Subtype in US Women, 2000-2012. *JAMA* 2018;319: 154-64.

31. Early Breast Cancer Trialists' Collaborative G. 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.

32. www.iknl.nl.

33. 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.

34. Vondeling GT, Menezes GL, Dvortsin EP, Jansman FGA, Konings IR, Postma MJ, Rozenbaum MH. Burden of early, advanced and metastatic breast cancer in The Netherlands. *BMC Cancer* 2018;18: 262.

35. Magnus MC, Ping M, Shen MM, Bourgeois J, Magnus JH. Effectiveness of mammography screening in reducing breast cancer mortality in women aged 39-49 years: a meta-analysis. *J Womens Health (Larchmt)* 2011;20: 845-52.

36. 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.

37. Moss SM, Wale C, Smith R, Evans A, Cuckle H, Duffy SW. Effect of mammographic screening from age 40 years on breast cancer mortality in the UK Age trial at 17 years' follow-up: a randomised controlled trial. *Lancet Oncol* 2015;16: 1123-32.

38. 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.

39. Gabrielson M, Eriksson M, Hammarstrom M, Borgquist S, Leifland K, Czene K, Hall P. Cohort Profile: The Karolinska Mammography Project for Risk Prediction of Breast Cancer (KARMA). *Int J Epidemiol* 2017;46: 1740-1g.

40. Giordano L, Gallo F, Petracci E, Chiorino G, Segnan N, Andromeda working g. The ANDROMEDA prospective cohort study: predictive value of combined criteria to tailor breast cancer screening and new opportunities from circulating markers: study protocol. *BMC Cancer* 2017;17: 785.

41. https://www.prisma-studie.nl/.

42. https://mypebs.eu/the-project/.

43. Sardanelli F, Podo F, D'Agnolo G, Verdecchia A, Santaquilani M, Musumeci R, Trecate G, Manoukian S, Morassut S, de Giacomi C, Federico M, Cortesi L, et al. Multicenter comparative multimodality surveillance of women at genetic-familial high risk for breast cancer (HIBCRIT study): interim results. *Radiology* 2007;242: 698-715.

44. Kuhl CK, Schrading S, Leutner CC, Morakkabati-Spitz N, Wardelmann E, Fimmers R, Kuhn W, Schild HH. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. *J Clin Oncol* 2005;23: 8469-76.

45. Kriege M, Brekelmans CT, Boetes C, Besnard PE, Zonderland HM, Obdeijn IM, Manoliu RA, Kok T, Peterse H, Tilanus-Linthorst MM, Muller SH, Meijer S, et al. Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition. *N Engl J Med* 2004;351: 427-37.

46. Emaus MJ, Bakker MF, Peeters PH, Loo CE, Mann RM, de Jong MD, Bisschops RH, Veltman J, Duvivier KM, Lobbes MB, Pijnappel RM, Karssemeijer N, et al. MR Imaging as an Additional Screening Modality for the Detection of Breast Cancer in Women Aged 50-75 Years with Extremely Dense Breasts: The DENSE Trial Study Design. *Radiology* 2015;277: 527-37.

47. Buist DSM, Abraham L, Lee CI, Lee JM, Lehman C, O'Meara ES, Stout NK, Henderson LM, Hill D, Wernli KJ, Haas JS, Tosteson ANA, et al. Breast Biopsy Intensity and Findings Following Breast Cancer Screening in Women With and Without a Personal History of Breast Cancer. *JAMA Intern Med* 2018;178: 458-68.

48. Baker JA, Lo JY. Breast tomosynthesis: state-of-the-art and review of the literature. *Acad Radiol* 2011;18: 1298-310.

49. Rafferty EA, Park JM, Philpotts LE, Poplack SP, Sumkin JH, Halpern EF, Niklason LT. Assessing radiologist performance using combined digital mammography and breast tomosynthesis compared with digital mammography alone: results of a multicenter, multireader trial. *Radiology* 2013;266: 104-13.

50. Bernardi D, Ciatto S, Pellegrini M, Anesi V, Burlon S, Cauli E, Depaoli M, Larentis L, Malesani V, Targa L, Baldo P, Houssami N. Application of breast tomosynthesis in screening: incremental effect on mammography acquisition and reading time. *Br J Radiol* 2012;85: e1174-8.

51. Tagliafico AS, Calabrese M, Bignotti B, Signori A, Fisci E, Rossi F, Valdora F, Houssami N. Accuracy and reading time for six strategies using digital breast tomosynthesis in women with mammographically negative dense breasts. *Eur Radiol* 2017;27: 5179-84.

52. Bernardi D, Macaskill P, Pellegrini M, Valentini M, Fanto C, Ostillio L, Tuttobene P, Luparia A, Houssami N. Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study. *Lancet Oncol* 2016;17: 1105-13.

53. Skaane P, Bandos AI, Gullien R, Eben EB, Ekseth U, Haakenaasen U, Izadi M, Jebsen IN, Jahr G, Krager M, Hofvind S. Prospective trial comparing full-field digital mammography (FFDM) versus combined FFDM and tomosynthesis in a population-based screening programme using independent double reading with arbitration. *Eur Radiol* 2013;23: 2061-71.

54. Lang K, Nergarden M, Andersson I, Rosso A, Zackrisson S. False positives in breast cancer screening with one-view breast tomosynthesis: An analysis of findings leading to recall, work-up and biopsy rates in the Malmo Breast Tomosynthesis Screening Trial. *Eur Radiol* 2016;26: 3899-907.

55. Friedewald SM, Rafferty EA, Rose SL, Durand MA, Plecha DM, Greenberg JS, Hayes MK, Copit DS, Carlson KL, Cink TM, Barke LD, Greer LN, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. *JAMA* 2014;311: 2499-507.

56. Weinstein MC. Recent developments in decision-analytic modelling for economic evaluation. *Pharmacoeconomics* 2006;24: 1043-53.

57. Caro JJ, Briggs AH, Siebert U, Kuntz KM, Force I-SMGRPT. Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. *Med Decis Making* 2012;32: 667-77.

58. Mandelblatt JS, Stout NK, Schechter CB, van den Broek JJ, Miglioretti DL, Krapcho M, Trentham-Dietz A, Munoz D, Lee SJ, Berry DA, van Ravesteyn NT, Alagoz O, et al. Collaborative Modeling of the Benefits and Harms Associated With Different U.S. Breast Cancer Screening Strategies. *Ann Intern Med* 2016;164: 215-25.

59. 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. *Prev Med* 2011;53: 134-40.

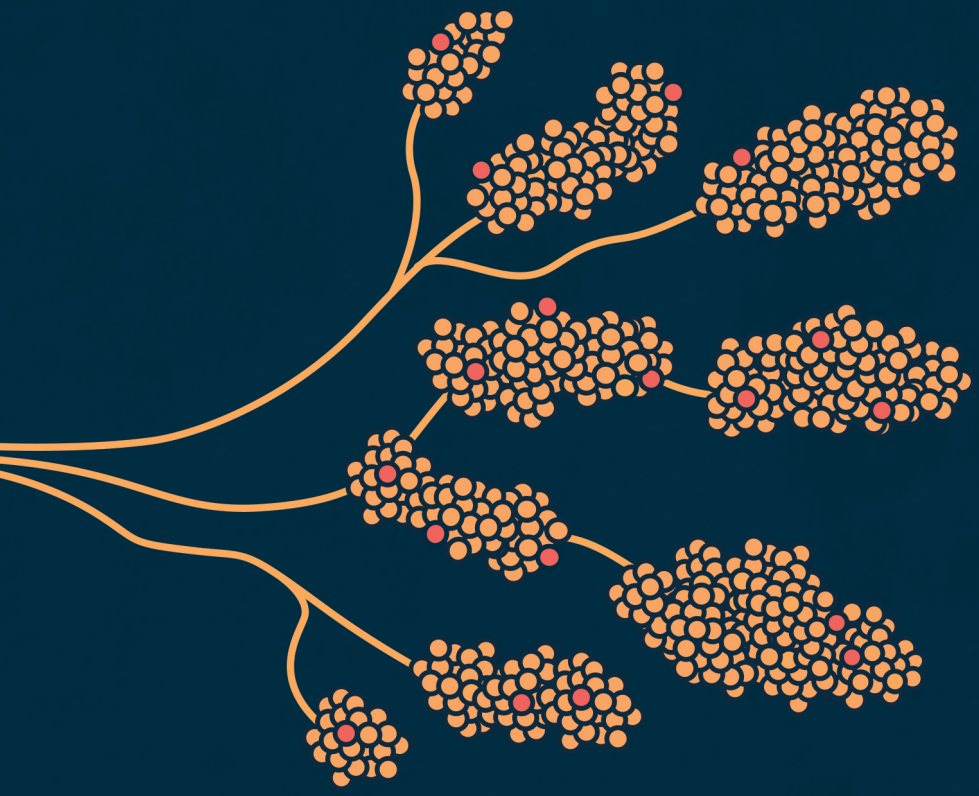
60. Rijnsburger AJ, van Oortmarssen GJ, Boer R, Draisma G, To T, Miller AB, de Koning HJ. Mammography benefit in the Canadian National Breast Screening Study-2: a model evaluation. *Int J Cancer* 2004;110: 756-62.

61. Sankatsing VD, Heijnsdijk EA, van Luijt PA, van Ravesteyn NT, Fracheboud J, de Koning HJ. Cost-effectiveness of digital mammography screening before the age of 50 in The Netherlands. *Int J Cancer* 2015;137: 1990-9.

62. 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.

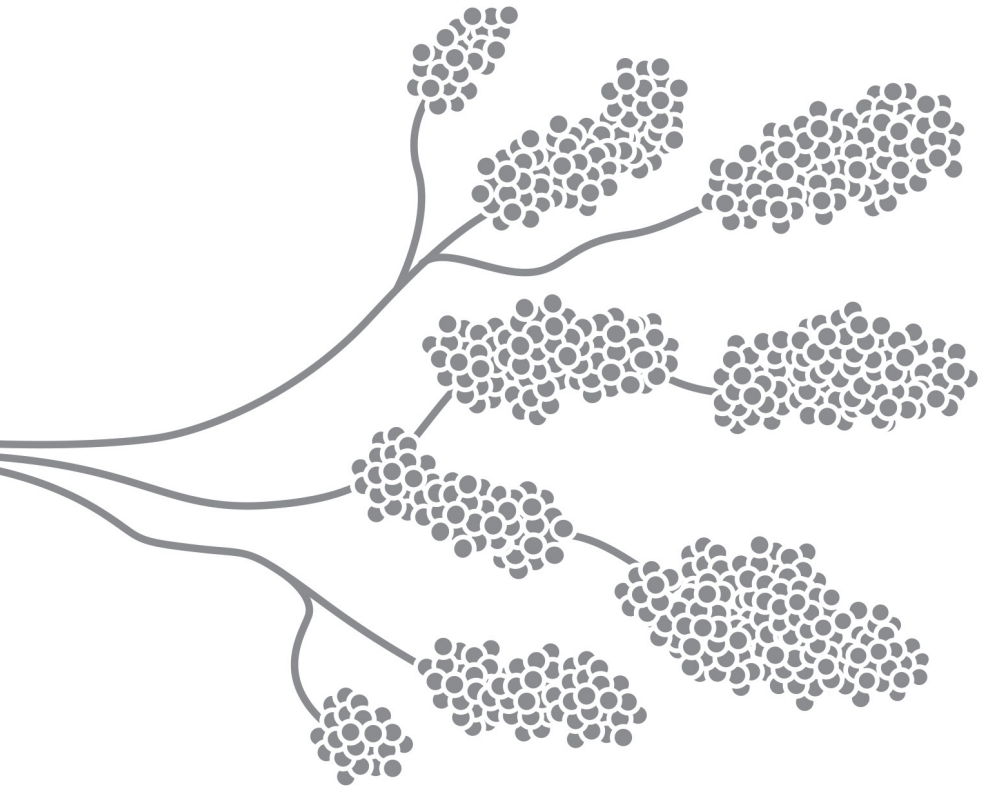
63. Giordano L, von Karsa L, Tomatis M, Majek O, de Wolf C, Lancucki L, Hofvind S, Nystrom L, Segnan N, Ponti A, Eunice Working G, Van Hal G, et al. Mammographic screening programmes in Europe: organization, coverage and participation. *J Med Screen* 2012;19 Suppl 1: 72-82.

64. De Koning HJ. Breast cancer screening; cost-effective in practice? *Eur J Radiol* 2000;33: 32-7.



PART I

Evaluation of current breast cancer screening in the Netherlands



Chapter 2

Detection and interval cancer rates during the transition from screen-film to digital mammography in population-based screening

Valérie D.V. Sankatsing, Jacques Fracheboud, Linda de Munck, Mireille J.M. Broeders, Nicolien T. van Ravesteyn, Eveline A.M. Heijnsdijk, André L.M. Verbeek, Johannes D.M. Otten, Ruud M. Pijnappel, Sabine Siesling, Harry J. de Koning

BMC Cancer 2018;18(1):256.

ABSTRACT

Background: Between 2003 and 2010 digital mammography (DM) gradually replaced screen-film mammography (SFM) in the Dutch breast cancer screening programme (BCSP). Previous studies showed increases in detection rate (DR) after the transition to DM. However, national interval cancer rates (ICR) have not yet been reported.

Methods: We assessed programme sensitivity and specificity during the transition period to DM, analysing nationwide data on screen-detected and interval cancers. Data of 7.3 million screens in women aged 49-74, between 2004-2011, were linked to the Netherlands Cancer Registry to obtain data on interval cancers. Age-adjusted DRs, ICRs and recall rates (RR) per 1000 screens and programme sensitivity and specificity were calculated by year, age and screening modality.

Results: 41,662 screen-detected and 16,160 interval cancers were analysed. The DR significantly increased from 5.13 (95% confidence interval (CI):5.00-5.30) in 2004 to 6.34 (95%CI:6.15-6.47) in 2011, for both in situ (2004:0.73;2011:1.24) and invasive cancers (2004:4.42;2011:5.07), whereas the ICR remained stable (2004: 2.16 (95%CI2.06-2.25);2011: 2.13 (95%CI:2.04-2.22)). The RR changed significantly from 14.0 to 21.4. Programme sensitivity significantly increased, mainly between ages 49-59, from 70.0% (95%CI:68.9-71.2) to 74.4% (95%CI:73.5-75.4) whereas specificity slightly declined (2004:99.1% (95%CI:99.09-99.13);2011:98.5% (95%CI:98.45-98.50)). The overall DR was significantly higher for DM than for SFM (6.24;5.36) as was programme sensitivity (73.6%;70.1%), the ICR was similar (2.19;2.20) and specificity was significantly lower for DM (98.5%;98.9%).

Conclusions: During the transition from SFM to DM, there was a significant rise in DR and a stable ICR, leading to increased programme sensitivity. Although the recall rate increased, programme specificity remained high compared to other countries. These findings indicate that the performance of DM in a nationwide screening programme is not inferior to, and may be even better, than that of SFM.

INTRODUCTION

Sensitivity and specificity are considered to be important quality assurance indicators for the performance of screening. The sensitivity of a breast cancer screening programme (BCSP) is calculated using the detection rate (DR) of screen-detected cancers and the interval cancer rate (ICR). The number of published studies that report interval cancers on a national level is low[1-4]. Data on nationwide interval cancers are difficult to obtain, as an accurate linkage between national screening data and the national cancer registry is required. In addition, because the number of interval cancers can only be determined at the end of an interval between screening rounds, there is an inherent delay in the availability of the data (usually two years), compared to data on cancers detected at screening.

In the past decade, many Western BCSPs made the transition from screen-film mammography (SFM) to digital mammography screening (DM)[5-9]. DM has been shown to influence the performance of BCSPs, leading to higher detection rates than SFM, through increased recall rates[6, 10-13]. In most studies, the increase in cancer detection was largely driven by a significant rise in the detection of DCIS. It has been argued that increased DCIS detection leads to a substantial rise in overdiagnosis of breast cancer without contributing to breast cancer mortality reduction. However, a recently published study showed an association between increased screen-detection of DCIS and fewer subsequent invasive interval cancer cases[14]. DM may thus also have the potential to lower ICRs.

In the Netherlands, the transition from SFM to DM was realised between 2003 and 2010[15, 16]. In the same period, the percentage of 2-view mammography at subsequent screens increased from 50% to over 90%[17, 18]. Several Dutch studies showed statistically significant improvements in cancer detection for DM compared to SFM[13, 19-22], whereas others found no significant differences[16, 23]. However, so far, only regional interval cancer rates during the transition to DM in the Netherlands have been published[16, 21] and programme sensitivity on a national level was therefore not calculated.

The objective of this study was to evaluate the national performance of the BCSP in the Netherlands during the transition period to DM by assessing programme sensitivity and specificity, using screen-detected and interval cancers between 2004 and 2011.

METHODS

Dutch Breast Cancer Screening Programme

The Dutch BCSP is carried out by 5 regional Cancer Screening Organisations (65 screening units), which invite all eligible women based on the population registry, aged 50-74 years, biennially to take part in screening. The attendance rate is around 80%. From 2003 onwards, a pilot phase started in which DM was introduced next to SFM, increasing the proportion of DM from 1% to 7% of all screens in 2007. This period was followed by a roll-out phase in which DM expanded from 10% in 2008, to 42% in 2009 and 100% in June 2010.

We collected data on all screens between 2004 and 2011. At initial screens 2-view mammography, with double reading, was performed. In 2004, about half of the subsequent screens had a second view and this proportion increased to 93% in 2010. The reading policy was double reading with consensus or arbitration. Women were only recalled if both independent readers concluded that the screening mammogram was positive or if a third reader came to this conclusion, in case of disagreement.

Data

All screen-detected and interval cancers between 2004 and 2011 were analysed. To classify cancers as screen-detected or interval cancers, records of all screening examinations were linked to the Netherlands Cancer Registry (NCR). Linkages were made using an algorithm to identify identical subjects with high probability. The NCR classified the positive matches (94% of all breast cancers) preliminarily into screen-detected and interval cancers. Unclassified cancers were checked manually by the Cancer Screening Organisations, using information from the

patient's medical file. A small fraction of all women screened (0.01%) did not give permission to link their records.

Information on whether DM or SFM was performed was derived from the separate screening units, following the rollout schedule for digitization.

Definitions

Screening examinations were defined as mammograms following an invitation to screening. These examinations were subdivided in initial screens, regular subsequent screens within 2.5 years after previous screening and irregular subsequent screens 2.5 years or later after previous screening (4% of all screens between 2004-2011). The latter were not used in this study: as the precise length of the irregular interval could not be determined from the aggregated dataset, including irregular subsequent screens would lead to distortion of (i.e. higher) detection- and interval cancer rates. Positive screens were considered to be screens with a suspicious mammographic lesion leading to recall and negative screens those without suspicious mammographic lesions, without any recommendation. Thus, screen-detected breast cancers were all diagnosed as a direct consequence of recall for further assessment, within one year after a positive screen.

All breast cancers diagnosed within two years after a negative screen were considered to be interval cancers. This concerned cancers arising from:

- Lesions that were screen-detectable at time of screening but were missed or not recalled
- Lesions that were present at screening but had minimal signs and were not recalled
- Lesions that were not present at screening and emerged within the screening interval

Interval cancers could also occur after a false-positive screen: if the cancer detected in the interval did not resemble the earlier detected lesion or was localized in the other breast, it was considered to be an interval cancer and coded accordingly.

Interval cancers were thus calculated using all screens and not only women with a negative screen. Both ductal carcinomas in situ (DCIS) and invasive cancers were included in the number of screen-detected and interval cancers.

We defined programme sensitivity as the number of screen-detected cancers expressed as a proportion of the total number of breast cancers diagnosed in women who were screened, within two years after screening (screen-detected cancers + interval cancers). Programme specificity was defined as the number of negative screens in women without breast cancer as a proportion of the total number of screens in women without a breast cancer diagnosis (true negatives + false-positives), within two years after screening. The false-positive rate was calculated as the number of recalls that did not lead to a breast cancer diagnosis per 1000 screens. As for some recalls the final diagnosis is not known, the numbers of true- and false-positives do not completely add up to the number of recalled women.

Age-adjusted recall (RR), false-positive (FPR), detection (DR) and interval cancer rates (ICR) per 1000 screens were calculated, using the total number of invitations during 2004-2011 as reference population. The positive predictive value (PPV) was calculated as the percentage screen-detected cancers (true positives) of all women recalled (true and false-positives). Performance indicators were based on all screening examinations (initial + regular subsequent), calculated by calendar year and age and presented with 95% confidence intervals (CI).

Analysis

Screening examinations performed at age 75 (N=9,507) and interval cancers diagnosed within two years after screening at age 75 (N=16) were excluded because of small numbers. Results are presented for the age group 49-74 and were calculated for the period 2004-2011, for all screening examinations and for DM and SFM screens separately.

Whether differences in outcomes were statistically significant was determined using the 95% confidence intervals. For proportions these intervals were calculated

using the standard formula (P ± 1.96*s.e.). The 95% confidence intervals for the rates were calculated using a log linear model (exp(b+ log(N)); Poisson distribution) and rates were calculated per 1000 screens.

RESULTS

All screens

Overall results

Between 2004 and 2011, 7343327 screens (initial + regular subsequent) were performed within the Dutch BCSP (Table 1). There were 41662 breast cancers detected by screening; the DR was 5.7 per 1000 screens, of which 0.94 were DCIS. The recall rate (RR) was 17.8 per 1000 screens and the FPR 12.1 (PPV:33.5%). The 16,160 interval cancers identified led to an ICR of 2.2 per 1000, of which 0.1 were DCIS (data not shown). The programme sensitivity was 71.4% and the programme specificity 98.8%.

Trends over time

The DR significantly increased by more than 20% over the study period, from 5.1 per 1000 to 6.3 and the ICR remained stable (Figure 1a; Supplementary material 1a). The DRs of both DCIS (+0.5) and invasive breast cancers (+0.7) increased (Supplementary material 1a). The detection rate increased for all age groups over the entire study period (Fig. 2a; Supplementary material 2a). Detection also increased with age from 55 years onward; in the youngest ages (in particular 49 years) the detection rate was relatively high due to prevalent screening.

The overall ICR remained stable over the study period (2004: 2.2 per 1000 screens; 2011: 2.1; Fig. 1a; Supplementary material 1a). The interval cancer rate showed a slightly decreasing tendency for the younger age groups over the study period and a slight increase in the trend for the older ages (Fig. 2b; Supplementary material 2b). The fluctuation in the overall interval cancer rate was mainly determined by the rate for invasive breast cancers (Fig. 3). There were slight decreases in the age-adjusted overall interval cancer rate in 2007, 2009 and 2011 relative to the previous year (not statistically significant), accompanied by a decline in invasive

interval cancers alone in 2007 and in both invasive and in situ interval cancers in 2009 and 2011 (Fig. 3; Supplementary material 3).

The programme sensitivity increased from 70.0% in 2004 to 74.4% in 2011 (Figure 4a; Supplementary material 1a) and increased statistically significant from 2010 (compared to 2004). The overall programme sensitivity was mainly determined

Table 1. Age-adjusted results for all, DM and SFM examinations between 2004 and 2011 (49-74)

	All (95% C.I.)	**DM** (95% C.I.)	**SFM** (95% C.I.)
No. screens	7343327	2620442	4722885
No. screen-detected cancers	41662	16400	25262
No. interval cancers	16160	5748	10412
No. false-positives	88862	38621	50241
No. recalls	130524	55021	75503
Recall rate	17.8 (17.7-17.9)	21.0 (20.8-21.2)	16.0 (15.9-16.1)
False positive rate	12.1 (12.0-12.2)	14.8 (14.7-15.28)	10.6 (10.5-10.7)
Detection rate (all)	5.7 (5.6-5.7)	6.2 (6.1-6.3)	5.4 (5.3-5.4)
Detection rate DCIS	0.94 (0.92-0.96)	1.1 (1.1-1.2)	0.83 (0.81-0.86)
Detection rate invasive	4.7 (4.7-4.8)	5.1 (5.0-5.2)	4.5 (4.5-4.6)
Interval cancer rate	2.2 (2.2-2.2)	2.2 (2.1-2.3)	2.2 (2.2-2.3)
Programme sensitivity (%)	71.4 (71.1-71.8)	73.6 (73.0-74.2)	70.1 (69.6-70.6)
Programme specificity (%)	98.8 (98.8-98.8)	98.5 (98.5-98.5)	98.9 (98.9-98.9)
Positive predictive value (%)	33.5 (33.2-33.7)	31.5 (31.1-31.9)	34.9 (34.5-35.2)

Rates are presented per 1000 screens

by SFM between 2004-2008 and increased steeply with the expansion of DM between 2008-2011 (Figure 4a; Supplementary material 1b, 1c). The programme sensitivity strongly varied by age in 2004, which attenuated with the expansion of DM due to a significant increase in programme sensitivity for women aged 49-59 (Supplementary material 4). Trends in programme sensitivity of all breast cancers and invasive cancers only were similar between 2004-2008 (Figure 5). In 2009-2010, there was an increase in the sensitivity of all cancers but not of invasive cancers only, which reflects an increased detection of DCIS. In 2011, there was a similar rise in both groups, thus reflecting an increased detection of invasive cancers.

The RR increased significantly over time from 14.0 to 21.4 (Supplementary material 1a). The programme specificity significantly declined slightly from 99.1% to 98.5% (Figure 4b; Supplementary material 1a). The difference in programme specificity between DM and SFM was most prominent in the beginning of the study period and decreased over time.

DM versus SFM

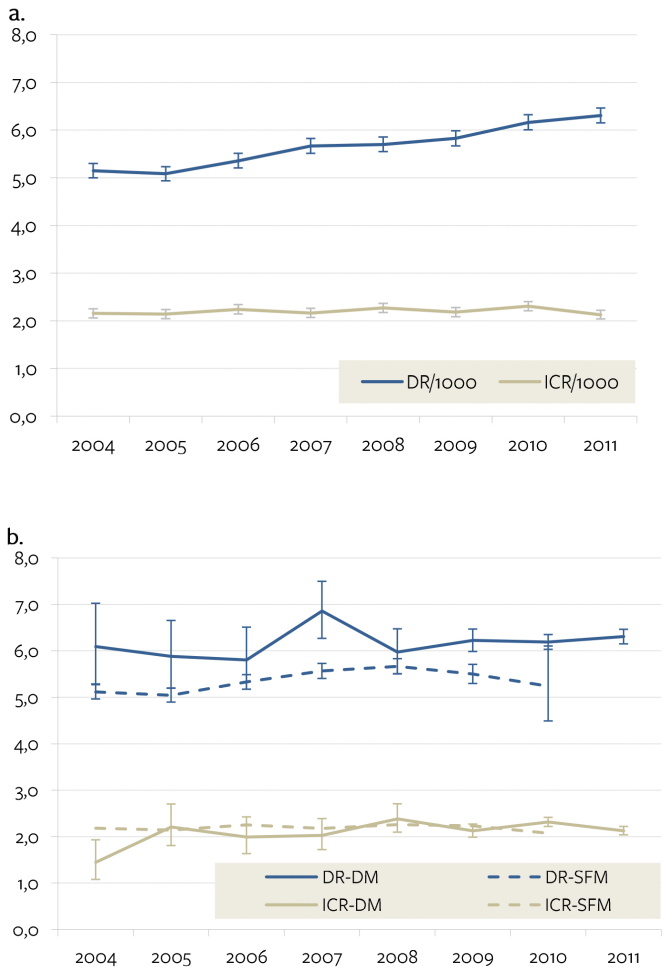
Overall results

Of all screens, 2620442 were DM (36%) and 4722885 SFM (64%; Table 1). The RR for DM was 1.3 times higher than the RR for SFM. The DR was significantly higher for DM than for SFM (6.2 vs. 5.4), leading to higher programme sensitivity (73.6% vs. 70.1%). Both the DR of DCIS and invasive cancers was significantly higher for DM (1.1 and 5.1 respectively) than for SFM (0.83 and 4.5) (Table 1). The PPV and programme specificity were significantly lower for DM (31.5% and 98.5% respectively) than for SFM (34.9% and 98.9%). The ICRs were equal (2.2).

Trends over time

The DR of DM was higher than that of SFM in all years, and significantly higher in 2004, 2007 and 2009 (Figure 1b; Supplementary material 1b, 1c). The ICRs were similar over the years (except for 2004).

Figure 1. Age-adjusted detection and interval cancer rates per 1000 women screened for all screens (a) and DM or SFMa (b)



[a]In 2011 all screens were DM screens. Abbreviations: detection rate (DR); interval cancer rate (ICR); digital mammography (DM); screen-film mammography (SFM)

Figure 2. Age-specific detection (a) and interval cancer rates (b) per 1000 women screened

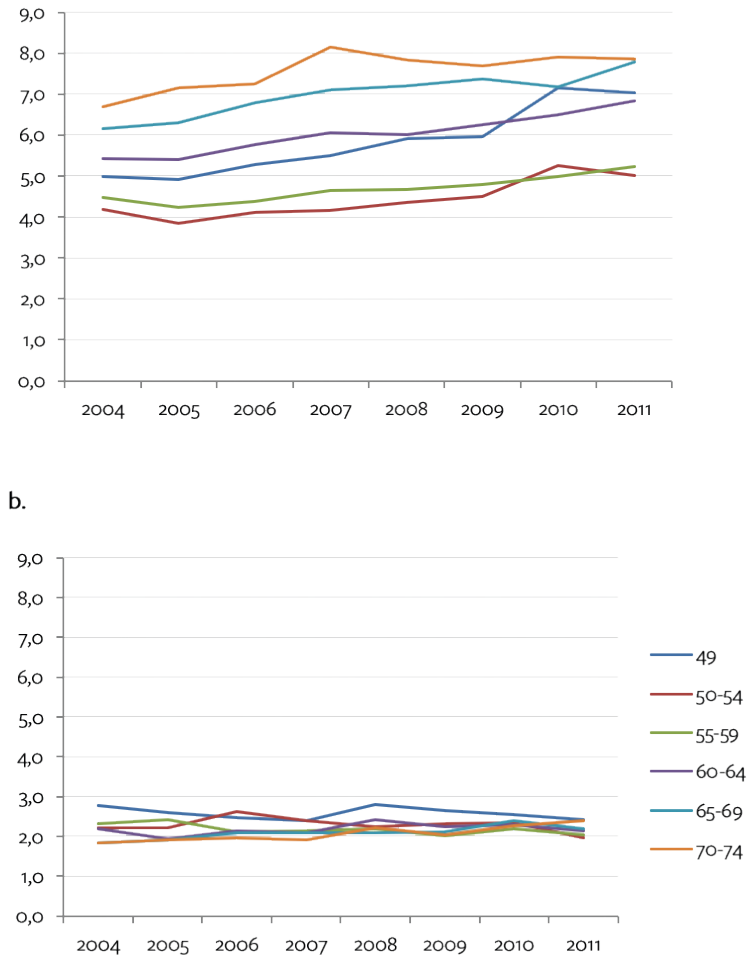


Figure 3. Age adjusted-interval cancer rate (per 1000 women screened) for all, invasive and in situ carcinomas

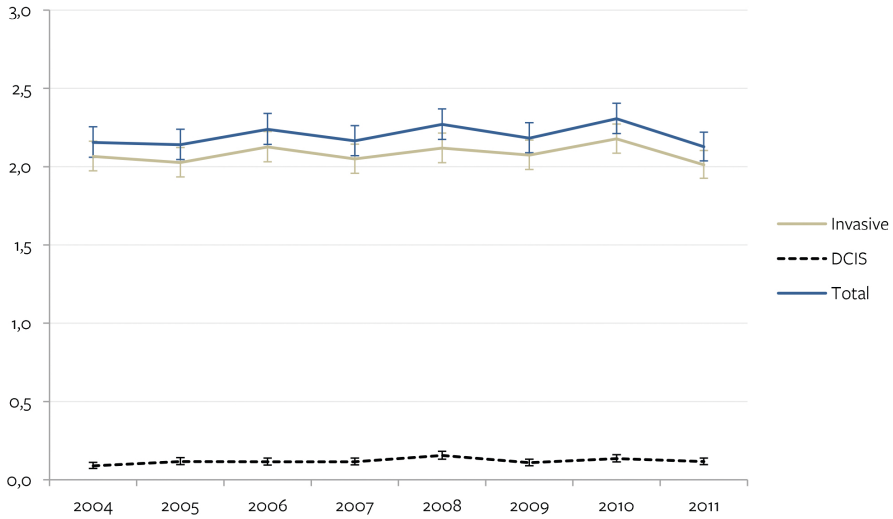
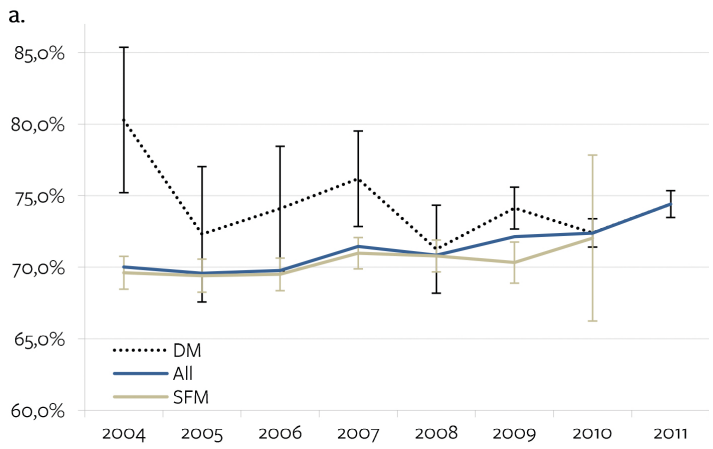
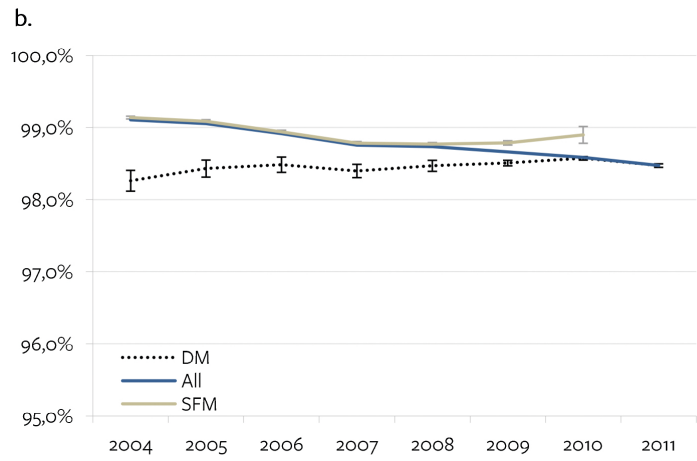


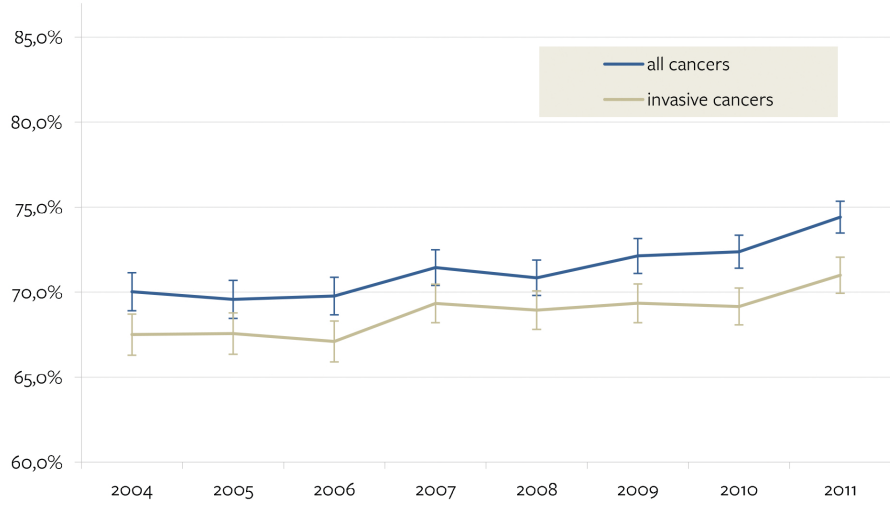
Figure 4. Aged-adjusted programme sensitivity (a) and programme specificity (b) for all screens, DMa and SFM (49–74).





[a]The percentage DM screens between 2004 and 2007 was considerably small; in 2011, all screens were DM screens. N.B. scale Y-axis differs between graph a and b. Abbreviations: digital mammography (DM); screen-film mammography (SFM).

Figure 5. Age-adjusted programme sensitivity for all (invasive + DCIS) and invasive breast cancers only (49–74)



DISCUSSION

This nationwide study shows that the detection rate and programme sensitivity in the Dutch BCSP significantly increased during the transition from SFM to DM. This rise was most prominent for women under age 60. Despite the substantial improvement in detection, there was no decrease in the overall ICR. The programme specificity declined slightly as a result of the increased recall rate. Slight decreases were observable in the trend in interval cancers for younger women. The detection of both DCIS and invasive cancers and programme sensitivity were significantly higher for DM than for SFM, whereas the ICR was similar and the programme specificity was slightly lower for DM.

The increase in cancer detection can be partially explained by the transition to DM. Other studies also reported higher DRs for DM[6, 10, 12, 13]. DM has been demonstrated to lead to a substantially higher DCIS detection compared to SFM in the Netherlands[13, 20, 22]. There have been concerns that the increase in screen-detection of DCIS leads to overdiagnosis rather than to a significant additional reduction in breast cancer mortality[24]. Therefore, some might argue that the rise in breast cancer detection in this study largely reflects overdiagnosis. However, the results of a recent study suggest that for every 1.5-3 screen-detected DCIS cases, one subsequent invasive interval cancer is averted; at levels of DCIS up to 1.5 per 1000 women screened (0.94 in our study)[14]. In addition, our findings show a significant increase in the detection of invasive breast cancers, which are less likely to be overdiagnosed than DCIS. Nevertheless, we recognize that a substantial rise in cancer detection may lead to a somewhat higher absolute number of overdiagnosed cases. Next to the transition to DM however, other factors also contributed to the increase in breast cancer detection. This increase already started in the mid-1990s, far before the introduction of DM[18]. First, the higher DR may also have resulted from an increase in the underlying breast cancer incidence over the years. It has been shown that the underlying breast cancer incidence in the Netherlands increased before the introduction of screening between 1975-1990 in women later invited to screening and in women not yet invited to screening (40-49) before and after the introduction of screening (1975-2004)[25], which has also been reported for other countries[26, 27]. It is reasonable to expect that the rise in background incidence

continued after implementation of screening, due to increases in risk factors for breast cancer, including older age at first pregnancy and menarche and breast feeding at a later stage in life[28-30]. For example, in the Netherlands, the average age at birth of first child has increased from 26 years in 1970 to 29 years in 2004[31].

Second, the significant increase in the percentage of 2-view mammography at subsequent screens during our study period (50% in 2005; >90% in 2011[18]) is likely to have contributed to higher breast cancer detection[17, 32, 33]. Finally, the DR may have increased due to changes in screening protocol. Following the outcomes of a study by Otten et al.[34], the national recall strategy was altered and the RR in the Netherlands increased from 0.9% in 2000 to 1.8% in 2007[18].

We think that the stable interval cancer rate with the increasing trend in detection could also in fact reflect a reduction in the ICR, given the increase in background breast cancer incidence. The rise in detection may have prevented the interval cancer rate to increase as a result of increased breast cancer incidence.

Our estimate for the overall ICR (2.2 per 1000 screens) is in line with earlier reported rates from the BCSP in Germany (2.3)[35] and Norway (1.8)[2]. We found that DM performed significantly better than SFM in terms of DR and programme sensitivity, at the expense of significantly higher RRs and FPRs and slightly lower programme specificity. These findings are also consistent with results of earlier studies[6, 10, 12, 13, 19]. We found RRs (expressed as the percentage of screens recalled for further assessment) of 1.6% for SFM and 2.1% for DM throughout the study period. Recently reported RRs for DM in other European BCSPs range from 2.9% to 6.1%[5-7, 9, 36, 37]. Therefore, RRs in the Netherlands are still rather low compared to other countries[6, 12, 36, 38].

We did not find a difference in ICR between DM and SFM. Similar ICRs for DM and SFM were also reported for other BCSPs[37, 39]. It might be too early to observe the full effect of the transition to DM on the ICR. We observed a small, non-significant, decrease in the overall ICR in 2011 but we need future data, after a few years of full DM screening, to determine whether or not this will turn into a

further statistically significant decline. Although we did not observe a significant difference in the overall ICR, looking at specific age groups we found that the ICR at initial screening in women aged 49-51 years was significantly lower for DM than for SFM (2.3 vs. 2.6 per 1000 screens; Additional file 1 S5). This finding corresponds to the results of the DMIST trial, which showed a higher diagnostic accuracy for DM than for SFM in pre- and perimenopausal women with dense breasts under the age of 50[10].

The major strength of this study was the availability of national data on a large number of interval cancers. Thus, this study is the first nationwide analysis of sensitivity and specificity in the Dutch BCSP during the transition to DM. Furthermore, DM expanded during the second half of the study period and the effect of the transition from SFM to DM could therefore be studied well.

This study also had some limitations. Single screening examinations were not labelled as DM or SFM at time of screening and information about the proportion DM and SFM, during the years in which both modalities were used, had to be obtained from the screening units. The screens for which it was uncertain whether they were performed using screen-film or digital mammography were added to the screen-film group. This could lead to underestimation of detection rates for DM and to increased apparent detections rates for SFM. The difference in detection of DM relative to SFM could thus be (somewhat) greater than we report and our estimates may therefore be conservative. In addition, 2% of all breast cancers in the NCR database could not be classified as screen-detected or interval cancer.

Conclusions

In conclusion, the detection rate in the Dutch breast cancer screening programme substantially increased between 2004 and 2011, whereas the interval cancer rate was stable over time. The recall rate increased over the study period, resulting in a decrease in programme specificity over time, even though the current specificity of the Dutch programme is still relatively high (in international context). DM resulted in higher detection rates than SFM, with similar interval cancer rates. The overall interval cancer rate, slightly, but non significantly declined in younger age groups

and a significant rise in programme sensitivity in women under age 60 years was observed, which may be partly attributable to the transition to DM. Particularly young women may therefore have benefited from the change to DM but further exploration is needed to confirm these findings.

REFERENCES

1. Fracheboud J, de Koning HJ, Beemsterboer PM, Boer R, Verbeek AL, Hendriks JH, van Ineveld BM, Broeders MJ, de Bruyn AE, van der Maas PJ: Interval cancers in the Dutch breast cancer screening programme. *Br J Cancer* 1999, 81(5):912-917.
2. Hofvind S, Yankaskas BC, Bulliard JL, Klabunde CN, Fracheboud J: Comparing interval breast cancer rates in Norway and North Carolina: results and challenges. *J Med Screen* 2009, 16(3):131-139.
3. Bennett RL, Sellars SJ, Moss SM: Interval cancers in the NHS breast cancer screening programme in England, Wales and Northern Ireland. *Br J Cancer* 2011, 104(4):571-577.
4. Dibden A, Offman J, Parmar D, Jenkins J, Slater J, Binysh K, McSorley J, Scorfield S, Cumming P, Liao XH *et al*: Reduction in interval cancer rates following the introduction of two-view mammography in the UK breast screening programme. *Br J Cancer* 2014, 110(3):560-564.
5. 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(4):860-866.
6. Hambly NM, McNicholas MM, Phelan N, Hargaden GC, O'Doherty A, Flanagan FL: Comparison of digital mammography and screen-film mammography in breast cancer screening: a review in the Irish breast screening program. *AJR Am J Roentgenol* 2009, 193(4):1010-1018.
7. Vinnicombe S, Pinto Pereira SM, McCormack VA, Shiel S, Perry N, Dos Santos Silva IM: Full-field digital versus screen-film mammography: comparison within the UK breast screening program and systematic review of published data. *Radiology* 2009, 251(2):347-358.
8. Timmers JM, den Heeten GJ, Adang EM, Otten JD, Verbeek AL, Broeders MJ: Dutch digital breast cancer screening: implications for breast cancer care. *Eur J Public Health* 2012, 22(6):925-929.
9. Domingo L, Romero A, Belvis F, Sanchez M, Ferrer J, Salas D, Ibanez J, Vega A, Ferrer F, Laso MS *et al*: Differences in radiological patterns, tumour characteristics and diagnostic precision between digital mammography and screen-film mammography in four breast cancer screening programmes in Spain. *Eur Radiol* 2011, 21(9):2020-2028.
10. Pisano ED, Hendrick RE, Yaffe MJ, Baum JK, Acharyya S, Cormack JB, Hanna LA, Conant EF, Fajardo LL, Bassett LW *et al*: Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. *Radiology* 2008, 246(2):376-383.

11. Skaane P, Skjennald A: Screen-film mammography versus full-field digital mammography with soft-copy reading: randomized trial in a population-based screening program--the Oslo II Study. *Radiology* 2004, 232(1):197-204.
12. Skaane P: Studies comparing screen-film mammography and full-field digital mammography in breast cancer screening: updated review. *Acta Radiol* 2009, 50(1):3-14.
13. van Luijt PA, Fracheboud J, Heijnsdijk EA, den Heeten GJ, de Koning HJ, National Evaluation Team for Breast Cancer Screening in Netherlands Study G: Nation-wide data on screening performance during the transition to digital mammography: observations in 6 million screens. *Eur J Cancer* 2013, 49(16):3517-3525.
14. Duffy SW, Dibden A, Michalopoulos D, Offman J, Parmar D, Jenkins J, Collins B, Robson T, Scorfield S, Green K *et al*: Screen detection of ductal carcinoma in situ and subsequent incidence of invasive interval breast cancers: a retrospective population-based study. *Lancet Oncol* 2016, 17(1):109-114.
15. Bluekens AM, Karssemeijer N, Beijerinck D, Deurenberg JJ, van Engen RE, Broeders MJ, den Heeten GJ: Consequences of digital mammography in population-based breast cancer screening: initial changes and long-term impact on referral rates. *Eur Radiol* 2010, 20(9):2067-2073.
16. de Munck L, de Bock GH, Otter R, Reiding D, Broeders MJ, Willemse PH, Siesling S: Digital vs screen-film mammography in population-based breast cancer screening: performance indicators and tumour characteristics of screen-detected and interval cancers. *Br J Cancer* 2016, 115(5):517-524.
17. van Breest Smallenburg V, Duijm LE, den Heeten GJ, Groenewoud JH, Jansen FH, Fracheboud J, Plaisier ML, van Doorne-Nagtegaal HJ, Broeders MJ: Two-view versus single-view mammography at subsequent screening in a region of the Dutch breast screening programme. *Eur J Radiol* 2012, 81(9):2189-2194.
18. National Evaluation Team for Breast cancer screening (NETB). National evaluation of breast cancer screening in the Netherlands 1990-2011/2012. *Thirteenth evaluation report.* Erasmus MC University Medical Center Rotterdam, Radboud University Nijmegen Medical Centre, 2014.
19. Bluekens AM, Holland R, Karssemeijer N, Broeders MJ, den Heeten GJ: Comparison of digital screening mammography and screen-film mammography in the early detection of clinically relevant cancers: a multicenter study. *Radiology* 2012, 265(3):707-714.
20. Nederend J, Duijm LE, Louwman MW, Groenewoud JH, Donkers-van Rossum AB, Voogd AC: Impact of transition from analog screening mammography to digital screening mammography on screening outcome in The Netherlands: a population-based study. *Ann Oncol* 2012, 23(12):3098-3103.

21. Nederend J, Duijm LE, Louwman MW, Coebergh JW, Roumen RM, Lohle PN, Roukema JA, Rutten MJ, van Steenbergen LN, Ernst MF *et al*: Impact of the transition from screen-film to digital screening mammography on interval cancer characteristics and treatment - a population based study from the Netherlands. *Eur J Cancer* 2014, 50(1):31-39.

22. Weber RJ, Nederend J, Voogd AC, Strobbe LJ, Duijm LE: Screening outcome and surgical treatment during and after the transition from screen-film to digital screening mammography in the south of The Netherlands. *Int J Cancer* 2015, 137(1):135-143.

23. Karssemeijer N, Bluekens AM, Beijerinck D, Deurenberg JJ, Beekman M, Visser R, van Engen R, Bartels-Kortland A, Broeders MJ: Breast cancer screening results 5 years after introduction of digital mammography in a population-based screening program. *Radiology* 2009, 253(2):353-358.

24. Welch HG, Passow HJ: Quantifying the benefits and harms of screening mammography. *JAMA Intern Med* 2014, 174(3):448-454.

25. 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(1):97-106.

26. Harmer C, Staples M, Kavanagh AM: Evaluation of breast cancer incidence: is the increase due entirely to mammographic screening? *Cancer Causes Control* 1999, 10(5):333-337.

27. Botha JL, Bray F, Sankila R, Parkin DM: Breast cancer incidence and mortality trends in 16 European countries. *Eur J Cancer* 2003, 39(12):1718-1729.

28. Yankaskas BC: Epidemiology of breast cancer in young women. *Breast Dis* 2005, 23:3-8.

29. Collaborative Group on Hormonal Factors in Breast C: Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. *Lancet* 2002, 360(9328):187-195.

30. Helewa M, Levesque P, Provencher D, Lea RH, Rosolowich V, Shapiro HM, Breast Disease C, Executive C, Council SoO, Gynaecologists of C: Breast cancer, pregnancy, and breastfeeding. *J Obstet Gynaecol Can* 2002, 24(2):164-180; quiz 181-164.

31. CBS (Centraal Bureau Statistiek, Statistics Netherlands.

32. Wald NJ, Murphy P, Major P, Parkes C, Townsend J, Frost C: UKCCCR multicentre randomised controlled trial of one and two view mammography in breast cancer screening. *BMJ* 1995, 311(7014):1189-1193.

33. Hackshaw AK, Wald NJ, Michell MJ, Field S, Wilson AR: An investigation into why two-view mammography is better than one-view in breast cancer screening. *Clin Radiol* 2000, 55(6):454-458.

34. Otten JD, Karssemeijer N, Hendriks JH, Groenewoud JH, Fracheboud J, Verbeek AL, de Koning HJ, Holland R: Effect of recall rate on earlier screen detection of breast cancers based on the Dutch performance indicators. *J Natl Cancer Inst* 2005, 97(10):748-754.

35. Heidinger O, Batzler WU, Krieg V, Weigel S, Biesheuvel C, Heindel W, Hense HW: The incidence of interval cancers in the German mammography screening program: results from the population-based cancer registry in North Rhine-Westphalia. *Dtsch Arztebl Int* 2012, 109(46):781-787.

36. Sala M, Comas M, Macia F, Martinez J, Casamitjana M, Castells X: Implementation of digital mammography in a population-based breast cancer screening program: effect of screening round on recall rate and cancer detection. *Radiology* 2009, 252(1):31-39.

37. Hofvind S, Skaane P, Elmore JG, Sebuodegard S, Hoff SR, Lee CI: Mammographic performance in a population-based screening program: before, during, and after the transition from screen-film to full-field digital mammography. *Radiology* 2014, 272(1):52-62.

38. Yankaskas BC, Klabunde CN, Ancelle-Park R, Renner G, Wang H, Fracheboud J, Pou G, Bulliard JL, International Breast Cancer Screening N: International comparison of performance measures for screening mammography: can it be done? *J Med Screen* 2004, 11(4):187-193.

39. Hoff SR, Abrahamsen AL, Samset JH, Vigeland E, Klepp O, Hofvind S: Breast cancer: missed interval and screening-detected cancer at full-field digital mammography and screen-film mammography-- results from a retrospective review. *Radiology* 2012, 264(2):378-386.

SUPPLEMENTARY MATERIAL

1a. Performance indicators (age-adjusted rates per 1000 screens) per calendar year for all screens (95% CI)

	Recall rate	Detection rate	Detection rate DCIS	Detection rate invasive	Interval cancer rate	Programme sensitivity	Programme specificity
2004	14.0 (13.8-14.2)	5.1 (5.0-5.3)	0.73 (0.68-0.79)	4.4 (4.3-4.5)	2.2 (2.1-2.3)	70.0% (68.9-71.2%)	99.1% (99.1-99.1%)
2005	14.5 (14.2-14.7)	5.1 (4.9-5.2)	0.74 (0.69-0.80)	4.3 (4.2-4.5)	2.1 (2.1-2.2)	69.6% (68.5-71.0%)	99.1% (99.0-99.1%)
2006	16.1 (15.9-16.4)	5.4 (5.2-5.5)	0.88 (0.82-0.94)	4.5 (4.3-4.6)	2.2 (2.1-2.3)	69.8% (68.7-70.9%)	98.9% (98.9-98.9%)
2007	18.0 (17.8-18.3)	5.7 (5.5-5.8)	0.84 (0.78-0.90)	4.8 (4.7-5.0)	2.2 (2.1-2.3)	71.4% (70.4-72.5%)	98.8% (98.7-98.8%)
2008	18.2 (18.0-18.5)	5.7 (5.6-5.9)	0.87 (0.81-0.93)	4.8 (4.7-5.0)	2.3 (2.2-2.4)	70.8% (69.8-71.9%)	98.7% (98.7-98.8%)
2009	19.1 (18.8-19.4)	5.8 (5.7-6.0)	1.02 (0.96-1.09)	4.8 (4.7-5.0)	2.2 (2.1-2.3)	72.1% (71.1-73.2%)	98.7% (98.6-98.7%)
2010	20.2 (19.9-20.5)	6.2 (6.0-6.3)	1.2 (1.08-1.22)	5.0 (4.9-5.2)	2.3 (2.2-2.4)	72.4% (71.4-73.4%)	98.6% (98.6-98.6%)
2011	21.4 (21.2-21.7)	6.3 (6.2-6.5)	1.2 (1.17-1.31)	5.1 (4.9-5.2)	2.1 (2.0-2.2)	74.4% (73.5-75.4%)	98.5% (98.5-98.5%)

1b. Performance indicators (age-adjusted rates per 1000 screens) per calendar year for SFM (95% CI)

	Recall rate	Detection rate	Detection rate DCIS	Detection rate invasive	Interval cancer rate	Programme sensitivity	Programme specificity
2004	13.7 (13.4-13.9)	5.1 (5.0-5.3)	0.73 (0.67-0.79)	4.4 (4.3-4.5)	2.2 (2.1-2.3)	69.6% (68.5-70.8%)	99.1% (99.1-99.2%)
2005	14.1 (13.9-14.4)	5.0 (4.9-5.2)	0.75 (0.69-0.81)	4.3 (4.2-4.4)	2.1 (2.0-2.2)	69.4% (68.2-70.6%)	99.1% (99.1-99.1%)
2006	15.8 (15.6-16.1)	5.3 (5.2-5.5)	0.89 (0.83-0.95)	4.4 (4.3-4.6)	2.3 (2.2-2.4)	69.5% (68.4-70.6%)	98.9% (98.9-99.0%)
2007	17.6 (17.3-17.9)	5.6 (5.4-5.7)	0.83 (0.77-0.90)	4.7 (4.6-4.9)	2.2 (2.1-2.3)	71.0% (69.9-72.1%)	98.8% (98.8-98.8%)
2008	17.9 (17.6-18.2)	5.7 (5.5-5.8)	0.88 (0.81-0.94)	4.8 (4.6-4.9)	2.3 (2.2-2.4)	70.8% (69.7-71.9%)	98.8% (98.8-98.8%)
2009	17.5 (17.2-17.9)	5.5 (5.3-5.7)	1.00 (0.91-1.09)	4.5 (4.3-4.7)	2.2 (2.1-2.4)	70.3% (68.9-71.8%)	98.8% (98.8-98.8%)
2010	16.2 (14.8-17.7)	5.2 (4.5-6.1)	0.74 (0.49-1.1)	4.5 (3.8-5.3)	2.1 (1.6-2.6)	72.0% (66.2-77.8%)	98.9% (98.8-99.0%)
2011[a]	-	-	-	-	-	-	-

[a]In 2011, all screens were digital. Abbreviations: screen-film mammography (SFM).

1c. Performance indicators (age-adjusted rates per 1000 screens) per calendar year for DM (95% CI)

	Recall rate	Detection rate	Detection rate DCIS	Detection rate invasive	Interval cancer rate	Programme sensitivity	Programme specificity
2004	23.3 (21.7-25.1)	6.1 (5.3-7.0)	0.87 (0.60-1.27)	5.2 (4.5-6.1)	1.4 (1.08-1.9)	80.3% (75.2-85.4%)	98.3% (98.1-98.4%)
2005	21.4 (20.1-22.9)	5.9 (5.2-6.7)	0.64 (0.44-0.93)	5.2 (4.6-6.0)	2.2 (1.8-2.7)	72.3% (67.6-77.0%)	98.4% (98.3-98.6%)
2006	20.8 (19.6-22.2)	5.8 (5.2-6.5)	0.73 (0.52-1.01)	5.1 (4.5-5.7)	2.0 (1.6-2.4)	74.1% (69.7-78.5%)	98.5% (98.4-98.6%)
2007	22.7 (21.6-23.9)	6.9 (6.3-7.5)	0.85 (0.66-1.1)	6.0 (5.5-6.6)	2.0 (1.7-2.4)	76.2% (72.8-79.5%)	98.4% (98.3-98.5%)
2008	21.1 (20.3-22.1)	6.0 (5.5-6.5)	0.79 (0.63-0.99)	5.2 (4.8-5.7)	2.4 (2.1-2.7)	71.3% (68.2-74.3%)	98.5% (98.4-98.6%)
2009	21.0 (20.6-21.5)	6.2 (6.0-6.5)	1.1 (0.96-1.2)	5.2 (5.0-5.4)	2.1 (2.0-2.3)	74.1% (72.7-75.6%)	98.5% (98.5-98.5%)
2010	20.3 (20.0-20.6)	6.2 (6.0-6.4)	1.2 (1.09-1.2)	5.0 (4.9-5.2)	2.3 (2.2-2.4)	72.4% (71.4-73.4%)	98.6% (98.6-98.6%)
2011	21.4 (21.2-21.7)	6.3 (6.2-6.5)	1.24 (1.2-1.3)	5.1 (4.9-5.2)	2.1 (2.0-2.2)	74.4% (73.5-75.4%)	98.5% (98.5-98.5%)

Abbreviations: digital mammography (DM).

2a. Age-specific detection rate per 1000 screens by calendar year

	2004	2005	2006	2007	2008	2009	2010	2011
49	4.98	4.91	5.27	5.50	5.92	5.95	7.16	7.04
50-54	4.18	3.84	4.11	4.17	4.35	4.51	5.26	5.01
55-59	4.47	4.23	4.38	4.65	4.67	4.79	4.99	5.23
60-64	5.43	5.41	5.76	6.05	6.01	6.25	6.49	6.83
65-69	6.16	6.31	6.79	7.11	7.20	7.38	7.19	7.78
70-74	6.69	7.14	7.25	8.15	7.82	7.68	7.91	7.87
C.I. low								
49	4.38	4.32	4.66	4.87	5.27	5.30	6.47	6.35
50-54	3.92	3.59	3.85	3.90	4.08	4.23	4.97	4.73
55-59	4.19	3.96	4.11	4.37	4.38	4.50	4.69	4.93
60-64	5.08	5.06	5.40	5.70	5.67	5.91	6.15	6.48
65-69	5.75	5.90	6.36	6.68	6.76	6.94	6.77	7.35
70-74	6.22	6.67	6.76	7.64	7.32	7.18	7.41	7.38
C.I. high								
49	5.67	5.58	5.96	6.21	6.64	6.68	7.93	7.80
50-54	4.76	4.11	4.39	4.45	4.63	4.80	5.57	5.31
55-59	5.09	4.51	4.66	4.95	4.97	5.10	5.30	5.55
60-64	6.18	5.78	6.14	6.43	6.37	6.61	6.84	7.19
65-69	7.01	6.75	7.25	7.57	7.66	7.85	7.63	8.24
70-74	7.61	7.66	7.78	8.70	8.36	8.21	8.44	8.39

Abbreviations: confidence interval (C.I.)

2b. Age-specific interval cancer rate per 1000 screens by calendar year

	2004	2005	2006	2007	2008	2009	2010	2011
49	2.77	2.60	2.47	2.38	2.80	2.64	2.54	2.42
50-54	2.21	2.21	2.63	2.38	2.25	2.31	2.33	1.96
55-59	2.31	2.41	2.12	2.14	2.19	2.01	2.19	2.03
60-64	2.20	1.93	2.14	2.09	2.41	2.25	2.30	2.15
65-69	1.83	1.92	2.09	2.10	2.09	2.12	2.39	2.20
70-74	1.82	1.91	1.95	1.90	2.22	2.02	2.27	2.38
C.I. low								
49	2.33	2.18	2.06	1.98	2.37	2.22	2.14	2.04
50-54	2.02	2.02	2.42	2.18	2.06	2.11	2.14	1.79
55-59	2.12	2.21	1.94	1.95	1.99	1.83	2.00	1.84
60-64	1.98	1.73	1.92	1.89	2.20	2.05	2.10	1.96
65-69	1.61	1.70	1.86	1.87	1.86	1.89	2.15	1.98
70-74	1.59	1.67	1.71	1.66	1.96	1.78	2.02	2.12
C.I. high								
49	3.30	3.10	2.96	2.87	3.32	3.14	3.02	2.89
50-54	2.63	2.42	2.85	2.59	2.46	2.52	2.54	2.15
55-59	2.75	2.62	2.32	2.35	2.40	2.22	2.40	2.23
60-64	2.61	2.16	2.38	2.31	2.65	2.47	2.52	2.35
65-69	2.18	2.17	2.36	2.36	2.34	2.38	2.65	2.45
70-74	2.17	2.18	2.23	2.17	2.52	2.31	2.57	2.67

Abbreviations: confidence interval (C.I.)

3. Age standardised interval cancer rates for all cancers, invasive cancers and in situ cancers (95% confidence interval)

	2004	2005	2006	2007	2008	2009	2010	2011
Invasive	2.07 (1.97-2.16)	2.03 (1.94-2.12)	2.13 (2.03-2.22)	2.05 (1.96-2.14)	2.12 (2.03-2.21)	2.07 (1.98-2.17)	2.18 (2.09-2.27)	2.01 (1.93-2.10)
DCIS	0.09 (0.07-0.11)	0.12 (0.10-0.14)	0.11 (0.09-0.14)	0.12 (0.10-0.14)	0.15 (0.13-0.18)	0.11 (0.09-0.13)	0.14 (0.11-0.16)	0.12 (0.10-0.14)
Overall	2.16 (2.06-2.25)	2.14 (2.05-2.24)	2.24 (2.14-2.34)	2.16 (2.07-2.26)	2.27 (2.17-2.37)	2.18 (2.09-2.28)	2.31 (2.21-2.40)	2.13 (2.04-2.22)

4. Age-specific programme sensitivity of all screens (based on first two years after a screen)

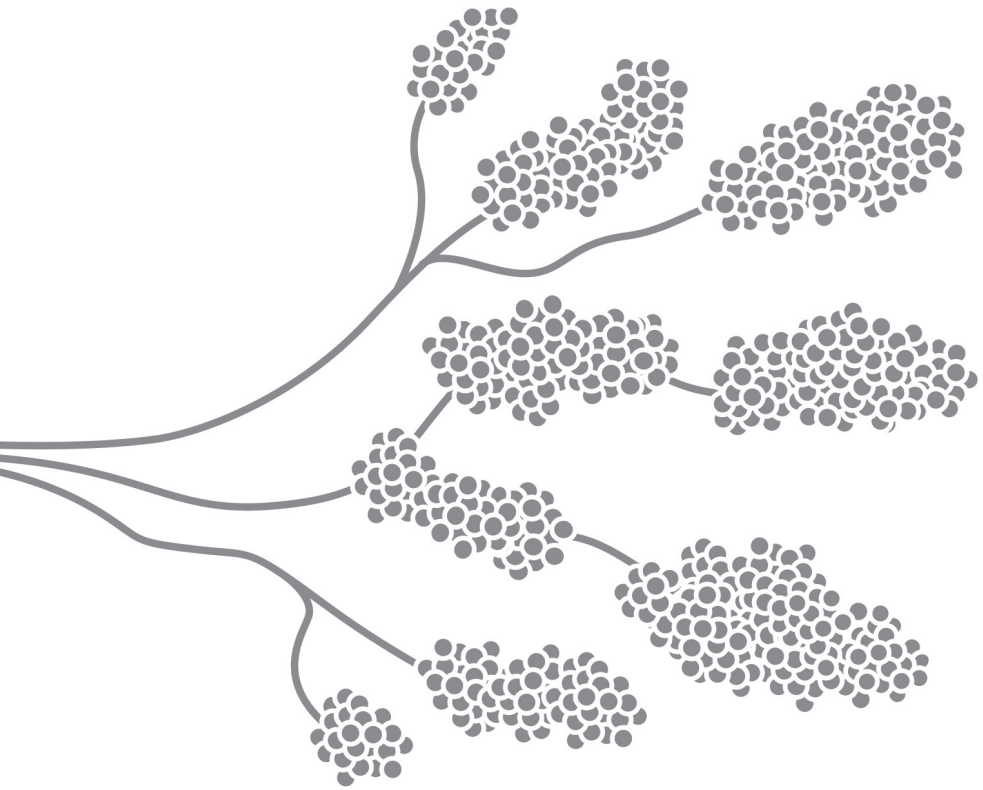
	2004	2005	2006	2007	2008	2009	2010	2011
49	64.2%	65.4%	68.1%	69.8%	67.8%	69.3%	73.8%	74.4%
50-54	65.4%	63.4%	61.0%	63.6%	65.9%	66.1%	69.3%	71.8%
55-59	65.9%	63.7%	67.3%	68.5%	68.1%	70.4%	69.4%	72.1%
60-64	71.2%	73.7%	72.9%	74.4%	71.3%	73.6%	73.8%	76.1%
65-69	77.1%	76.7%	76.4%	77.2%	77.5%	77.7%	75.1%	78.0%
70-74	78.6%	78.9%	78.8%	81.1%	77.9%	79.1%	77.7%	76.8%
C.I. low								
49	59.3%	60.5%	63.4%	65.1%	63.4%	64.9%	69.9%	70.5%
50-54	62.9%	60.8%	58.5%	61.1%	63.4%	63.7%	67.1%	69.6%
55-59	63.5%	61.3%	64.9%	66.1%	65.7%	68.0%	67.1%	69.9%
60-64	68.6%	71.2%	70.5%	72.1%	69.1%	71.4%	71.8%	74.2%
65-69	74.5%	74.2%	74.0%	74.9%	75.2%	75.4%	72.8%	75.9%
70-74	76.0%	76.4%	76.3%	78.8%	75.4%	76.7%	75.3%	74.4%
C.I. high								
49	69.2%	70.3%	72.8%	74.4%	72.3%	73.7%	77.7%	78.2%
50-54	67.9%	66.0%	63.5%	66.1%	68.3%	68.5%	71.5%	74.0%
55-59	68.3%	66.2%	69.8%	70.9%	70.5%	72.8%	71.8%	74.3%
60-64	73.8%	76.2%	75.4%	76.6%	73.5%	75.7%	75.8%	78.0%
65-69	79.6%	79.2%	78.9%	79.5%	79.8%	79.9%	77.3%	80.0%
70-74	81.2%	81.4%	81.3%	83.4%	80.3%	81.6%	80.0%	79.1%

Abbreviations: confidence interval (C.I.)

5. Age-adjusted results for all, DM and SFM between 2004 and 2011 for women aged 49-51

	All (95% C.I.)	**DM** (95% C.I.)	**SFM** (95% C.I.)
No. screens	777908	273266	504642
No. screen-detected cancers	4667	1920	2747
No. interval cancers	1936	648	1288
No. false-positives	23682	10762	12920
Recall rate	33.6 (33.2-34.0)	47.7 (46.9-48.6)	31.9 (31.4-32.4)
False positive rate	27.7 (27.3-28.1)	40.4 (39.9-40.9)	26.3 (25.9-26.8)
Detection rate (all)	5.9 (5.7-6.1)	7.3 (6.97-7.61)	5.7 (5.5-5.9)
Detection rate DCIS	1.1 (1.03-1.2)	1.9 (1.81-2.00)	1.2 (1.1-1.3)
Detection rate invasive	4.8 (4.7-5.0)	5.4 (5.2-5.6)	4.5 (4.3-4.7)
Interval cancer rate	2.5 (2.4-2.7)	2.3 (2.2-2.4)	2.6 (2.5-2.8)
Programme sensitivity (%)	69.7 (68.6-70.8)	76.3 (74.6-77.9)	68.5 (67.1-70.0)
Programme specificity (%)	97.2 (97.2-97.2)	95.9 (95.8-96.0)	97.4 (97.3-97.4)
Positive predictive value (%)	16.5 (16.2-16.8)	15.1 (14.9-15.4)	17.5 (17.2-17.8)

Rates are presented per 1000 screens



Chapter 3

The effect of population-based mammography screening in Dutch municipalities on breast cancer mortality: 20 years of follow-up

Valérie D.V. Sankatsing, Nicolien T. van Ravesteyn, Eveline A.M. Heijnsdijk, Caspar W.N. Looman, Paula A. van Luijt, Jacques Fracheboud, Gerard J. den Heeten, Mireille J.M. Broeders, Harry J. de Koning

International Journal of Cancer 2017;141(4):671-677.

ABSTRACT

Background: Long-term follow-up data on the effects of screening are scarce and debate exists on the relative contribution of screening versus treatment to breast cancer mortality reduction. Our aim was therefore to assess the long-term effect of screening by age and time of implementation.

Methods: We obtained data on 69,630 breast cancer deaths between 1980 and 2010 by municipality (N=431) and age of death (40-79) in the Netherlands. Breast cancer mortality trends were analysed by defining the municipality-specific calendar year of introduction of screening as year 0. Additionally, log-linear Poisson regression was used to estimate the turning point in the trend after year 0, per municipality, and the annual percentage change (APC) before and after this point.

Results: Twenty years after introduction of screening breast cancer mortality was reduced by 30% in women aged 55-74 and by 34% in women aged 75-79, compared to year 0. A similar and significant decrease was present in municipalities that started early (1987-1992) and late (1995-1997) with screening, despite the difference in availability of effective adjuvant treatment. In the age groups 55-74 and 75-79 the turning point in the trend in breast cancer mortality was estimated in year 2 and 6 after the introduction of screening respectively, after which mortality decreased significantly by 1.9% and 2.6% annually.

Conclusions: These findings show that the implementation of mammography screening in Dutch municipalities is associated with a significant decline in breast cancer mortality in women aged 55-79, irrespective of time of implementation.

INTRODUCTION

Mammography screening has been shown to reduce breast cancer mortality in various randomised controlled trials[1-4]. Following the outcomes of these trials, population-based screening programmes have been implemented in many European countries[5]. In the Netherlands, mammography screening for women aged 50 to 69 was gradually implemented between 1987 and 1997 and extended to age 74 between 1998 and 2001.

Numerous observational studies have assessed the impact of population-based screening on breast cancer mortality[6-10]. The International Agency for Research on Cancer (IARC) has recently released a report on breast cancer screening in which the value of these studies was assessed[11]. Based on observational studies that were considered to be "informative for evaluating the effectiveness of mammographic screening programmes", the experts concluded that there is sufficient evidence for the effectiveness of mammography screening in reducing breast cancer mortality in women aged 50 to 69 and 70 to 74. The studies supported an average 23% reduction in breast cancer mortality in 50 to 69 year old women invited to screening and a reduction of approximately 40% in women who attended screening[11].

Trends in breast cancer mortality rates are often analysed to estimate the effectiveness of screening programmes. In order to assess the effect properly, long follow-up after full screening coverage is crucial. However, trend studies with adequate follow-up are scarce[12]. The aim of this study was to assess the effect of population-based screening on breast cancer mortality rates in the Netherlands between 1980 and 2010 using a long follow-up of 20 years after the start of nationwide screening (corresponding to 13 years after full screening coverage was achieved) and approximately 10 years after the extension of screening from age 69 to age 74. By analysing breast cancer mortality between 1980 and 2010, this study will add 10 years of follow-up to previous work[6]. In addition, the long(er) follow-up after the extension of screening allowed us to assess the full impact of screening between age 50 and 74. Furthermore, breast cancer mortality was analysed at the municipality level and we could, therefore, account explicitly for the municipality-specific introduction year of screening.

Apart from screening, adjuvant therapy also contributes to a reduction in breast cancer mortality[13-16]. Adjuvant therapy was introduced in the Netherlands in the early 1980's for women aged 50 years and older[17] and has improved over time[14]. The proportion of breast cancer patients receiving adjuvant therapy in the Netherlands has increased substantially since the early 1980's[17-20]. It has been argued that more effective therapy will lead to a smaller screening effect[21]. We therefore also compared trends in breast cancer mortality rates in municipalities that started early with screening to trends in late starting municipalities, in which more effective treatment was available.

METHODS

Dutch breast cancer screening programme

The breast cancer screening programme for women aged 50-69 in the Netherlands was initiated in 1987-89 in municipalities adjacent to two pilot municipalities (Utrecht and Nijmegen). From 1990 onwards, the screening programme was gradually implemented throughout the country, reaching full coverage in 1997. The programme was extended to age 74 between 1998 and 2001. Due to the gradual implementation, there were differences in calendar year of introduction of screening between municipalities. In 2003, a pilot study was conducted in which digital mammography was introduced. The proportion of digital mammography screens steadily increased from 1% in 2003 to 100% in June 2010.

Women aged 50-74 years are invited to screening once every two years in the Netherlands. Mammograms are read independently by two readers. Initially, 2-view mammography was only performed at first screening with 1-view mammography at subsequent screens. However, 2-view mammography was increasingly performed at subsequent screens over the years, from around 50% of all screens in 2004 to over 90% in 2010[22, 23]. The attendance rate in the Dutch screening programme has been around 80% since the introduction of screening, with only small differences between age groups[22]. The recall rate was 9.9 per 1000 women screened during the introduction of screening, between 1990-1997, and increased to 21.4 per 1000 in 2011[22].

Between 1980 and 2010, the number of municipalities decreased from 812 to 431, due to mergers of municipalities. Our dataset included all municipalities, and thus consisted of municipalities that remained unchanged between 1980 and 2010, municipalities that were merged, but had the same calendar year of introduction of screening and a small number of municipalities that were merged and had different calendar years of introduction of screening. In the latter group the linking of breast cancer deaths occurring after the merger of municipalities to the correct calendar year of introduction of screening is not straightforward and is described in the Methods of the Supplementary material.

Data

Breast cancer deaths between 1980 and 2010 and population data by calendar year, age at death (ages 40-79) and municipality were obtained from Statistics Netherlands[24]. All deaths in the Netherlands are recorded by Statistics Netherlands by cause of death and last known address. The cause of death is assessed by use of the death certificate with the indication of death determined by the attending physician.

Based on evidence from the literature, many have argued to expect a delay of 3-5 years after first invitation to screening (age 50) in the full effect of screening on breast cancer mortality[25]. We therefore used the following age groups: 55-74 (referred to as 'women invited to screening'); 40-54 ('younger women'); 75-79 ('older women').

Statistical analyses

To correct for the gradual implementation of screening, we transformed calendar years (1980-2010) to years relative to the municipality-specific year of introduction of screening, defined as year 0. Preceding and subsequent years, before and after the introduction of screening, were indicated as year -1, -2, -3 etc. and 1, 2, 3 etc. respectively. Results are thus presented by year since introduction of screening instead of calendar year. Because of the gradual implementation, the maximum follow-up period of a municipality after the introduction of screening ranges from

13 (introduction of screening in 1997) to 23 (introduction of screening in 1987) years.

Breast cancer mortality trends were analysed using log-linear Poisson regression. We calculated breast cancer mortality rates per 100 000 woman-years, using the number of breast cancer deaths as the numerator and the mid-year female population size as the denominator. Age-standardisation[26] was done to allow for (visual) comparability of mortality rates. In addition, we calculated breast cancer mortality rates relative to the rate at time of introduction of screening in women invited to screening (55-74), for each year before and after the introduction of screening, by using the breast cancer mortality rate in year 0 as a reference value (value 1 in the graph). The relative breast cancer mortality rate in a given year thus reflects the proportion of breast cancer mortality compared to the breast cancer mortality rate in year 0, in which screening was introduced. Corresponding 95% confidence intervals (CIs) were also calculated.

We performed an additional analysis by using log-linear Poisson regression (on the number of breast cancer deaths and woman-years) to assess the best estimate for a knot (referred to as 'turning point') in the trend in breast cancer mortality after the introduction of screening (year 0), using the municipality-specific calendar year of introduction of screening. This turning point reflects the most probable year in which the mortality trend changes after the introduction of screening. Slopes before and after the turning point were estimated to calculate the annual percentage change (APC) and the corresponding 95% CI. The methods and results of this analysis are extensively described in the Methods of the Supplementary material. All analyses were performed using R 3.0.2.

Early, intermediate and late starters

We divided municipalities into three groups of comparable size: early starters (introduction of screening between 1987 and 1992), intermediate starters (introduction between 1993 and 1994) and late starters (introduction between 1995 and 1997). Adjuvant therapy was administered to 15% of patients in the Netherlands between 1975 and 1984 versus 49% of patients between 1995 and

2004[19]. We therefore assumed that late starters have higher treatment coverage and more effective treatment available in the years before and after the introduction of screening than early and intermediate starters. We compared breast cancer mortality trends in early, intermediate and late starting municipalities for the age group invited to screening (55-74) and we used log-linear Poisson regression to estimate the turning point and the slopes before and after this point (For results turning points see Supplementary material).

RESULTS

Our dataset consisted of 431 municipalities in 2010 and 69 630 breast cancer deaths between 1980 and 2010 (Table 1). The distribution of breast cancer deaths among age groups and early, intermediate and late starters is also shown in Table 1.

Table 1. Distribution of breast cancer deaths among age groups and starter groups, before and after the municipality-specific introduction of screening

	Number of breast cancer deaths		
	Total number	Before introduction of screening (since 1980)	After introduction of screening (until 2010)[a]
All ages (40-79)	69 630	26 557	43 073
40-54	17 911	6 541	11 370
55-74	40 559	15 932	24 627
75-79	11 160	4 084	7 076
All starters (1987-1997)[b]	39 563	15 932	23 631[c]
1987-1992 (early)	17 430	6 077	11 353
1993-1994 (intermediate)	13 380	5 668	7 712
1995-1997 (late)	8 753	4 187	4 566

[a]Including the year of introduction (year 0). [b]Restricted to age 55-74 years. [c]Lower number than for the group 55-74 because Utrecht and Nijmegen (introduction in 1974 and 1975) were not included.

The age-standardised breast cancer mortality rates by age group relative to the year in which screening was introduced at the municipality level (year 0) are displayed in Figure 1. In women invited to screening (55-74), there is a slight increase in breast cancer mortality before the introduction of screening at the municipality level followed by a strong decrease afterwards. A similar decline is visible for women aged 75-79, a few years later (women aged 70-74 were on average invited to screening 5 years later than women aged 50-69). In women aged 40-54, the mortality rate starts to decline moderately a few years before the introduction of screening (Figure 1). The underlying rates per year relative to the introduction of screening are listed in the Supplementary material (Supplementary material 1a, 1b).

The relative breast cancer mortality rate for women aged 55-74 shows a reduction in breast cancer mortality 20 years after the introduction of screening of 30% (Figure 2). In the age group 75-79, there was a significant 34% reduction in year 20, compared to year 0 (Figure 1).

The turning point in the breast cancer mortality trend was estimated in year 2 after the introduction of screening in women aged 55-74 years and in year 6 after introduction in women aged 75-79 years, with corresponding significant annual decreases of 1.9% and 2.6% respectively from these points onwards (Supplementary material 2).

Early, intermediate and late starters

The age-standardised mortality rates, for women invited to screening (age-group 55-74), are similar for early, intermediate and late starters before and after the introduction of screening. All three groups show a strong and similar decline in breast cancer mortality related to the introduction of screening at the municipality level, despite the (expected) difference in therapy regimens (Figure 3).

The trend in breast cancer mortality was estimated to change from a stable to a significantly declining rate in year 2 (early starters; 1987-1992) and year 4 (intermediate and late starters; 1993-1994 and 1995-1997) after the introduction of

screening. After the turning point the mortality rates in the three groups declined significantly between 2.2% and 2.9% per year (Supplementary material 2).

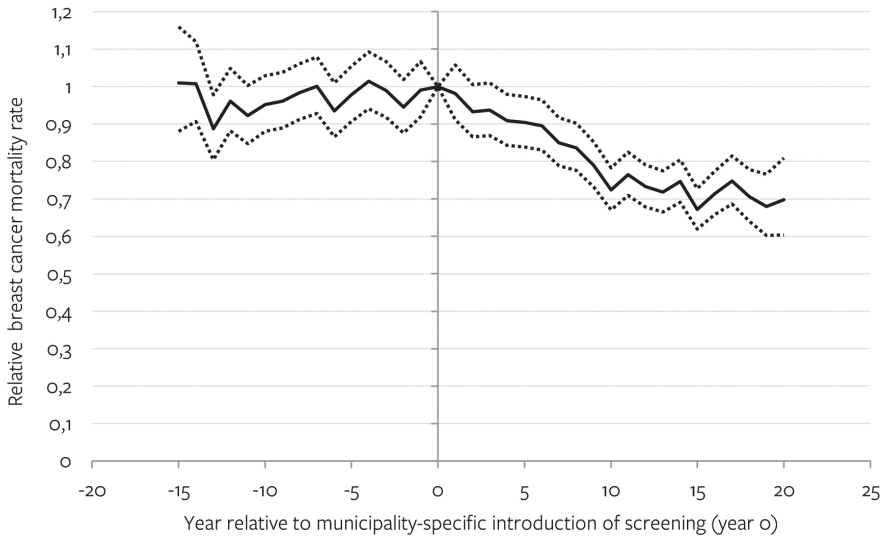
Figure 1. Age-standardised breast cancer mortality rates for age groups 40-54, 55-74 and 75-79 by year relative to introduction of screening

DISCUSSION

Our results show a strong change in breast cancer mortality trends in women aged 55-74 and 75-79, related to the municipality-specific introduction of screening in the age groups 50-69 and 70-74. The results support that mammography screening from age 50 to 74 contributes to a reduction in breast cancer mortality. The decline in breast cancer mortality in women aged 55-74 was estimated to be 30% after 20 years of follow-up since the introduction of screening. For these women we estimated a significant annual decline of 1.9% over 11-21 years of follow-up, from year 2 after the introduction of screening onwards. Breast cancer mortality in older women (75-79) was reduced by 34% after 20 years of follow-up and was estimated

to decrease significantly by 2.6% annually over 7-17 years of follow-up, from year 6 after the introduction of screening onwards.

Figure 2. Relative breast cancer mortality rates for age group 55-74 by year relative to introduction of screening

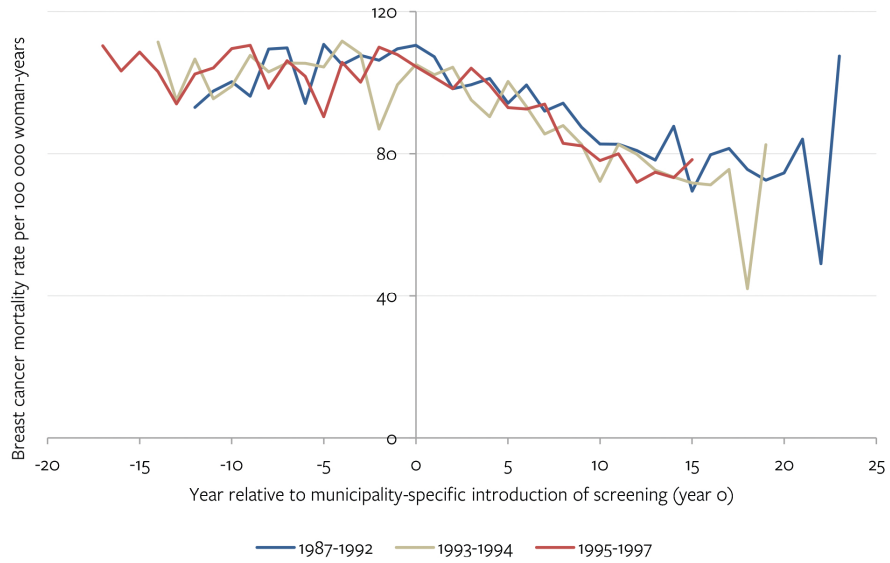


Dashed lines represent 95% CI.

The observed start of the decline in breast cancer mortality (and the estimated turning point in year 2) might seem early and in conflict with the expectation that the full effect of screening cannot be observed before 3-5 years after the first invitation to screening (we only accounted for a time delay in age, not for a delay in calendar time). However, the estimate of the delay in the effect of screening is largely based on results of the randomized controlled trials of breast cancer screening, which started in the early 1980's. As the number of breast cancer deaths in our study is substantially higher than in any of the trials, which had around 750-2000 breast cancer deaths[1, 27, 28], it is plausible that the effect of screening already occurs before 3 years after its introduction. Also, at the time of introduction of screening, there may have been women with metastasized breast cancer who were not yet diagnosed and who would have died within two years in the absence

of screening. Screen-detection may have averted death in some of these cases, which could have led to a rapid effect on breast cancer mortality rates in terms of mortality reduction. More importantly, a change in the trend in breast cancer mortality, from a stable to a downwards trend, after the introduction of screening is likely to occur before the point in time at which the full effect of screening can be observed.

Figure 3. Age-standardised breast cancer mortality rates for early (1987-1992), intermediate (1993-1994) and late (1995-1997) starters, for age group 55-74



A major strength of our study is the long period of follow-up of more than 10 years, both after full coverage of the screening programme was achieved in 1997 and from the extension of screening from age 69 to age 74 in 1998. Another strength is that we had nationwide coverage, as we obtained mortality data of all 431 municipalities in the Netherlands in 2010, and 69 630 breast cancer deaths.

Trend analyses of population mortality rates over time are often considered to have drawbacks with respect to the quantification of the effect of screening[12]. Our study

was carefully designed to account for some of these drawbacks. We accounted for differences in calendar year of introduction of screening between municipalities by analysing our data by year relative to introduction of screening at the municipality level rather than by calendar year. An observed effect of screening may be diluted by breast cancer deaths in women diagnosed with breast cancer before screening was introduced (prevalent tumours). This may, however, not be a major issue in our analysis as we assessed breast cancer mortality in the age group 55-74 (rather than 50-74) and we calculated the annual percentage change in breast cancer mortality rates after the estimated turning point (rather than directly after the introduction of screening). These aspects of study design have been argued to limit the bias resulting from inclusion of death from prevalent tumours[12]. Nevertheless, there will be some dilution of the screening effect in the first years after the introduction of screening. However, it is plausible that the turning point in the trend occurs before the full effect of screening can be observed, even if there is dilution in the early years after the introduction of screening.

Our study had some limitations. As women who participated in screening, at the time it was introduced, in a certain municipality may have moved to another municipality afterwards, the correlation between the municipality-specific introduction year of screening and female inhabitants will become smaller over time. In addition, a study based on individual data would have been more accurate and reliable. However this was not possible due to privacy legislation. Furthermore, in the Netherlands participation is relatively high (on average 80% in each screening round)[22], but, because we had no individual data on screening participation, the screening effect is diluted by non-participation. Finally, the comparisons between the trends in the age group invited to screening (55-74) and younger women (40-54) should be considered with caution as the administration of adjuvant treatment before and after the introduction of screening differed between these age groups in the Netherlands[17].

After introduction of screening, we observed a 30% reduction in breast cancer mortality in women invited to screening which is consistent with reduction estimates in the age group invited to population-based screening from other

studies[29, 30] and in agreement with the 25% reduction shown by the randomized trials of mammography screening[25]. Our results are in line with trend studies, with adequate follow-up after complete coverage of screening, which demonstrate a strong association between the introduction of mammography screening and breast cancer mortality reduction[8, 29]. Furthermore, we estimated an annual decline in breast cancer mortality of 1.9% from year 2 after the introduction of screening onwards in women aged 55-74, which corresponds to the previously reported 1.7% annual decline, after the introduction of screening in the Netherlands[6]. Our estimate is also within the range of reductions of 1 – 9% per year, reported by earlier European trend studies[12, 31].

The effect of the extension of the upper age limit of screening to age 74 between 1998 and 2001 on breast cancer mortality in women aged 75 and older was not (yet) reflected in the results of earlier research[6], because of limited follow-up after this extension. Women aged 70-74 were, on average, invited 5 years later to screening than women aged 50-69. Having long-term follow-up data now, our results support a downturn in breast cancer mortality for women aged 75-79, showing a steep decline in this age group. The size of the change in the breast cancer mortality trend, after the introduction of screening, observed in this study was larger in the oldest age group (75-79) than in younger women aged 55-74 (APC of 2.6 vs APC of 1.9), which is in line with earlier findings[10, 32]. The decline in breast cancer mortality in this group reflects a long-term accumulation of screening between ages 50 and 69 and screening between ages 70 and 74. Also, the benefit of screening, in terms of breast cancer mortality reduction, has been shown to increase with the number of screens attended[33, 34]. In addition, as older women tend to have slower tumour growth rates and less dense breasts[35-37], test sensitivity of mammography is higher at older age[38-40].

In women aged 40-54, breast cancer mortality rates started to decline moderately a few years before the introduction of screening. The most probable cause for this downturn is the increasing use of adjuvant therapy, which younger women received significantly earlier than women in the age group 55-74 in the Netherlands[22]. Before the introduction of adjuvant therapy in the 1980's for postmenopausal women

(above the age of 55 years), premenopausal node-positive women (under age 55 years) already had received adjuvant chemotherapy for ten years[17,18]. Moreover, the proportion of node-positive patients under age 55 years that received adjuvant chemotherapy increased substantially before and around the introduction of screening in the Netherlands[17]. Opportunistic screening (for high-risk women) outside the screening programme may also lower breast cancer mortality[41]. As we expected beforehand, we did not find a turning point related to screening for the age group 40-54 (Appendix) as only a small part of this group is invited to screening (i.e. 50-54) and even in this subgroup the full effect of screening may not be observable yet.

The effect of adjuvant treatment on breast cancer mortality is clearly also intertwined in the trend in women aged 55-74. A recent modelling study predicted that adjuvant treatment caused a 13.9% reduction in breast cancer mortality in the total Dutch population in 2008, compared to a setting without adjuvant treatment[16]. Simulation of biennial screening from 50-74 resulted in an additional 15.7% reduction in breast cancer mortality. Similar contributions of adjuvant treatment and screening to the observed breast cancer mortality reduction were also predicted by an earlier modelling study from the United States[42].

It is difficult to disentangle the effect of screening and treatment after the introduction of screening. Adjuvant endocrine therapy (Tamoxifen) was introduced in the 1980s in the Netherlands, before the introduction of screening, for postmenopausal patients with node-positive and estrogen receptor (ER) positive breast cancer[17]. Over the years, adjuvant treatment has improved and its usage has increased significantly[14, 17-20]. For example, from 1998 onwards, node-negative patients also received adjuvant systemic therapy, which is likely to have contributed to the decline in breast cancer mortality afterwards. Our results show that breast cancer mortality did not decrease before the introduction of screening in women invited to screening and older women, despite these improvements in therapy, which is in line with previous work[8, 32]. We did not observe differences in breast cancer mortality rates between municipalities in which screening was introduced rather early and municipalities in which screening was introduced

rather late, with the latter having more effective adjuvant treatment available in the years before and after the introduction of screening. More importantly, a strong significant decline in breast cancer mortality shortly after the introduction of screening was present in both early, intermediate and late starters, suggesting a crucial role for screening in this downturn, irrespective of the quality and the extent of treatment. The annual decline after the turning point is stronger in late starters (2.9%) than in early starters (2.2%), which may indicate a synergistic effect between early detection and improved adjuvant treatment.

In conclusion, our results show a significant decrease in breast cancer mortality shortly after the introduction of screening at the municipality level in women invited to screening and in older women, similar in early starting municipalities and late starting municipalities (with more effective adjuvant treatment available), indicating that this decrease is strongly associated with the implementation of the national screening programme.

REFERENCES

1. 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.
2. Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2013;6:CD001877.
3. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. Br J Cancer 2013;108:2205-40.
4. Nelson HD, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L. Effectiveness of Breast Cancer Screening: Systematic Review and Meta-analysis to Update the 2009 U.S. Preventive Services Task Force Recommendation. Ann Intern Med 2016;164:244-55.
5. Von Karsa 2008 von Karsa L AA, Ronco G, et al, Cancer Screening in the European Union. Report on the Implementation of the Council Recommendation on Cancer Screening First report. European Commission, 2008.
6. 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.
7. 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.
8. Ascunce EN, Moreno-Iribas C, Barcos Urtiaga A, Ardanaz E, Ederra Sanz M, Castilla J, Egues N. Changes in breast cancer mortality in Navarre (Spain) after introduction of a screening programme. J Med Screen 2007;14:14-20.
9. 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 2011;343:d4411.
10. Otto SJ, Fracheboud J, Verbeek AL, Boer R, Reijerink-Verheij JC, Otten JD, Broeders MJ, de Koning HJ, National Evaluation Team for Breast Cancer S. Mammography screening and risk of breast cancer death: a population-based case-control study. Cancer Epidemiol Biomarkers Prev 2012;21:66-73.
11. Lauby-Secretan B, Scoccianti C, Loomis D, Benbrahim-Tallaa L, Bouvard V, Bianchini F, Straif K, International Agency for Research on Cancer Handbook Working G. Breast-cancer screening--viewpoint of the IARC Working Group. N Engl J Med 2015;372:2353-8.

12. Moss SM, Nystrom L, Jonsson H, Paci E, Lynge E, Njor S, Broeders M, Euroscreen Working G. The impact of mammographic screening on breast cancer mortality in Europe: a review of trend studies. J Med Screen 2012;19 Suppl 1:26-32.
13. 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.
14. Early Breast Cancer Trialists' Collaborative G. 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.
15. Berry DA, Inoue L, Shen Y, Venier J, Cohen D, Bondy M, Theriault R, Munsell MF. Modeling the impact of treatment and screening on U.S. breast cancer mortality: a Bayesian approach. J Natl Cancer Inst Monogr 2006:30-6.
16. de Gelder R, Heijnsdijk EA, Fracheboud J, Draisma G, de Koning HJ. The effects of population-based mammography screening starting between age 40 and 50 in the presence of adjuvant systemic therapy. Int J Cancer 2014.
17. Voogd AC, van Beek MW, Crommelin MA, Kluck HM, Repelaer van Driel OJ, Coebergh JW. Management of early breast cancer in southeast Netherlands since 1984. A population-based study. Regional Breast Cancer Study Group. Acta Oncol 1994;33:753-7.
18. Nab HW HW, Crommelin MA, Kluck HM, van der Heijden LH, Coebergh JWW. Changes in long term prognosis for breast cancer in a Dutch cancer registry. BMJ 1994;309:83-86
19. 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.
20. 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.
21. Welch HG. Screening mammography--a long run for a short slide? N Engl J Med 2010;363:1276-8.
22. National Evaluation Team for Breast cancer screening (NETB). National evaluation of breast cancer screening in the Netherlands 1990-2011/2012. Thirteenth evaluation report. Erasmus MC University Medical Center Rotterdam, Radboud University Nijmegen Medical Centre, 2014.
23. van Breest Smallenburg V, Duijm LE, den Heeten GJ, Groenewoud JH, Jansen FH, Fracheboud J, Plaisier ML, van Doorne-Nagtegaal HJ, Broeders MJ. Two-view versus single-view mammography at subsequent screening in a region of the Dutch breast screening programme. Eur J Radiol 2012;81:2189-94.

24. Statistics Netherlands (CBS). www.cbs.nl. Accessed:April 2013.

25. International Agency for Research on Cancer. Breast Cancer Screening. IARC Handbooks of Cancer Prevention.ed., vol. 15. Lyon: IARCPress, 2014.

26. Waterhouse JAH MC, Correa P, Powell J eds. Cancer incidence in five continents. Lyon: IARC 1976; 3:456;1976; 3:456.

27. 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.

28. Miller AB, Wall C, Baines CJ, Sun P, To T, Narod SA. Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. BMJ 2014;348:g366.

29. Duffy SW, Tabar L, Olsen AH, Vitak B, Allgood PC, Chen TH, Yen AM, Smith RA. 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;17:25-30.

30. Weedon-Fekjaer H, Romundstad PR, Vatten LJ. Modern mammography screening and breast cancer mortality: population study. BMJ 2014;348:g3701.

31. Broeders M, Moss S, Nystrom L, Njor S, Jonsson H, Paap E, Massat N, Duffy S, Lynge E, Paci E, Group EW. The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. J Med Screen 2012;19 Suppl 1:14-25.

32. 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.

33. Miltenburg GA, Peeters PH, Fracheboud J, Collette HJ. Seventeen-year evaluation of breast cancer screening: the DOM project, The Netherlands. Diagnostisch Onderzoek (investigation) Mammacarcinoom. Br J Cancer 1998;78:962-5.

34. Fielder HM, Warwick J, Brook D, Gower-Thomas K, Cuzick J, Monypenny I, Duffy SW. A case-control study to estimate the impact on breast cancer death of the breast screening programme in Wales. J Med Screen 2004;11:194-8.

35. Peer PG, van Dijck JA, Hendriks JH, Holland R, Verbeek AL. Age-dependent growth rate of primary breast cancer. Cancer 1993;71:3547-51.

36. Weedon-Fekjaer H, Lindqvist BH, Vatten LJ, Aalen OO, Tretli S. Breast cancer tumor growth estimated through mammography screening data. Breast Cancer Res 2008;10:R41.

37. Checka CM, Chun JE, Schnabel FR, Lee J, Toth H. The relationship of mammographic density and age: implications for breast cancer screening. AJR Am J Roentgenol 2012;198:W292-5.

38. Rosenberg RD, Hunt WC, Williamson MR, Gilliland FD, Wiest PW, Kelsey CA, Key CR, Linver MN. Effects of age, breast density, ethnicity, and estrogen replacement therapy on screening mammographic sensitivity and cancer stage at diagnosis: review of 183,134 screening mammograms in Albuquerque, New Mexico. Radiology 1998;209:511-8.

39. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. Radiology 2002;225:165-75.

40. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, Geller BM, Abraham LA, Taplin SH, Dignan M, Cutter G, Ballard-Barbash R. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Ann Intern Med 2003;138:168-75.

41. Lynge E, Braaten T, Njor SH, Olsen AH, Kumle M, Waaseth M, Lund E. Mammography activity in Norway 1983 to 2008. Acta Oncol 2011;50:1062-7.

42. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, Mandelblatt JS, Yakovlev AY, Habbema JD, Feuer EJ, Cancer I, Surveillance Modeling Network C. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med 2005;353:1784-92.

SUPPLEMENTARY MATERIAL

1a. Age-standardised breast cancer mortality rates per 100 000 woman-years for the age group 40-54, 55-74 and 75-79 years per year relative to the introduction of screening

Year	40-54	55-74	75-79
-15	53.0	108.6	145.3
-14	46.0	107.1	131.2
-13	41.9	94.7	145.7
-12	47.5	102.5	160.8
-11	43.9	98.5	161.0
-10	47.0	102.0	146.7
-9	47.2	103.4	171.5
-8	43.8	104.9	161.5
-7	46.2	107.5	158.9
-6	44.9	99.7	143.0
-5	46.9	104.1	160.8
-4	45.8	107.5	171.8
-3	47.2	106.1	160.7
-2	47.0	100.5	158.0
-1	46.4	105.8	163.0
0	45.7	107.4	163.6
1	43.1	104.3	168.3
2	42.3	100.4	166.9
3	43.4	99.1	157.5
4	38.8	96.8	169.6
5	42.7	96.0	166.0
6	41.2	95.6	165.9
7	38.4	90.1	169.2
8	39.9	89.3	164.9
9	38.3	84.5	152.5
10	37.2	77.9	144.8
11	36.9	82.0	137.5
12	33.9	78.5	123.7
13	32.2	76.5	132.1
14	34.6	79.6	124.7
15	32.9	71.6	117.5
16	33.2	75.9	106.2
17	29.8	79.8	114.6
18	34.2	75.3	113.6
19	29.2	72.7	116.5
20	32.5	74.6	107.7

1b. Age-standardised breast cancer mortality rates per 100 000 woman-years for early, intermediate and late starters (age 55-74 years) per year relative to the introduction of screening

	1987-1992	1993-1994	1995-1997
-17			110.3
-16			103.3
-15			108.6
-14		111.4	103.1
-13		95.0	94.0
-12	93.1	106.6	102.4
-11	97.6	95.5	104.2
-10	100.3	99.0	109.6
-9	96.2	107.7	110.5
-8	109.5	103.0	98.4
-7	109.8	105.5	106.2
-6	94.1	105.5	101.8
-5	110.8	104.4	90.4
-4	105.0	111.7	105.7
-3	107.7	108.1	100.1
-2	106.3	86.9	110.0
-1	109.5	99.4	107.9
0	110.5	105.1	104.6
1	107.3	102.2	101.4
2	98.3	104.4	98.3
3	99.4	95.1	104.1
4	101.2	90.4	99.4
5	94.2	100.3	92.9
6	99.3	93.2	92.6
7	92.0	85.5	93.9
8	94.2	87.9	82.9
9	87.5	82.6	82.2
10	82.7	72.2	78.0
11	82.7	82.6	80.0
12	80.9	79.8	71.9
13	78.2	75.3	74.8
14	87.7	73.4	73.2
15	69.5	71.7	78.3
16	79.7	71.2	
17	81.4	75.5	
18	75.5	42.0	
19	72.6	82.5	
20	74.6		
21	84.1		
22	49.0		
23	107.5		

ADDITIONAL ANALYSIS

Methods

We used log-linear Poisson regression (on the number of breast cancer deaths and woman-years) to assess the best estimate for a knot (referred to as 'turning point') in the trend in breast cancer mortality after the introduction of screening (year 0), using the municipality-specific calendar year of introduction of screening. Two connecting linear splines were estimated for all possible positive integer values for the turning point (year 0, 1, 2, etc.) for each municipality separately, resulting in scaled deviances for all possible values for the turning point. The deviances were summed for all municipalities and the year with the smallest summed deviance was considered the best fit for the turning point. Thus, the turning point reflects the most probable year in which the mortality trend changes after the introduction of screening. Slopes before and after the turning point were estimated to calculate the annual percentage change (APC) and the corresponding 95% CI.

Models

To correct for the gradual implementation of screening, we transformed calendar years to years relative to the municipality-specific year of introduction of screening (range -t to t, with the municipality-specific year of introduction of screening defined as t = 0).

In order to determine the turning point (the best estimate for the lag time between introduction of screening and the turning point in the trend), the linear splines before and after the turning point and the corresponding deviances were estimated for all possible integer values of the turning point, using a log-linear Poisson regression model. The turning points (k) were restricted to occur after the introduction of screening and assumed to be integers, thereby ranging from 0 (introduction screening, t = 0) to n (the number of municipality-specific follow-up years since the introduction of screening until 2010, which depends on the calendar year of introduction of screening). The deviances of all municipalities were aggregated per possible turning point, resulting in a summed deviance for each turning point. The year with the smallest summed deviance was considered the best fit for the turning point (k^*).

For municipalities that remained unchanged during our study period and for municipalities without differences in calendar year of introduction of screening that were merged we used the following model (a 'Municipality-specific Model') to determine the turning point in the trend:

$$y_{ik} = \exp(\alpha_i + \beta_1 J_1 + \beta_2 J_2 + \log(PY_{it})) \qquad Y \sim \text{Poisson}$$

$$J_1 = \begin{cases} t-k & if\ t < k \\ 0 & otherwise \end{cases}$$

$$J_2 = \begin{cases} t-k & if\ t \geq k \\ 0 & otherwise \end{cases}$$

where:

y_{ik}	number of deaths in age class i and year k relative to start screening
$y_{\hat{i}k}$	fitted value for the number of deaths in age class i and year k relative to start screening
α_i	log of the mortality rate in age class i in start year of screening
β_1	slope parameter before the turning point
β_2	slope parameter after the turning point
PY_{it}	number of person-years in age class i and year j
$log\ (PY_{it})$	offset parameter: log of the number of person-years in age class i and year j
J_1;J_2	linear spline relative to year since start screening per municipality
t	time in years since start screening per municipality
k	location of possible turning point relative to t (year since start screening)

Using the correct calendar year of introduction of screening is complicated if municipalities with a different calendar year of introduction year of screening merge. Because we only had information on address at time of death it is not possible to link a breast cancer death that occurs after the merger of municipalities to the municipality that the woman who died from breast cancer lived in at the time screening was introduced and, thus, to the corresponding calendar year of introduction of screening. Therefore, the model described above was not applicable to municipalities with different calendar years of introduction of screening that were merged at some point between 1980 and 2010, as after the merger only the combined data can be observed. For these municipalities we therefore used a more complex model ('Municipality-specific Mixture Model'), which takes into account all municipality-specific years of introduction of screening of former municipalities.

$$\varphi_{ij1} = \exp(\alpha_i + \beta_1 J_{11} + \beta_2 J_{21} + \log(PY_{ijm}))$$

$$\varphi_{ij2} = \exp(\alpha_i + \beta_1 J_{12} + \beta_2 J_{22} + \beta_4 + \log(PY_{ijm}))$$

$$y_{\hat{i}\hat{j}m} = \mu_i \varphi_{ij1} + (1 - \mu_i)\varphi_{ij2}$$

$$\mu_i = \begin{cases} 1 & \text{if } m=1 \\ 0 & \text{if } m=2 \\ \frac{PY_{i\lambda 1}}{PY_{i\lambda 1} + PY_{i\lambda 2}} & \text{if } m=1+2 \end{cases}$$

where:

φ_{ijm}	estimated numbers in the original municipality m (age class i and year j). This refers to both before and after the municipal merger
$J_{11};J_{21};J_{12};J_{22}$	linear splines relative to year since start screening per municipality
PY_{ijm}	number of person-years in the original municipality m in age class i and year j
$log\ (PY_{ijm})$	offset parameter: log of the number of person-years in the original municipality m in age class i and year j
y_{ijm}	fitted value for the number of deaths in original municipality m in age class i and year j
μ_i	fraction of the population that belongs to the (former) municipality m in ageclass i
λ	last year before the merger

This model uses the assumption that, after the merger of municipalities, woman-years at risk after the merger are divided proportionally between the former municipalities based on the age specific woman-years at risk in the last year before the merger (μ_i). After the merger only the combined data of the merged municipalities can be observed and the combined data are therefore the sum of the expected numbers from the original municipalities. Slope parameters are assumed to be equal in the former municipalities before the merger but differences in breast cancer mortality level (relative to the year of introduction of screening) were modelled using a single parameter for each municipality, assuming that hazard ratios are equal for all age classes (β_4).

After the turning point was determined, slopes before and after this point were estimated using an aggregated Poisson regression model:

$$y_{ik} = \exp(\alpha_i + \beta_1 J_1 + \beta_2 J_2 + \log(PY_{it}) \qquad Y \sim \text{Poisson}$$

$$J_1 = \begin{cases} t - k^* & if\ t < k^* \\ 0 & otherwise \end{cases}$$

$$J_2 = \begin{cases} t - k^* & if\ t \geq k^* \\ 0 & otherwise \end{cases}$$

where:

y_{ik}	number of deaths in age class *i* and year *k* relative to start screening
y_{ik}	fitted value for the number of deaths in age class *i* and year *k* relative to start screening
α_i	log of the mortality rate in age class *i* in start year of screening
β_1	slope parameter before the turning point
β_2	slope parameter after the turning point
PY_{it}	number of person-years in age class *i* and year *j*
$log\ (PY_{it})$	offset parameter: log of the number of person-years in age class *i* and year *j*
$J_1;J_2$	linear spline relative to year since start screening per municipality
t	time in years since start screening per municipality
k^*	location of turning point (best fit) relative to *t* (year since start screening)

It was not feasible to take into account differences in municipality-specific year of introduction of screening of merged municipalities in this analysis. We therefore considered the year of introduction of screening of the largest former municipality as the year of introduction of screening for the whole municipal merger and thus neglected the introduction years of the other smaller former municipalities. We could therefore aggregate the number of breast cancer deaths and person-years at risk for all municipalities by age class (μ) and years relative to the introduction of screening (-t to +t). Thus, in this model, mortality levels (relative to the year of introduction of screening) and slopes before and after the turning point are

assumed to be equal in municipalities with different calendar years of introduction of screening that merge. The model uses age class as intercept and a linear spline before and after the turning point (relative to the year since introduction of screening (-t to +t) and generates the overall estimate of the slope before and after the turning point.

The more complicated regression model (Mixture Model) failed to fit for a few small municipalities that emerged from a complicated merger and these municipalities were therefore removed from the dataset.

Results

For women invited to screening (55-74), the turning point in the breast cancer mortality trend was estimated in year 2 after the introduction of screening (Table 2). The trend changed from a non-significant trend (APC -0.19, 95% CI: -0.43 – 0.05) before year 2 to a statistically significant annual decline of 1.93% (95% CI: 2.14 – 1.72) after year 2. The turning point was estimated in year 6 after the introduction of screening for women aged 75-79 (women aged 70-74 were invited to screening 5 years later on average than women aged 50-69). There was no significant trend in breast cancer mortality before year 6 (APC 0.23%, 95% CI: -0.12 – 0.58); hereafter an annual decrease of 2.56% (95% CI: 3.09 – 2.04) was estimated. For women aged 40-54, the turning point in the trend after the introduction of screening was estimated in year 10, with a statistically significant 1.08% (95% CI: 1.30 – 0.87) annual decline before year 10 and a stronger annual decrease of 1.97% (95% CI: 2.55 – 1.39) after year 10.

The turning point in the trend was estimated in year 2 after the introduction of screening for early starters (1987-1992) and in year 4 for intermediate (1993-1994) and late starters (1995-1997). Before the turning point there was no significant trend in breast cancer mortality estimated in these groups (APC 0.18, 95% CI: -0.27 – 0.62; -0.15, 95% CI: -0.56 – 0.26; 0.01, 95% CI: -0.44 – 0.45 respectively). After the turning point breast cancer mortality rates declined significantly with 2.15% (95% CI: 2.46 – 1.85), 2.55% (95% CI: 2.98 – 2.12) and 2.93% (95% CI: 3.55 – 2.30) per year, respectively (Table 2).

2. Estimated turning point after municipality-specific introduction of screening and annual percentage change (APC) before and after the turning point for different age groups and for early, intermediate and late starters (age group 55-74)

	Turning point[a] (CI)	APC before turning point (95% CI)	APC after turning point (95% CI)	p[b]
Age group				
55-74	2 (1-3)	-0.19 (-0.43 to 0.05)	-1.93 (-2.14 to -1.72)	<0.001
75-79	6[c] (4-7)	0.23 (-0.12 to 0.58)	-2.56 (-3.09 to -2.04)	<0.001
40-54	10 (10-10)	-1.08 (-1.30 to -0.87)	-1.97 (-2.55 to -1.39)	<0.001
Introduction year				
1987-1992 (early)	2 (1-2)	0.18 (-0.27 to 0.62)	-2.15 (-2.46 to -1.85)	<0.001
1993-1994 (intermediate)	4 (1-4)	-0.15 (-0.56 to 0.26)	-2.55 (-2.98 to -2.12)	<0.001
1995-1997 (late)	4 (1-7)	0.01 (-0.44 to 0.45)	-2.93 (-3.55 to -2.30)	<0.001

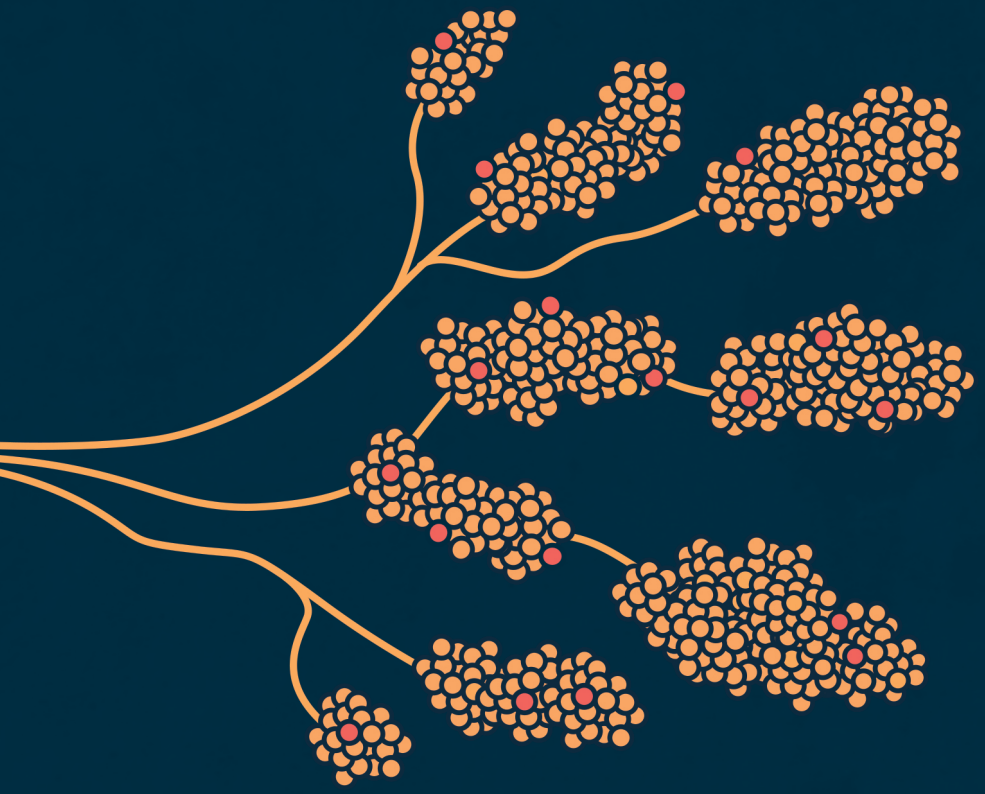
[a] In year after introduction of screening. [b] p value for difference in slope before and after estimated turning point. [c] Women aged 70-74 were invited to screening 5 years later on average than women aged 50-69.

Discussion

We estimated an annual decline in breast cancer mortality of 1.9% from year 2 after the introduction of screening onwards in women aged 55-74, which corresponds with the previously reported 1.7% annual decline, after the introduction of screening.[6] Our estimate is also within the range of reductions of 1 – 9% per year, reported by earlier European trend studies[12, 30].

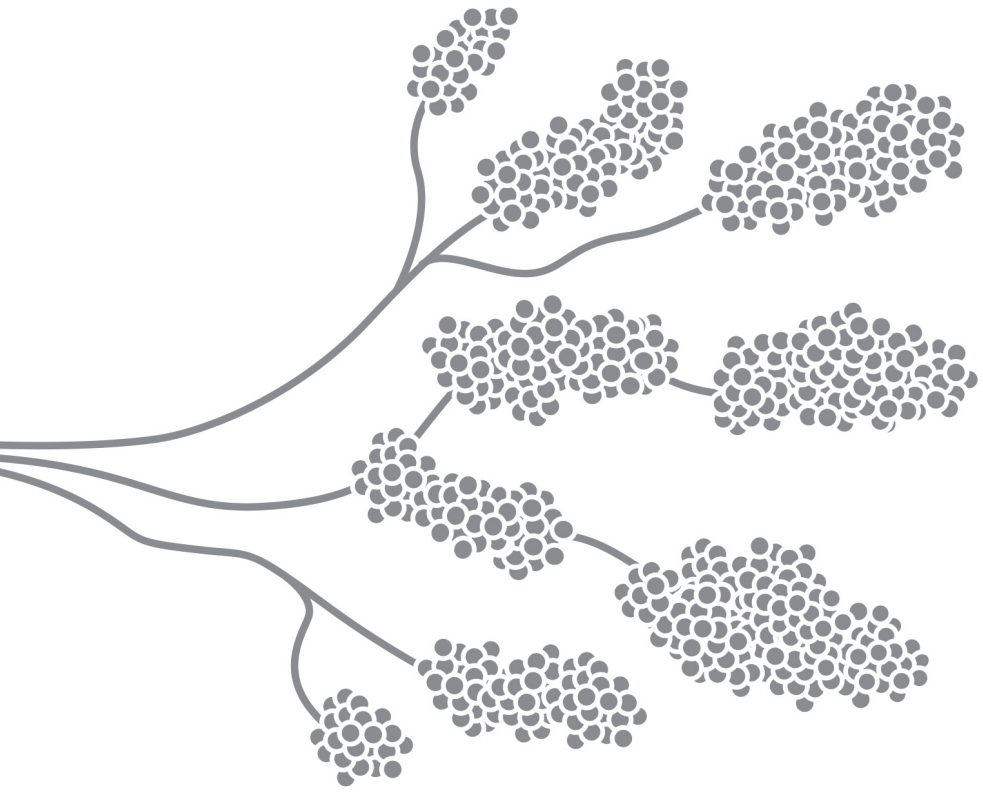
The turning point for women aged 75-79 was estimated in year 6 after the introduction of screening (for women aged 50-69), which is later than for women invited to screening (year 2) and which can be explained by the fact that women aged 70-74 years were not invited to screening until the extension of the upper-age limit of the programme between 1998 and 2001. However, a turning point in year 6 is still quite early keeping in mind that in a significant part of the municipalities screening was already introduced in the early 90s and the upper age limit was not extended until 1998. This rather early downturn in breast cancer mortality in women aged 75 and older may partly be attributable to screening participation of these women between ages 50 and 69.

We did not expect to find a turning point related to screening for the age group 40-54 as only a small part of this group is invited to screening (i.e. 50-54) and even in this subgroup the full effect of screening may not be observable yet. Indeed, the turning point in the trend in breast cancer mortality for this age group in year 10 is unlikely to be related to the introduction of screening. Of course, screening is likely to have contributed to some extent to the decline in breast cancer mortality after the introduction of screening in this age group, however it is plausible that this contribution does not translate into a turning point that is strongly associated with the introduction of screening. Nevertheless, we chose to assess the turning point in the trend after the introduction of screening for the age group 40-54 to make sure that no turning point similar to the one for the age group 55-74 (in year 2 after the introduction of screening) was found. If this would be the case, it would be plausible that a factor other than screening had a significant role in the turning of the trend.



PART II

Quantifying the cost-effectiveness, harms and benefits of different screening strategies and screening modalities in the Netherlands using microsimulation modelling



Chapter 4

Cost-effectiveness of digital mammography screening before the age of 50 in the Netherlands

Valérie D.V. Sankatsing, Eveline A.M. Heijnsdijk, Paula A. van Luijt, Nicolien T. van Ravesteyn, Jacques Fracheboud, Harry J. de Koning

International Journal of Cancer 2015;137(8):1990-9.

ABSTRACT

Background: In the Netherlands, routine mammography screening starts at age 50. This starting age may have to be reconsidered because of the increasing breast cancer incidence among women aged 40 to 49 and the recent implementation of digital mammography.

Methods: We assessed the cost-effectiveness of digital mammography screening that starts between age 40 and 49, using a microsimulation model. Women were screened before age 50, in addition to the current programme (biennial 50-74). Screening strategies varied in starting age (between 40 and 50) and frequency (annual or biennial). The numbers of breast cancers diagnosed, life-years gained (LYG) and breast cancer deaths averted were predicted and incremental cost-effectiveness ratios (ICERs) were calculated to compare screening scenarios.

Results: Biennial screening from age 50 to 74 (current strategy) was estimated to gain 157 life years per 1,000 women with lifelong follow-up, compared to a situation without screening, and cost €3,376/LYG (3.5% discounted). Additional screening increased the number of LYG, compared to no screening, ranging from 168 to 242. The costs to generate one additional LYG (i.e. ICER), comparing a screening strategy to the less intensive alternative, were estimated at €5,329 (biennial 48-74 vs. current strategy), €7,628 (biennial 45-74 vs. biennial 48-74), €10,826 (biennial 40-74 vs. biennial 45-74) and €18,759 (annual 40-49 + biennial 50-74 vs. biennial 40-74). Other strategies (49 + biennial 50-74 and annual 45-49 + biennial 50-74) resulted in less favourable ICERs.

Conclusions: These findings show that extending the Dutch screening programme by screening between age 40 and 49 is cost-effective, particularly for biennial strategies.

INTRODUCTION

Breast cancer is the most commonly diagnosed form of cancer among women aged 30 and older in the Netherlands [1]. Mammography screening allows for early detection and early treatment of breast cancer, with the aim of averting breast cancer death. In the Netherlands, women aged 50 to 74 are invited biennially to screening. Various randomised controlled trials demonstrated a statistically significant breast cancer mortality reduction due to mammography screening in this age group [2-4]. Furthermore, Otto et al. [5, 6] showed that the Dutch population-based screening programme is effective in reducing breast cancer mortality. The evidence for the benefit of mammography screening for younger women is less conclusive[7-9], however an effect for this age group is supported by several studies. Although the UK Age Trial showed a non-significant 17% reduction in breast cancer mortality[9], a statistically significant breast cancer mortality reduction of 15 to 18% associated with screening for women aged 39 to 49 or 40 to 49 at entry was demonstrated by several meta-analyses of randomised controlled trials[10-13]. In addition, a recent Swedish observational study that compared breast cancer mortality rates between women aged 40 to 49 who were invited to screening and women who were not invited to screening demonstrated a 26% statistically significant breast cancer mortality reduction[14]. Younger women may benefit less from mammography screening because of factors associated with younger age, including a lower breast cancer incidence [1] and a lower test sensitivity of mammography due to higher breast density and, possibly, faster growing tumours [15, 16]. However, results from the DMIST trial show that digital mammography improves test sensitivity for younger women with dense breasts compared to film mammography[17].

Due to the controversy over screening before age 50 there is no consensus on whether or not to offer screening to women aged 40 to 49. The American Cancer Society recommends screening starting at age 40 [18], whereas the U.S. Preventive Services Task Force stated that “the decision to start regular, biennial screening mammography before the age of 50 years should be an individual one and take patient context into account, including the patient’s values regarding specific benefits and harms” [19]. There is also no wide agreement in Europe, although the

majority of European national screening programmes do not invite women under age 50 years. Routine mammography screening is currently extended to age 47 in the UK, as a result of the outcomes of the UK Age Trial[20].

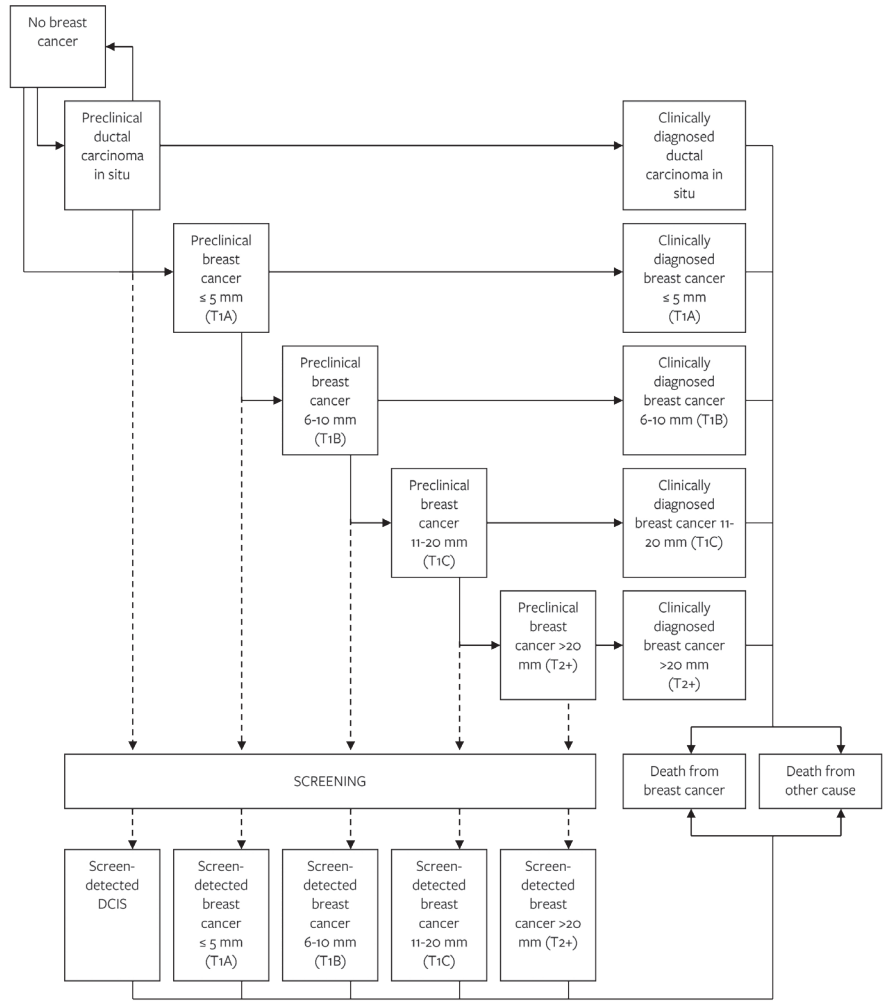
Biennial mammography screening for women aged 50 to 74 in the Netherlands has been shown to reduce breast cancer mortality at reasonable cost[21]. Two important reasons to consider a lower starting age of routine screening in the Netherlands are the increasing incidence of breast cancer among women aged 40 to 49 (2.4% annually between 1995 and 2004) [22] and the recent implementation of digital mammography. It is unclear, however, whether extending screening to women younger than 50 years is cost-effective. This study therefore assesses the cost-effectiveness of digital mammography screening between ages 40 and 49 in addition to current screening in the Netherlands.

METHODS

Model overview

The effects of screening were assessed using the MISCAN microsimulation model[21, 23]. MISCAN simulates individual life histories of women and the natural history of breast cancer in a subset of these women. First, breast cancer incidence and breast cancer mortality are estimated in a situation without screening. Subsequently, mammography screening and treatment related improvements in survival are simulated, in order to determine the impact of screening and treatment on the life histories. Breast cancer starts with the onset of a preclinical ductal carcinoma in situ (DCIS) and continues with its progression through the invasive successive stages T1A, T1B, T1C and T2+. At each stage, a tumour may become screen-detected (if screening is present), clinically detected (if symptoms are present) or may progress to the next preclinical stage (Figure 1)[24]. Screening leads to the detection of smaller tumours, which may improve survival after diagnosis. Women with a screen- or clinically detected cancer may receive adjuvant treatment, which also improves survival.

Figure 1. Transitions in the MISCAN model



Adopted from de Gelder et al., 2009[24]. The arrows represent the possible transitions.

Model parameters and assumptions

The MISCAN model was updated earlier by de Gelder et al., using Dutch screening and treatment data from 1975 to 2008 and international data[23]. We updated this model with regard to test sensitivity of mammography and background breast cancer incidence (described below). Other parameters were adopted from the

earlier model, including the mean duration of preclinical screen-detectable cancer, transition probabilities between tumour stages and survival rates after clinical diagnosis and screen-detection (Supplementary material 1).

All parameters were specified by age (including ages 40-49) and tumour stage and survival rates were also specified by lymph node status. Dutch data from screening organisations and comprehensive cancer centres were used to estimate the mean duration of preclinical screen-detectable cancer and the transition probabilities between tumour stages[1, 25]. Survival rates after screen-detection were estimated using data from the Swedish randomised controlled trials[3, 4, 26, 27]. Probabilities of receiving adjuvant treatment (endocrine therapy, chemotherapy or a combination of the two) and survival rates after receiving adjuvant treatment were incorporated using data from Dutch regional comprehensive cancer centers[1, 25] (data by age, stage and calendar year) and data from the EBCTCG meta-analysis[28] respectively.

Test sensitivity of mammography was estimated earlier using Dutch screening data on rates of screen-detected and interval cancers between 1990 and 2007 and between 1990 and 2005 respectively [23]. As digital mammography completely replaced film mammography in 2010 in the Netherlands and women in our analysis are screened from 2014 onwards, we refitted the model for age- and stage specific test sensitivity of digital mammography. The model was recalibrated using Dutch digital screening data including age-specific detection- and interval cancer rates [29], stage distribution and stage-specific detection rates [30] and breast cancer incidence between 2007 and 2011[1]. We also incorporated future trends in breast cancer incidence. The trend in background breast cancer incidence (in the absence of screening) between 1975 and 2008 was modelled previously with MISCAN by assuming an annual rise in breast cancer incidence of 1.4%[23]. Because extrapolating this annual rise to the period of time of our analysis would lead to an unlikely high incidence, we assumed that the rise diminished from 2008 to a constant incidence from the year 2028 onwards (Supplementary material 1).

The effect of screening might be smaller for women aged 40 to 49 due to, amongst others, higher breast density[15, 16]. Test sensitivity of mammography and positive

predictive value (PPV) of a screening mammogram are therefore likely to be lower for younger women. Recent Dutch data on screening characteristics of women aged 40 to 49 are unavailable. Based on different trials, we estimated test sensitivity for women aged 40 to 49 to be up to 25% lower than test sensitivity for women aged 55 and older [31]. In order to account for a gradual change, test sensitivity was linearly interpolated between age 50 and 55. Stage-specific test sensitivities are shown in the Supplementary material 1. The PPV of a digital screening mammogram was estimated to be 30% for women aged 50 and older, based on findings of a recent Dutch study [30]. The PPV for women aged 40 to 49 was assumed to be 12% (40% of the PPV for women aged 50 and older) [17, 32, 33].

In the Netherlands, mammography is often repeated in the hospital after recall for further assessment because of a suspicious screening mammogram[34]. Therefore, we assumed that all women recalled for further assessment would undergo an additional (diagnostic) mammogram.

Screening strategies

A cohort of 10 million Dutch women was simulated and women were followed from age 40 to death (from 2014 to 2074). Screening strategies varied in starting age (between 40 and 50) and frequency (annual or biennial). Women were screened biennially between age 50 and 74 in all strategies, in accordance with the current programme. Additional effects of strategies could therefore directly be linked to screening before age 50. Women aged 40 to 49 have been shown to have higher interval cancer rates than older women, due to greater mammographic breast density and higher tumour growth rates[16, 35]. Therefore, we considered both biennial and annual screening for women aged 40 to 49. The following strategies were simulated: 1) biennial screening from age 50 to 74 (current strategy), 2) one screen at age 49 and biennial screening from age 50 to 74, 3) biennial screening from age 48 to 74, 4) biennial screening from age 45 to 74, 5) annual screening from age 45 to 49 and biennial screening from age 50 to 74, 6) biennial screening from age 40 to 74, 7) annual screening from age 40 to 49 and biennial screening from age 50 to 74. In addition, to compare the cost-effectiveness of lowering the starting age of screening to the cost-effectiveness of increasing the stopping age, biennial

screening from age 50 to 76 was simulated. An 80% attendance to screening was assumed [29].

For each strategy, the number of invitations, mammograms, screen- and clinically detected tumours (stage- and age specific), total life-years and breast cancer deaths were predicted. Results are presented per 1,000 women, aged 40 in 2014, with lifelong follow-up. The number of false positive findings was calculated using the number of screen-detected cancers and a PPV of 30% (women aged 50 and older) or a PPV of 12% (women aged 40 to 49).

Costs and effects

We adopted a health care payer perspective[36] and calculated direct medical costs including costs of screening, diagnostics and treatment. An overview of all costs is shown in Table 1[29, 37-39].

Both the effects of screening and adjuvant treatment are simulated with MISCAN. In order to predict the effect of screening, breast cancer mortality is estimated in a scenario with adjuvant treatment and screening and compared to breast cancer mortality in a scenario with adjuvant therapy but without screening. The effect of screening was estimated by predicting life-years gained (LYG). Costs and effects were calculated from the lowest starting age of screening (40 years) until death. Both effects and costs were discounted at 3.5% per year to take time preference into account [40]. In order to meet Dutch standards, we also used a discount rate of 1.5% for effects and 4% for costs (Supplementary material 2).

Screening strategies were ranked according to their effectiveness (number of LYG). The cost-effectiveness ratio was calculated for the current screening strategy (biennial 50-74), the least effective scenario, as the difference in costs between current screening and no screening divided by the number of LYG by current screening. Subsequently, in order to compare screening scenarios, incremental cost-effectiveness ratios (ICERs) were calculated for strategies that screen additionally between ages 40 and 49 as the difference in costs divided by the difference in LYG between a strategy and the previous, less effective, strategy in

the ranking. The ICER of a strategy therefore reflects the costs required to generate one additional LYG, compared to the previous strategy. Strategies were defined as dominated if an alternative, more effective strategy existed that required lower costs to generate an additional LYG. Non-dominated strategies were considered to be efficient. We compared ICERs to a cost-effectiveness threshold of £20,000 - £30,000 (approximately €24,000 - €36,000) per quality-adjusted life-year (QALY) gained [40]. Strategies that did not exceed this threshold were considered to be cost-effective.

Sensitivity analyses

Univariate sensitivity analyses were performed in order to assess to what extent parameter values and assumptions affected the costs per LYG and the ranking of efficient screening strategies. First, we varied the test sensitivity of digital mammography for women aged 40 to 49 (50% and 100% of the test sensitivity for women aged 55 and older). Second, we varied the PPV of digital mammography for women aged 40 to 49 (9% and 15%). Third, we simulated a constant background breast cancer incidence (equal to incidence in 2008) as well as an annual increase in incidence of 1%. Fourth, in our analysis, all women who are screened and recalled for further assessment were assumed to undergo an additional mammogram (in addition to the screening mammogram) in the hospital. This may however not always be true in practice and we therefore also tested the assumption that instead of all women, only 50% of the women would undergo an additional mammogram in the hospital.

Finally, in order to assess the effect of screening strategies on quality of life we estimated quality-adjusted life-years (QALYs), using utility estimates with a value between 0 (worst imaginable health state) and 1 (healthy state). We included reductions in utility associated with screening participation and a positive screen of 0.006 for 1 week and 0.105 for 5 weeks respectively[41]. To take into account reductions in utility from breast cancer treatment, we used adjusted health utilities reported by Stout et al.[42], with minor adjustments in the application of these utilities as our model cannot discriminate between regional and distant breast cancer stages (Supplementary material 1). Women who do not die of breast

cancer were assumed to experience a loss in quality of life of 0.1 for 2 years from diagnosis if diagnosed with in situ or localised breast cancer and a loss of 0.25 for 2 years from diagnosis if diagnosed with regional or distant breast cancer. For women who die of breast cancer, a reduction of 0.4 was calculated from diagnosis until breast cancer death.

Table 1. Costs associated with breast cancer screening, diagnosis and treatment[29,37,38]

Procedure	Costs[a]
Screening	
Invitation	2
Digital screening mammogram	58
Diagnosis	
Magnetic Resonance Imaging (MRI)[b]	367.59
Consultation after recall[c]	69.05
Ultrasound	77.08
Fine Needle Aspiration (FNA)	141.73
Biopsy[d]	175.86
Additional mammogram[e]	103.23
Treatment by stage[f]	
DCIS	4,569
T1a, N-	4,333
T1b, N-	5,057
T1c, N-	11,146
T2, N-	10,815
T1a, N+	16,103
T1b, N+	6,744
T1c, N+	20,822
T2, N+	15,063
Advanced disease	
Palliative therapy	18,000[h]

[a]Costs are in Euros.
[b]MRI is assumed to be performed to measure the effect of neo-adjuvant therapy for all T2+ cancers.
[c]The number of consultations after recall in the presence of screening is calculated by using the number of screen-detected cancers and a positive predictive value (PPV) of 30%[30] (ages 50-74) or 12% (ages 40-49). The number of consultations after recall in the absence of screening is calculated by using the number of clinically detected cancers and a PPV of 58.3%[39].
[d]All biopsies are assumed to be image-guided and therefore coincide with an ultrasound. The number of biopsies is calculated by using the number of breast cancers diagnosed and a PPV of 66.7%[29]. Costs of biopsy are calculated as the mean costs of FNA and biopsy.
[e]All women recalled for further assessment are assumed to undergo an additional mammogram.
[f]Mean treatment costs per tumour stage.
[h]Estimate from de Koning et al. 1992[38]; indexed to current price levels.

RESULTS

Model validation

As women in our analysis are screened and followed in the future (2014-2074), we were not able to compare model predictions to observed data. However, our model predicted trends in breast cancer incidence and mortality from 1989 to 2011 quite well (Supplementary material 3, 4).

Effects of screening

Without screening, 135 cases of breast cancer and 45 breast cancer deaths were predicted to occur among 1,000 women, aged 40 years, followed over their lifetimes (undiscounted) (Table 2).

Screening these women biennially from age 50 to 74 (current strategy) averted 12 breast cancer deaths and gained 157 life-years (13 LYG per breast cancer death averted), compared to no screening. In total, 138 breast cancer cases were diagnosed, of which 53 were screen-detected and 3 were overdiagnosed (138 - 135). Screening led to 124 false positive findings.

Strategies that screened additionally before age 50 gained life-years ranging from 168 (49 + biennial 50-74) to 242 (annual 40-49 + biennial 50-74), compared to no screening. Offering annual screening from age 40 to 49, in addition to the current

Table 2. Effects and costs[a] of screening

Undiscounted	No screening	B 50-74 (current)	49 + B 50-74	48 + B 50-74	B 45-74	A 45-49 + B 50-74	B 40-74	A 40-49 + B50-74
Screening data								
Women invited to screening		+10,710	+11,604	+11,671	+13,129	+15,471	+15,591	+20,385
Screening tests		+8,653	+9,377	+9,431	+10,613	+12,511	+12,608	+16,495
Effects								
Breast cancers diagnosed[b]	135	+3	+3	+3	+3	+3	+3	+4
Screen-detected cancers		+53	+56	+55	+59	+62	+62	+67
Breast cancer deaths	45	-12	-12	-13	-13	-14	-14	-15
Life-years (%)[c]	41,382	+157 (0.4)	+168 (0.4)	+173 (0.4)	+191 (0.5)	+205 (0.5)	+215 (0.5)	+242 (0.6)
Quality-adjusted life-years	41,171	+198	+211	+217	+239	+256	+268	+300
False-positive findings		+124	+139	+147	+171	+189	+198	+229

Table 2. (Continued)

	No screening	B 50-74 (current)	49 + B 50-74	48 + B 50-74	B 45-74	A 45-49 + B 50-74	B 40-74	A 40-49 + B50-74
Costs (€)								
Invitations		+21,420	+23,208	+23,342	+26,258	+30,942	+31,181	+40,770
Screening mammograms		+501,848	+543,847	+547,017	+615,531	+725,639	+731,280	+956,713
Diagnosis	117,839	+5,090	+7,067	+8,493	+11,667	+14,041	+15,414	+19,458
Treatment	1,714,552	-173,202	-180,836	-181,749	-192,980	-207,184	-207,062	-230,906
Breast cancer deaths	814,882	-212,256	-222,712	-225,364	-239,243	-252,526	-255,868	-278,564
Total (%)[c]	2,647,273	+142,900 (5)	+170,574 (6)	+171,740 (6)	+221,232 (8)	+310,912 (12)	+314,945 (12)	+507,470 (19)
3.5% Discounted								
Total costs	1,161,008	+137,057	+158,680	+163,704	+210,234	+281,643	+306,590	+475,420
Life-years	20,834	+41	+44	+46	+52	+56	+61	+70

[a]Results represent the difference in effects or costs between a screening strategy and a situation without screening and are presented per 1,000 women, aged 40 years, followed over their lifetimes.
[b]The additional cancers diagnosed by screening, compared to no screening, reflect overdiagnosis.
[c]Percentage change compared to no screening.
Abbreviations: biennial screening (B); annual screening (A).

strategy, would lead to three additional deaths averted, 85 additional LYG and 105 additional false-positive findings (Table 2), by performing 7,842 additional mammograms, per 1000 women, aged 40 years, with life-long follow-up.

Biennial screening from age 50 to 76 gained 161 undiscounted life-years and averted 12 breast cancer deaths (data not shown).

Incremental cost-effectiveness

Additional screening at age 49 and additional annual screening from age 45 to 49 were dominated by biennial strategies that were more effective and required lower costs to generate an additional LYG. Biennial screening from age 50 to 76 was dominated by biennial screening from age 48 to 74, which gained more additional life-years for similar costs (data not shown). All other strategies were on the efficiency frontier (Figure 2).

Total costs due to breast cancer diagnosis, treatment and death in the absence of screening were estimated at €1,161,008 per 1,000 women, followed over their lifetimes (3.5% discounted; Table 2). The estimated costs of the current screening program were €3,376 per LYG (3.5% discounted) (Table 3). One additional screening round at age 48 was predicted to gain 5 additional life-years, per 1,000 women, and to cost €5,329 per additional LYG. Biennial screening from age 45 instead of 48 would gain 6 more life-years that cost €7,628 per LYG. Lowering the starting age of biennial screening from age 45 to age 40 would gain 9 additional life-years that would cost €10,826 each. Screening annually instead of biennially from age 40 would result in a gain of 9 more life-years that cost €18,759 per LYG. All ICERs were below the cost-effectiveness threshold.

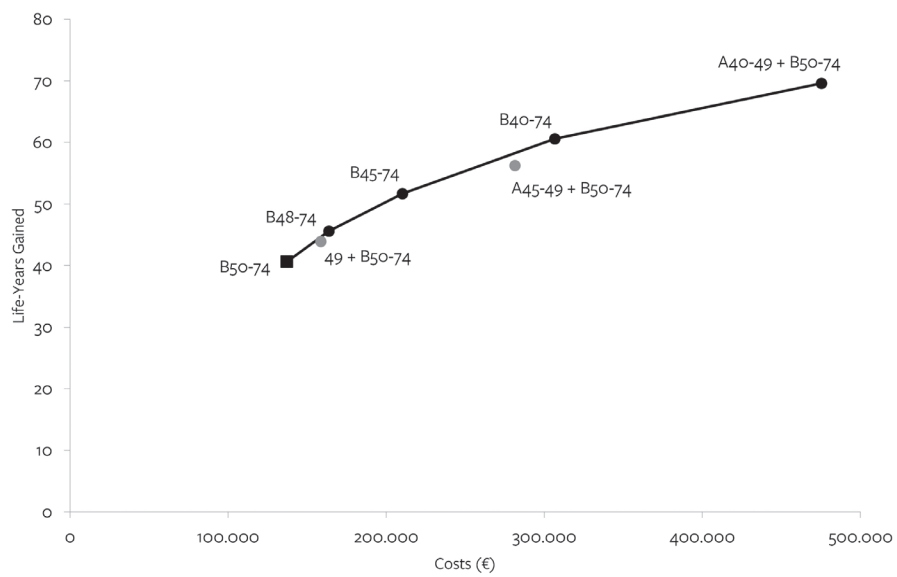
Sensitivity analyses

Estimated costs per (additional) LYG were only slightly influenced by varying the PPV for women under the age of 50 or by assuming that only 50% of women recalled for further assessment would undergo an additional mammogram (Supplementary material 5). Assuming a constant breast cancer incidence or lower test sensitivity of digital mammography for women aged 40 to 49 led to more unfavourable ratios

of costs and effects (maximum €21,459) whereas a constant annual increase in incidence or higher test sensitivity led to more favourable ratios. Adjustment for quality of life resulted in a slightly higher reduction in life-years in the absence of screening than in the presence of screening, which led to a higher number of QALYs gained than the number of life-years gained for the different screening scenarios (Table 2). ICERs calculated by dividing the change in costs by the change in QALYs gained are therefore slightly more favourable than the ICERs based on the change in life-years gained (Supplementary material 6).

The ranking of strategies was not affected by any of the sensitivity analyses performed and ICERs were below the cost-effectiveness threshold (£20,000, €24,000) in all scenarios.

Figure 2. Efficiency frontier of screening strategies



Both costs and LYG are relative to a situation without screening and are presented per 1,000 women (3.5% discounted).The current screening strategy (B 50-74) is displayed by a square, all other points reflect ICERs of additional screening strategies. Points on the frontier represent efficient strategies (i.e. no alternative strategy exists that gains more life years for fewer costs per additional LYG). Dominated strategies are represented by grey dots. Strategies consist of annual (A), biennial (B) or combined (A + B) screening.

Table 3. Incremental cost-effectiveness ratios of additional screening before age 50

Screening strategy	Additional screening rounds	Costs[a] (€)	Life-years gained[b]	ICER[b]
No screening		1,161,008		
B 50-74 (current)		137,057	41	3,376
49 + B 50-74	1	158,680	44	dominated
B 48-74	1	163,704	46	5,329
B 45-74	2.5	210,234	52	7,628
A 45-49 + B 50-74	5	281,643	56	dominated
B 40-74	5	306,590	61	10,826
A 40-49 + B 50-74	10	475,420	70	18,759

[a]Both costs and life-years gained are 3.5% discounted and are presented per 1,000 women, aged 40 years, followed over their lifetimes and are relative to a situation without screening.
[b]calculated as the difference in costs divided by the difference in life-years gained between a strategy and the previous, less effective, non-dominated strategy in the ranking (no screening for the current strategy).
Abbreviations: life-year gained (LYG); incremental cost-effectiveness ratio (ICER); biennial screening (B); annual screening (A).

DISCUSSION

This study shows that digital mammography screening between age 40 and 49 in the Netherlands, in addition to the current screening strategy, is cost-effective. Our results indicate that, from a cost-effectiveness perspective, biennial screening from age 40 to 49 is more efficient in addition to current screening than annual screening from age 45 to 49. Furthermore, our analysis suggests that one additional screening round is more efficient at age 48 than at age 49 and that extending the lower age-limit of screening by one additional screening round is more effective and will result in lower costs per additional LYG than extending the upper age-limit by one screening round.

Our model predicted that both the number of LYG and breast cancer mortality reduction due to screening would increase with decreasing starting age. However, aside from benefits, additional screening before age 50 years is also associated with additional harms. We estimated that the number of false positive findings would increase by 74 per 1,000 women (60%), if biennial screening starts at age 40 instead of age 50. A higher false positive rate may result in more unnecessary biopsies. However, false positive rates of current screening in the Netherlands are considerably lower than rates in other countries[43] and will therefore probably remain relatively low if women are screened from age 40, despite the substantial increase in absolute number of false positive findings. False positive mammograms were recently shown to increase short-term anxiety [44]. However, long-term anxiety and health utility scores did not differ between women with false positive mammograms and women with negative mammograms. The number of overdiagnosed cancers was estimated to increase by 0.33 per 1,000 women (11%) if biennial screening starts at age 40 instead of age 50 (cannot be read from Table 2 because numbers are rounded). Overdiagnosis is defined as screen-detection of a cancer that would never have presented clinically in the absence of screening, during a woman's lifetime. As overdiagnosed cancers would not have led to symptoms in the absence of screening, treatment of these cancers is considered to be harmful.

This study shows that the current screening programme in the Netherlands is highly cost-effective (€3,376 per LYG), which corresponds with findings of an earlier conducted Dutch study [21]. Our model predicted a 26% breast cancer mortality reduction due to screening from age 50 to 74, which is in line with the outcomes of randomized controlled trials [4]. Our results show that lowering the starting age of biennial screening to age 40 would reduce breast cancer mortality additionally by 5%. This finding is comparable to results of a previous modelling study in which the additional mortality reduction due to biennial screening from age 40 was determined using 6 different models (range, 1 to 6%) [45]. This previous study also showed that most strategies on the efficiency frontier are biennial strategies. Correspondingly, in our analysis 4 of the 5 efficient strategies have a biennial screening interval. The only annual strategy on the frontier, annual

screening from age 40 (+ biennial 50-74), is efficient because it is the most intensive strategy that we simulated and it therefore yields the largest effect (i.e. could not be dominated by an alternative strategy). However, the ICER increases considerably shifting from biennial to annual screening between ages 40 and 49 (from €10,826 per LYG to €18,759 per LYG), which has been reported before[46].

A micro simulation study based on data from the US showed that biennial mammography screening from age 40 to 79 is cost-effective only for women with either BIRADS breast density categories 3 or 4 or a previous breast biopsy as well as a family history of breast cancer and that annual mammography from age 40 to 49 is not cost-effective for any high-risk group (using a threshold of $100,000 per QALY gained) [47]. In contrast, we found that both annual and biennial mammography screening from age 40 are cost-effective, regardless of risk factors. However, Schousboe et al. [47] focused on film mammography, which has been shown to be less cost-effective than digital mammography for younger women [48]. Furthermore, sensitivity analyses showed that results were sensitive to the proportion of false positive findings assumed and false positive rates in the Netherlands are significantly lower than in the US [43]. This probably (partly) accounts for the more favourable ratio of costs and effects in our analysis. A more recent US modelling study showed that extending biennial digital mammography screening to all women aged 40 to 49, regardless of breast density, is cost-effective [49], which is in line with our findings. However, results were sensitive to decreases in quality of life associated with screening and false positives.

Although our results are primarily based on Dutch screening data, outcomes regarding the efficiency of screening strategies are likely to be translatable to other countries.

To our knowledge, our study is the first cost-effectiveness analysis, using Dutch population data and including digital mammography screening from age 40 to 49. One of the strengths of our study is that we used digital mammography screening data for the calibration of our model. An advantage of using a model to determine

the effectiveness of screening is that long-term effects are predicted, as women are followed over their lifetime.

This study also had a few limitations. First, our model outcomes depend on assumptions and input values. We assumed a constant PPV for women aged 40 to 49 and a higher constant PPV for women aged 50 and older. In reality there may be a gradual increase in PPV with increasing age and changes in PPV over time. However, sensitivity analyses showed that differences in PPV were of little influence on our outcomes. Differences in test sensitivity did affect model outcomes. We estimated test sensitivity for women aged 40 to 49 to be up to 25% lower, using studies that were based on the use of film mammography, [31] because digital data on age-specific test sensitivity is scarce. The ratio in sensitivity of digital mammography between younger and older women may be different due to improved sensitivity associated with digital mammography for younger but not for older women [17]. The estimated difference in sensitivity may therefore be overestimated and the number of additional LYG by screening before age 50 could be a conservative estimate. Model outcomes were also sensitive to the assumed trend in background breast cancer incidence. Extrapolating the earlier trend in incidence to the time period of our analysis would lead to an extremely high and unlikely incidence. We therefore assumed that the earlier annual rise in incidence would decrease over time. Although differences in test sensitivity and background breast cancer incidence were of influence on costs per additional LYG, the ranking of strategies and whether a strategy was efficient or dominated remained unchanged. Furthermore, in the worst-case scenarios (low test sensitivity or constant background breast cancer incidence) ICERs did not exceed the cost-effectiveness threshold (£20,000, €24,000).

Another limitation is that we compare costs per LYG to a threshold expressed as costs per QALY gained, in our base case analysis. However, our sensitivity analysis showed that adjusting for quality of life has little impact on cost-effectiveness estimates, which is in line with earlier findings[41, 50]. In addition, when we adjust for quality of life, the incremental cost-effectiveness ratios are slightly more favourable and therefore remain below the threshold of £20,000 per QALY gained

(approximately €24,000 per QALY gained). When we compare the ICERs to the willingness-to-pay threshold of €20,000 per QALY gained, often cited in Dutch cost-effectiveness analyses[51-53], all screening scenarios remain cost-effective.

Apart from age, we did not consider risk factors for breast cancer (e.g. breast density). The cost-effectiveness of risk-based screening strategies that start before age 50 is therefore an important area for further research.

Finally, the effect of adjuvant treatment was modelled using the most recent data of the EBCTCG trial[28]. Future improvements in therapy may reduce the effectiveness and therefore the cost-effectiveness of screening.

In conclusion, our results indicate that additional screening between age 40 and 49 in the Netherlands is cost-effective. However, the decision about whether or not to implement screening before age 50 years will also depend on the balance of benefits and harms. If it is decided to extend the screening programme, our findings provide information that could be useful for selecting an appropriate screening strategy, by taking into account the cost-effectiveness of different starting ages between age 40 and 49 and suggesting that biennial strategies have more favourable ratios of costs and effects.

REFERENCES

1. Nederlandse Kankerregistratie. http://www.cijfersoverkanker.nl/. (Accessed: January, 2014).
2. Shapiro S. Periodic screening for breast cancer: the HIP Randomized Controlled Trial. Health Insurance Plan. J Natl Cancer Inst Monogr 1997:27-30.
3. 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.
4. 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.
5. 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.
6. Otto SJ, Fracheboud J, Verbeek AL, Boer R, Reijerink-Verheij JC, Otten JD, Broeders MJ, de Koning HJ, National Evaluation Team for Breast Cancer S. Mammography screening and risk of breast cancer death: a population-based case-control study. Cancer Epidemiol Biomarkers Prev 2012;21:66-73.
7. Alexander FE, Anderson TJ, Brown HK, Forrest AP, Hepburn W, Kirkpatrick AE, Muir BB, Prescott RJ, Smith A. 14 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. Lancet 1999;353:1903-8.
8. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. Ann Intern Med 2002;137:305-12.
9. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L, Trial Management G. 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.
10. Hendrick RE, Smith RA, Rutledge JH, 3rd, Smart CR. Benefit of screening mammography in women aged 40-49: a new meta-analysis of randomized controlled trials. J Natl Cancer Inst Monogr 1997;22:87-92.
11. Smith RA, Duffy SW, Gabe R, Tabar L, Yen AM, Chen TH. The randomized trials of breast cancer screening: what have we learned? Radiol Clin North Am 2004;42:793-806, v.
12. 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.

13. Magnus MC, Ping M, Shen MM, Bourgeois J, Magnus JH. Effectiveness of mammography screening in reducing breast cancer mortality in women aged 39-49 years: a meta-analysis. J Womens Health (Larchmt) 2011;20:845-52.

14. 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.

15. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, Geller BM, Abraham LA, Taplin SH, Dignan M, Cutter G, Ballard-Barbash R. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Ann Intern Med 2003;138:168-75.

16. Buist DS, Porter PL, Lehman C, Taplin SH, White E. Factors contributing to mammography failure in women aged 40-49 years. J Natl Cancer Inst 2004;96:1432-40.

17. 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.

18. American Cancer Society. http://www.cancer.org/. (Accessed: February, 2014).

19. U. S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2009;151:716-26, W-236.

20. NHS Breast Cancer Screening Programme. http://www.cancerscreening.nhs.uk. (Accessed: May, 2014).

21. Groenewoud JH, Otten JD, Fracheboud J, Draisma G, van Ineveld BM, Holland R, Verbeek AL, de Koning HJ, Netb. Cost-effectiveness of different reading and referral strategies in mammography screening in the Netherlands. Breast Cancer Res Treat 2007;102:211-8.

22. 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.

23. In press: de Gelder R, Heijnsdijk EA, Fracheboud J, Draisma G, de Koning HJ. The effects of population-based mammography screening starting between age 40 and 50 in the presence of adjuvant systemic therapy. Int J Cancer 2015;137:165-172

24. 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;45:127-38.

25. 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.

26. 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.

27. 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.

28. 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.

29. National Evaluation Team for Breast cancer screening (NETB). National evaluation of breast cancer screening in the Netherlands 1990-2011/2012. Thirteenth evaluation report. Erasmus MC University Medical Center Rotterdam, Radboud University Nijmegen Medical Centre, 2014.

30. van Luijt PA, Fracheboud J, Heijnsdijk EA, den Heeten GJ, de Koning HJ, National Evaluation Team for Breast Cancer Screening in Netherlands Study G. Nation-wide data on screening performance during the transition to digital mammography: observations in 6 million screens. Eur J Cancer 2013;49:3517-25.

31. International Agency for Research on Cancer. Breast Cancer Screening, volume 7. In: Vainio H, Bianchini F. IARC Handbooks of Cancer Prevention, eds. IARC Handbooks of Cancer and Prevention. Lyon: IARCPress, 2002.

32. 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.

33. Perry NM, Patani N, Milner SE, Pinker K, Mokbel K, Allgood PC, Duffy SW. The impact of digital mammography on screening a young cohort of women for breast cancer in an urban specialist breast unit. Eur Radiol 2011;21:676-82.

34. Timmers JM, van Doorne-Nagtegaal HJ, Zonderland HM, van Tinteren H, Visser O, Verbeek AL, den Heeten GJ, Broeders MJ. The Breast Imaging Reporting and Data System (BI-RADS) in the Dutch breast cancer screening programme: its role as an assessment and stratification tool. Eur Radiol 2012;22:1717-23.

35. Bailey SL, Sigal BM, Plevritis SK. A simulation model investigating the impact of tumor volume doubling time and mammographic tumor detectability on screening outcomes in women aged 40-49 years. J Natl Cancer Inst 2010;102:1263-71.

36. Neumann PJ. Costing and perspective in published cost-effectiveness analysis. Med Care 2009;47:S28-32.

37. Saadatmand S, Tilanus-Linthorst MM, Rutgers EJ, Hoogerbrugge N, Oosterwijk JC, Tollenaar RA, Hooning M, Loo CE, Obdeijn IM, Heijnsdijk EA, de Koning HJ. Cost-effectiveness of screening women with familial risk for breast cancer with magnetic resonance imaging. J Natl Cancer Inst 2013;105:1314-21.

38. de Koning HJ, van Ineveld BM, de Haes JC, van Oortmarssen GJ, Klijn JG, van der Maas PJ. Advanced breast cancer and its prevention by screening. Br J Cancer 1992;65:950-5.

39. Huisarts en Wetenschap 2002;4:194-98.

40. National Institute for Health and Care Excellence. http://www.nice.org.uk/. (Accessed:January, 2014).

41. 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.

42. 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.

43. Otten JD, Fracheboud J, den Heeten GJ, Otto SJ, Holland R, de Koning HJ, Broeders MJ, Verbeek AL. Likelihood of early detection of breast cancer in relation to false-positive risk in life-time mammographic screening: population-based cohort study. Ann Oncol 2013;24:2501-6.

44. Tosteson AN, Fryback DG, Hammond CS, Hanna LG, Grove MR, Brown M, Wang Q, Lindfors K, Pisano ED. Consequences of false-positive screening mammograms. JAMA Intern Med 2014;174:954-61.

45. 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.

46. 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.

47. 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.

48. 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.

49. Stout NK, Lee SJ, Schechter CB, Kerlikowske K, Alagoz O, Berry D, Buist DS, Cevik M, Chisholm G, de Koning HJ, Huang H, Hubbard RA, et al. Benefits, harms, and costs for breast cancer screening after US implementation of digital mammography. J Natl Cancer Inst 2014;106:dju092.

50. Chapman RH, Berger M, Weinstein MC, Weeks JC, Goldie S, Neumann PJ. When does quality-adjusting life-years matter in cost-effectiveness analysis? Health Econ 2004;13:429-36.

51. de Kok IM, van Rosmalen J, Dillner J, Arbyn M, Sasieni P, Iftner T, van Ballegooijen M. Primary screening for human papillomavirus compared with cytology screening for cervical cancer in European settings: cost effectiveness analysis based on a Dutch microsimulation model. BMJ 2012;344:e670.

52. de Boer PT, Pouwels KB, Cox JM, Hak E, Wilschut JC, Postma MJ. Cost-effectiveness of vaccination of the elderly against herpes zoster in The Netherlands. Vaccine 2013;31:1276-83.

53. Verhoef TI, Redekop WK, Hasrat F, de Boer A, Maitland-van der Zee AH. Cost effectiveness of new oral anticoagulants for stroke prevention in patients with atrial fibrillation in two different European healthcare settings. Am J Cardiovasc Drugs 2014;14:451-62.

SUPPLEMENTARY MATERIAL

1. Model parameters on natural history of breast cancer, and survival after adjuvant treatment and screening

Annual increase in background breast cancer incidence (without screening)[a]	%
2008-2012	1.4
2013-2017	1.0
2018-2022	0.6
2023-2027	0.2
From 2028	0

Mean duration (years) of screen-detectable preclinical stage by age and stage[1, 2 b]	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

Test sensitivity of digital mammography by age and preclinical stage[1-3 b]	Stage								
Age	DCIS	T1AN–	T1AN+	T1BN–	T1BN+	T1CN–	T1CN+	T2+N–	T2+N+
<50	100%	70%	70%	83%	83%	80%	80%	93%	93%
≥55	100%	94%	94%	100%	100%	100%	100%	100%	100%

Long-term relative survival by clinical stage and age, without adjuvant treatment[4-9 c]	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[2, 10, 11 b,c]	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[2, 10, 11 b,c]	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[2, 10, 11 b,c]

Age	Stage								
	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[6, 12-14 c]

Age	Stage								
	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%

Quality-of-life effects by stage[15]

Stage	Quality-of-life adjustments
In situ or localised	0.9 for 2 years
Regional or distant	0.75 for 2 years
Breast cancer death	0.6 until death

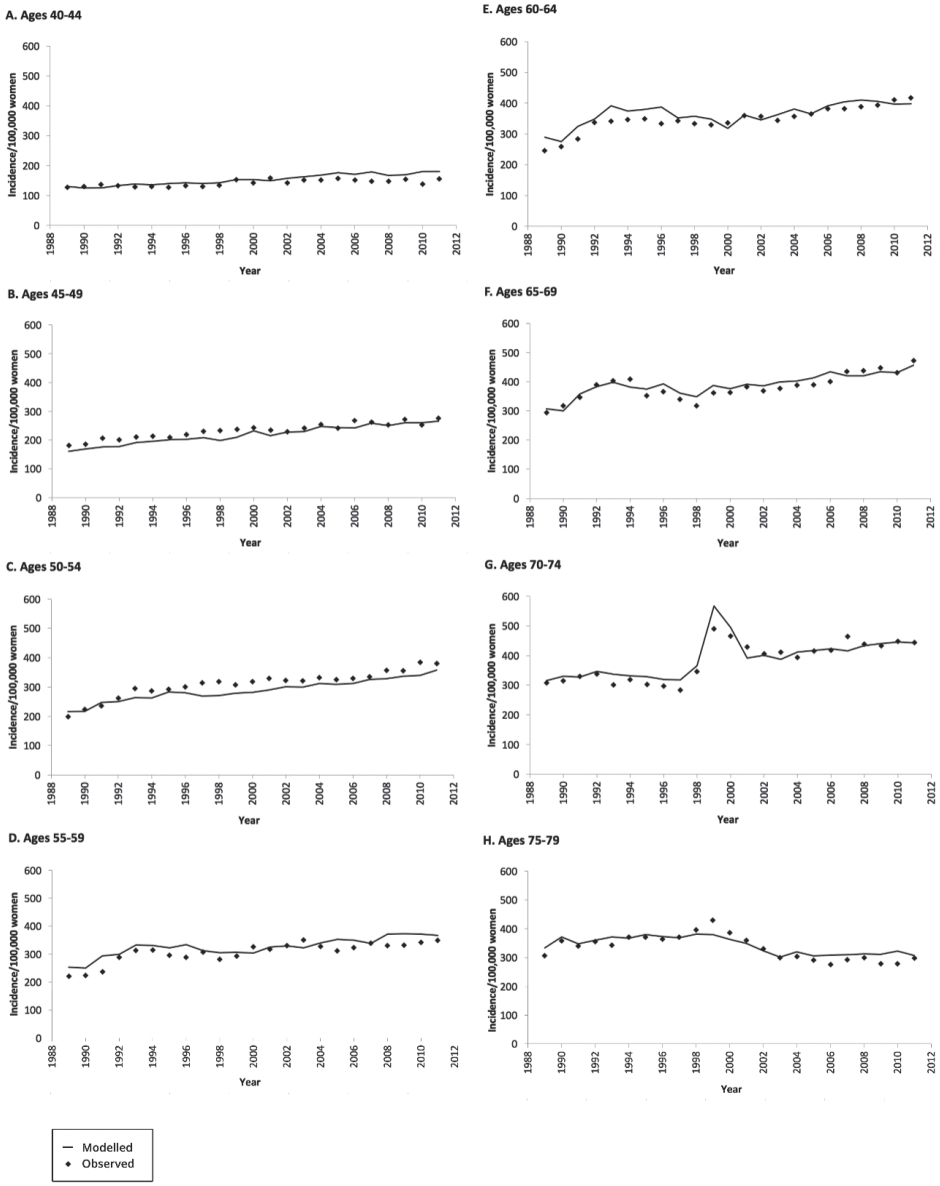
Adapted from de Gelder et al., 2014[16]. Abbreviations: ductal carcinoma *in situ* (DCIS); lymph node negative breast cancer with diameter of 5 mm or smaller (T1AN-); lymph node negative breast cancer with diameter of 6-10 mm (T1BN-); lymph node negative breast cancer with diameter of 11-20 mm (T1CN-); 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+).
[a] based on assumptions, [b] data used for model fit (calibration), [c] data used directly as input

2. Costs, effects and incremental cost-effectiveness ratios, discounted in accordance with Dutch standards (effects 1.5% discounted, costs 4% discounted)

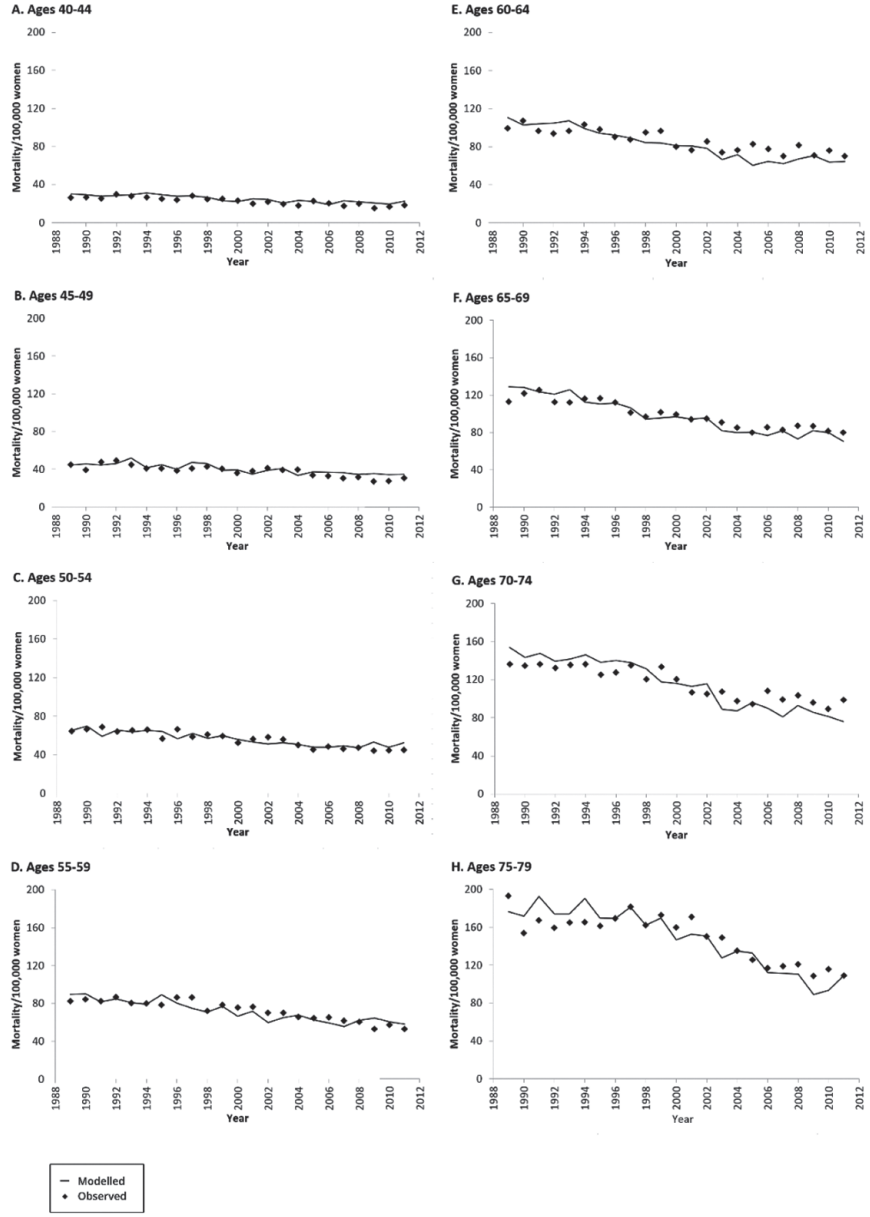
Screening strategy	Costs[a] (€)	Life-years gained[b]	ICER[c] (€/LYG)
No screening	1,049,093		
B 50-74 (current)	131,158	87	1,509
49 + B 50-74	151,956	93	dominated
B 48-74	157,237	96	2,774
B 45-74	203,062	107	4,128
A 45-49 + B 50-74	272,069	116	dominated
B 40-74	299,430	123	6,258
A 40-49 + B 50-74	464,913	139	10,029

[a]The costs of screening strategies reflect the additional costs relative to a situation without screening.
[b]Life-years gained relative to a situation without screening.
[c]ICERs are calculated as the difference in costs divided by the difference in life-years gained between a strategy and the previous, less effective, strategy in the ranking (no screening for the current strategy).
Abbreviations: life-year gained (LYG); biennial screening (B); annual screening (A).

3. Observed and modelled breast cancer incidence between 1989 and 2011 in five-year age groups, per 100,000 women



4. Observed and modelled breast cancer mortality between 1989 and 2011 in five-year age groups, per 100,000 women



5. Results of univariate sensitivity analyses (3.5% discounted)

Scenario	CER[a] (€/LYG) current screening	ICERs (€/LYG) additional screening[b]			
		48 + B 50-74	B 45-74	B 40-74	A 40-49 + B50-74
Baseline	3,376	5,329	7,628	10,826	18,759
50% of women undergo additional mammogram[c] after recall for further assessment	3,269	5,129	7,443	10,653	18,595
PPV for women aged <50 is 30% of PPV for women aged ≥50	3,376	5,584	7,833	11,020	18,938
PPV for women aged <50 is 50% of PPV for women aged ≥50	3,376	5,177	7,505	10,710	18,651
Breast cancer incidence is constant over the years	4,139	6,404	8,509	11,666	21,459
Breast cancer incidence increases annually with 1%	2,737	5,298	7,326	9,800	18,691
Test sensitivity of DM for women aged <50 is 50% of test sensitivity for women aged ≥55	3,655	8,076	11,075	15,299	21,170
Test sensitivity of DM for women aged <50 is equal to test sensitivity for women aged ≥55	3,263	4,457	6,596	9,204	18,020

[a]The cost-effectiveness ratio of current screening is calculated as the difference in costs between current screening and no screening divided by the life-years gained by current screening.
[b]ICERs of additional screening are calculated as the difference in costs divided by the difference in life-years gained between a strategy and the previous, less effective, strategy in the ranking.
[c] 50% of women who were screened and recalled for further assessment undergo a second mammogram in the hospital, in addition to the screening mammogram
Abbreviations: cost-effectiveness ratio (CER); incremental cost-effectiveness ratio (ICER); life-year gained (LYG); biennial screening (B); annual screening (A); positive predictive value (PPV); below age 50 (<50); age 50 and older (≥50); age 55 and older (≥55).

6. Effects and incremental cost-effectiveness ratios adjusted for quality of life

Screening strategy	Additional screening rounds	Costs[a] (€)	QALYs gained[b]	ICER[c] (€/QALY gained)
No screening		1,161,008		
B 50-74 (current)		137,057	54	2,529
49 + B 50-74	1	158,680	58	dominated
B 48-74	1	163,704	61	4,060
B 45-74	2.5	210,234	69	5,802
A 45-49 + B 50-74	5	281,643	75	dominated
B 40-74	5	306,590	80	8,259
A 40-49 + B 50-74	10	475,420	92	14,300

[a]Both costs and QALYs gained are 3.5% discounted and are presented per 1,000 women, followed over their lifetime from age 40. The costs of screening strategies reflect the additional costs relative to a situation without screening.
[b]QALYs gained relative to a situation without screening.
[c]ICERs are calculated as the difference in costs divided by the difference in QALYs gained between a strategy and the previous, less effective, strategy in the ranking (no screening for the current strategy).
Abbreviations: quality-adjusted life-year (QALY); biennial screening (B); annual screening (A).

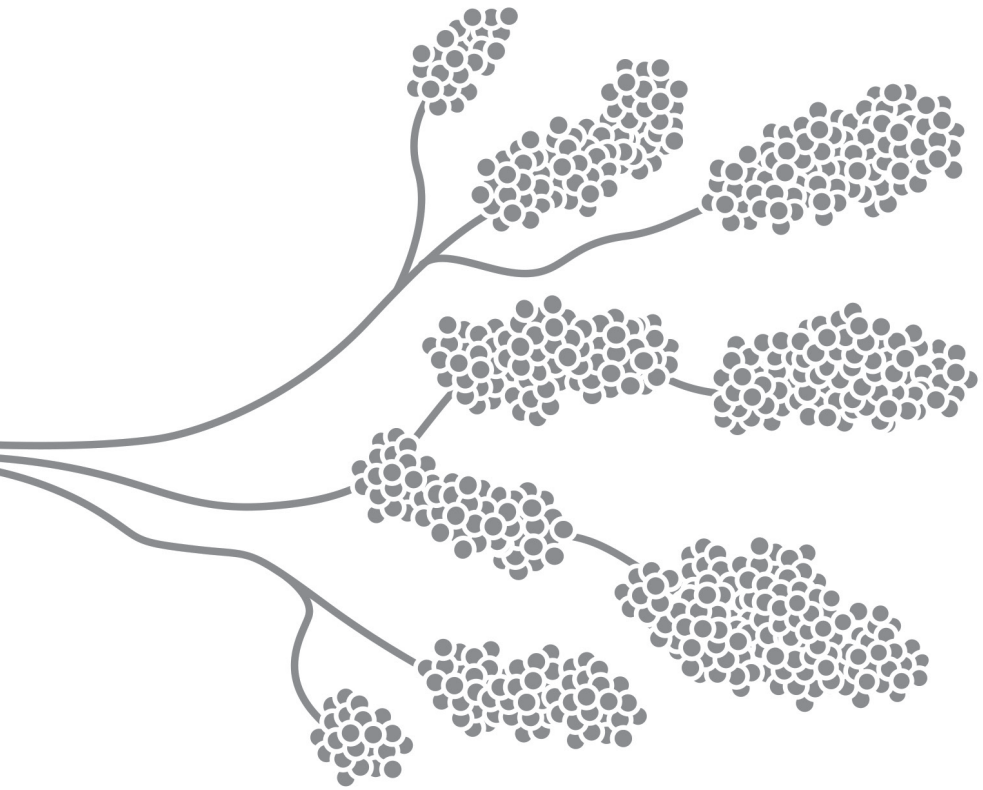
7. Pathway of screening and treatment

Women with a positive screening mammogram will be referred by their GP to a hospital. In order to discriminate between true positive and false positive results, these women will undergo a consultation, an additional mammogram and a proportion of these women will undergo ultrasound-guided biopsy. Women with a false positive or true negative result will be invited to the subsequent screening round. Women with a false negative result are invited to subsequent screening if the tumour is not clinically detected during the interval between screening rounds. Women with a true positive result may receive (adjuvant) treatment. In our model, treatment starts immediately once a tumour is screen- or clinically detected.

References

1. 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.
2. Nederlandse Kankerregistratie. http://www.cijfersoverkanker.nl/. (Accessed: December, 2014)
3. van Luijt PA, Fracheboud J, Heijnsdijk EA, den Heeten GJ, de Koning HJ, National Evaluation Team for Breast Cancer Screening in Netherlands Study G. Nation-wide data on screening performance during the transition to digital mammography: observations in 6 million screens. Eur J Cancer 2013;49:3517-25.
4. 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.
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. 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.
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. 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. Sant M, Allemani C, Capocaccia R, Hakulinen T, Aareleid T, Coebergh JW, Coleman MP, Grosclaude P, Martinez C, Bell J, Youngson J, Berrino F, et al. Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. Int J Cancer 2003;106:416-22.
10. 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.
11. Early Breast Cancer Trialists' Collaborative G. 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.
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. 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.
14. 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.
15. 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.
16. de Gelder R, Heijnsdijk EA, Fracheboud J, Draisma G, de Koning HJ. The effects of population-based mammography screening starting between age 40 and 50 in the presence of adjuvant systemic therapy. Int J Cancer 2014.



Chapter 5

Risk stratification in breast cancer screening: cost-effectiveness and harm-benefit ratios for low-risk and high-risk women

Valérie D.V. Sankatsing, Nicolien T. van Ravesteyn, Eveline A.M. Heijnsdijk, Mireille J.M. Broeders, Harry J. de Koning

International Journal of Cancer 2020;147(11):3059-3067.

ABSTRACT

Background: In mammography screening programmes, women are screened according to a one-size-fits-all principle. Tailored screening, based on risk levels, may lead to a better balance of benefits and harms.

Methods: With microsimulation modelling, we determined optimal mammography screening strategies for women at lower (relative risk (RR)0.75) and higher (RR1.8) than average risk of breast cancer, eligible for screening, using the incremental cost-effectiveness ratio (ICER) of current uniform screening in the Netherlands (biennial (B) 50-74) as a threshold ICER. Strategies varied by interval (annual (A), biennial (B), triennial (T)) and age range. The number of life-years gained (LYG), breast cancer deaths averted, overdiagnosed cases, false-positive mammograms, ICERs and harm-benefit ratios were calculated.

Results: Optimal risk-based screening scenarios, below the threshold ICER of €8,883/LYG, were T50-71 (€7,840/LYG) for low-risk and B40-74 (€6,062/LYG) for high-risk women. T50-71 screening in low-risk women resulted in a 33% reduction in false-positive findings, a similar reduction in costs and improved harm-benefit ratios compared to the current screening schedule. B40-74 in high-risk women led to an increase in screening benefit, compared to current B50-74 screening, but a relatively higher increase in false-positive findings.

Conclusions: In conclusion, optimal screening consisted of a longer interval and lower stopping age than current uniform screening for low-risk women, and a lower starting age for high-risk women. Extending the interval for women at lower risk from biennial to triennial screening reduced harms and costs while maintaining most of the screening benefit.

INTRODUCTION

Most European countries offer uniform mammography screening between age 50 and 69 or 74 to all eligible women, regardless of their individual risk of breast cancer[1]. This is in agreement with the European guidelines for quality assurance in breast cancer screening and diagnosis[2]. However, besides sex and age, there are other known risk factors for breast cancer that may affect the benefits and harms of screening, including breast density, family history of breast cancer, having had a previous breast biopsy and lifestyle-related factors[3, 4]. A risk-based screening approach may positively affect the balance between benefits and harms. In such an approach, the target age range and the screening interval are adjusted to different risk levels. There are several ongoing studies with respect to risk-based breast cancer screening in Europe, which collect information on risk factors for breast cancer for risk stratification[5-7].

The potential of risk-based screening has also been explored by modelling studies. Studies from the United States showed that biennial screening from age 40 onward in women at increased risk of breast cancer is cost-effective[8] and results in a similar ratio of false-positive (FP) findings and life-years gained (LYG) as biennial screening in average-risk women between age 50 and 74[9]. There are only a few European modelling studies that addressed risk-based mammography screening strategies. These studies also found that risk-based screening is cost-effective compared with uniform screening and that it reduces overall harms[10, 11]. Only one of these studies searched for optimal risk-based strategies for specific risk groups[10]. However, the authors had limited access to national data on screening performance and did not include ductal carcinoma in situ (DCIS) in their model[10].

In this study, we aimed to determine the optimal (i.e. most effective in terms of life-years gained) screening strategy for low-risk and high-risk women eligible for digital mammography screening in the Netherlands (no BRCA1/2 mutation carriers), under the condition that the cost-effectiveness did not exceed that of current uniform screening. We also evaluated harm-benefit ratios of these strategies.

METHODS

Model parameters and assumptions

We used the MIcrosimulation SCreening ANalysis (MISCAN) model, which simulates individual life histories [12]. Subsequently, the natural history of breast cancer (without screening) is simulated resulting in the onset of breast cancer in some women, which may be diagnosed and may eventually lead to breast cancer death. In MISCAN, breast cancer starts with the development of preclinical ductal carcinoma in situ (DCIS), which may progress through the invasive successive stages T1a, T1b, T1c and T2+ (semi-Markov process). A very small fraction of DCIS is assumed to regress[13].

Model outcomes are first estimated for a situation without screening, in which breast cancers can be detected when in the reach the clinical detectable phase or progress to the next preclinical stage. Hereafter, mammography screening and improvements in survival after screen-detection are modelled. In the presence of screening, tumours can also become screen-detected during the preclinical phase. Screening can therefore lead to earlier detection and treatment of breast cancer and possibly to aversion of breast cancer death. In the model, clinically and screen-detected cancers are treated with primary therapy and may also be treated with adjuvant therapy, which improves survival.

Recently, we recalibrated several parameters of the model, including stage- and age-specific sensitivity of screen-film and digital mammography, breast cancer background incidence, stage- and age-specific mean duration of preclinical screen-detectable breast cancer and progression and regression rate of DCIS. These parameters were calibrated to recent data from the Dutch screening programme on interval cancers (2004-2011), screen-detected cancers (2004-2013) and stage distribution at detection of cancers[14, 15]. Simultaneously, parameters were calibrated to data from the Netherlands Comprehensive Cancer Organisation[14] on breast cancer incidence between 1975-2013 by five-year age groups.

In addition, probabilities of being treated with adjuvant treatment (endocrine-, chemo-, a combination or a combination with targeted therapy) by age, stage

and calendar year were updated using data from the Netherlands Comprehensive Cancer Organisation over the years 2004-2011[14]. Values of important model parameters are listed in Table 1.

We simulated a cohort of women born in 1974. The time horizon for estimating effects was from age 40 until death. In order to fairly compare strategies consisting of different screening intervals, we simulated 100% attendance to screening. We also simulated 100% adherence to primary treatment.

Risk groups

We simulated different risk groups: low, average (total population) and high, based on common risk factors (other than breast density) and their relative risk (RR) estimate (Supplementary material 1). Combining these factors led to a combined low RR of 0.75 and a combined high RR of 1.8. Risk was defined based on RR rather than absolute risk to make results more comparable to countries with different average absolute breast cancer risk. BRCA1 and BRCA2 mutation carriers were not included in the high-risk group.

Our standard model represented the total Dutch female population with an overall average risk for breast cancer (RR=1). The underlying breast cancer incidence in the absence of screening (by age and calendar year) was modified for the low- and high-risk group by multiplying the onset hazards of average risk by the RR of low- and high-risk (Supplementary material 2). A woman's RR was assumed to be constant from birth until death.

Besides current screening (biennial 50-74), low and high-risk women were simulated to be screened with a range of strategies, which varied by interval, starting and stopping age. Intervals were annual (A), biennial (B) or triennial (T). Digital mammography was the screening test in all scenarios. Outcomes of all strategies were compared to no screening and outcomes of optimal screening also to current B50-74 screening in the respective risk-group. Strategies for women at low risk were less intensive and strategies for women at high risk more intensive than B50-74:

Table 1. Model input parameters with current uniform screening

Input parameters	Calibrated or fixed value
Stage-specific sensitivity of digital mammography DCIS	0.87
Stage-specific sensitivity of digital mammography T1a	0.55
Stage-specific sensitivity of digital mammography T1b	0.48
Stage-specific sensitivity of digital mammography T1c	0.86
Stage-specific sensitivity of digital mammography T2+	1
Relative sensitivity‡ of younger ages digital mammography	0.75
Breast cancer onset	0.291
Onset hazard age 30 years	7.049E-05
Onset hazard age 50 years	0.010
Onset hazard age 70 years	0.019
Onset hazard age 100 years	0.024
Stage-specific duration† (years) screen-detectable preclinical stage DCIS	3.5
Stage-specific duration† (years) screen-detectable preclinical stage T1a	0.1
Stage-specific duration† (years) screen-detectable preclinical stage T1b	0.5
Stage-specific duration† (years) screen-detectable preclinical stage T1c	1.5
Stage-specific duration† (years) screen-detectable preclinical stage T2+	0.9
Probability of DCIS progressing into next stage (T1A) immediately	0.842
Probability of DCIS progressing into next stage (T1A) slowly	0.038
Stage-specific reduction in risk of dying from breast cancer after screen-detection* DCIS	100%
Stage-specific reduction in risk of dying from breast cancer after screen-detection* T1a, N-	88%
Stage-specific reduction in risk of dying from breast cancer after screen-detection* T1b, N-	84%
Stage-specific reduction in risk of dying from breast cancer after screen-detection* T1c, N-	75%
Stage-specific reduction in risk of dying from breast cancer after screen-detection* T2+, N-	56%

Table 1. (Continued)

Input parameters	Calibrated or fixed value
Stage-specific reduction in risk of dying from breast cancer after screen-detection* T1a, N+	72%
Stage-specific reduction in risk of dying from breast cancer after screen-detection* T1b, N+	64%
Stage-specific reduction in risk of dying from breast cancer after screen-detection* T1c, N+	49%
Stage-specific reduction in risk of dying from breast cancer after screen-detection* T2+, N+	24%
Costs late stage breast cancer	€18000
Stage-specific treatment costs DCIS	€4569
Stage-specific treatment costs T1a, N-	€4333
Stage-specific treatment costs T1b, N-	€5057
Stage-specific treatment costs T1c, N-	€11146
Stage-specific treatment costs T2+, N-	€10815
Stage-specific treatment costs T1a, N+	€16103
Stage-specific treatment costs T1b, N+	€6744
Stage-specific treatment costs T1c, N+	€20822
Stage-specific treatment costs T2+, N+	€15063
Positive predictive value of recall with digital mammography screening	28%
Positive predictive value of biopsy	67%
Positive predictive value of recall by a general practitioner	58%

†Given values for the stage-specific durations represent values for women aged 50 years. In the model, different durations are used for different age groups ‡The relative sensitivity is 1 for age 55 and older, 0.75 for <50 and interpolated between 0.75 and 1 for 50-55 *Given values for the stage-specific reduction in risk of dying from breast cancer after screen-detection represent average values across all ages. In the model, different values are used for different age groups N.B. Onset hazards were adjusted for risk based screening; the positive predictive value of recall with digital mammography screening was adjusted for annual and triennial screening

Low risk (RR=0.75; 101 strategies):
- Starting age: between 50-60 years
- Stopping age: between 64-74 years
- Interval: B or T

High risk (RR=1.8; 182 strategies):
- Starting age: between 40-50 years
- Stopping age: between 74-84 years
- Interval: A or B

Predicted lifetime outcomes included LYG, breast cancer deaths averted, FP findings, overdiagnosis and costs.

Cost-effectiveness

Total costs (represented in Euros) were based on costs associated with screening, additional diagnostics and stage-specific treatment, using data from the MRISC study[16] and represent the additional costs compared with no screening. Costs and LYG were discounted at 3.5% per year. in agreement with the recommendations of the National Institute for health and Care Excellence (NICE).[17]

Strategies were ranked according to their effectiveness (in terms of LYG). Incremental cost-effectiveness ratios (ICERs) were calculated as the difference in costs divided by the difference in LYG between two consecutive strategies (i.e. costs required to yield one additional LYG, compared to previous strategy).

B50-74 (current) in the total population was very close to but not on the frontier and an ICER for this strategy could therefore not be calculated. We used the ICER of B48-72 (same number of screening rounds (13) and similar cost-effectiveness ratio; €8,883 per LYG; Table 2) as a proxy for that of current screening and as the threshold to compare risk-based strategies to. The most effective strategy under this ICER was considered optimal. Selected risk-based strategies were therefore optimal in terms of the ratio of costs to effects, the cost-effectiveness. Optimal strategies were thus not selected based on the highest possible benefit.

Harm-benefit ratios

We compared harm-benefit ratios between optimal screening in risk groups and current uniform screening. FP findings were calculated using screen-detected cancers (model output) and the positive predictive value (PPV_1) of recall (PPV=TP/(TP+FP)).

The PPV of biennial screening in the Dutch screening programme was recently reported to be 28%[18]. Based on the literature, we estimated that the PPV for women aged 40-49 years is 12%[19, 20]. The ratios of the PPV of annual and triennial screening to that of biennial screening were calculated based on findings in the U.S., Norway and Spain published by Domingo et al.[21] Subsequently, these ratios were used to estimate the PPV for annual and triennial screening in the Netherlands, which resulted in a PPV of 24% (≥50 years) and 10% (40-49 years) for annual screening and a PPV of 31% (≥50 years) for triennial screening[21].

Overdiagnosis was calculated by comparing the number of breast cancers detected in women in the presence and in the absence of screening, using lifelong follow-up and accounting for an increase in the background incidence of breast cancer over the years.

Scenario analyses

To assess the influence of realistic attendance on our outcomes, we performed a scenario analysis using observed attendance in the Netherlands (80%), instead of full attendance.

In a second scenario analysis, the RR of low and high-risk women was modelled to vary by age as opposed to a constant RR. We modelled a substantial attenuation of the RR with increasing age for both risk groups – i.e. both the RR of 1.8 of high risk women and the RR of 0.75 of low risk women attenuate gradually with increasing age and approach RR 1 at higher age.

RESULTS

Offering all eligible women uniform current B50-74 screening gained 206 life-years (non-discounted) and averted 16 breast cancer deaths at the expense of 187 FP findings and 5 overdiagnosed cases per 1000 women followed from age 40 until death, compared to no screening (Table 3).

Low-risk

Considering the threshold ICER of €8,883 per LYG, optimal screening for low-risk women was T50-71 (ICER:€7,840 per LYG), with 5 fewer screening rounds than current B50-74 (Table 2, 3). T50-71 led to 134 LYG and 10 breast cancer deaths averted per 1000 women (non-discounted), compared to no screening, at the expense of 102 FP findings and 3 overdiagnosed cases. Harm-benefit ratios of T50-71 screening for low-risk women were better than that of uniform B50-74 screening for all women (RR=1) and B50-74 screening for low-risk women (Table 2, 3). T50-71 led to 50 fewer FP findings than B50-74 in low-risk women (-33%) and 31 fewer LYG (-19%; Table 1), thus also resulting in a better balance between benefits and harms. Optimal screening reduced cost by 37% (Table 2). Triennial strategies yielded higher (or similar) benefit for lower costs compared to biennial alternatives (Figure 1).

High-risk

Optimal screening for high-risk women was B40-74 (ICER:€6,062 per LYG), which consists of 5 additional rounds before the starting age of current uniform screening. B40-74 resulted in increased benefits but also in an increase in harms (Table 3). Harm-benefit ratios for B40-74 in high-risk women were slightly more favourable than for current uniform B50-74 screening in the total population when overdiagnosis was considered and considerably less favourable with FP findings (Table 3). B40-74 led to 77 additional LYG (+25%) compared to B50-74 in high-risk women and 113 additional FP findings (+44%; Table 2). Harm-benefit ratios (FP findings/LYG) were less favourable for annual than biennial scenarios. Under a threshold of net costs of approximately €500,000, annual strategies were dominated by biennial strategies (Figure 2).

Table 2. Efficient strategies in women at low, average or high risk of breast cancer with corresponding screening outcomes† and incremental harm-benefit and cost-effectiveness ratios (ICER)

	No. rounds	False-positives	Life-years gained	Harm-benefit ratio‡	Net costs (€)§	Life-years gained§	ICER§
Low risk (RR=0.75)							
T 58-64	3	46	58	0.80	70,325	16	4,423
T 55-64	4	56	78	0.72	100,886	22	4,851
T 54-66	5	69	93	0.73	125,326	26	5,819
T 53-65	5	66	95	0.69	131,485	27	6,158
T 52-67	6	78	110	0.72	156,917	31	6,358
T 50-68	7	87	125	0.70	191,191	36	6,995
T 50-71	8	102	134	0.76	207,655	38	7,840
T 50-74.	9	116	140	0.82	223,101	40	11,881
B 50-72	12	140	161	0.87	316,521	46	14,829
B 50-74	13	152	166	0.92	331,785	47	15,264
Average Risk (total population; RR=1)							
B 48-72	13	204	219	0.93	336,863	64	8,883
B 50-74	13	187	206	0.90	302,904	59	-

Table 2. (Continued)

	No. rounds	False-positives	Life-years gained	Harm-benefit ratio‡	Net costs (€)§	Life-years gained§	ICER§
High risk (RR=1.8)							
B 50-74	13	258	303	0.85	232,285	88	2,655
B 48-74	14	310	333	0.93	275,425	98	4,109
B 46-74	15	344	358	0.96	323,141	107	5,187
B 40-74	18	371	380	0.98	376,488	116	6,062
B 41-75	18	378	383	0.99	382,416	117	9,880
A 40-74	35	551	488	1.13	843,045	151	13,429
A 40-75	36	561	491	1.14	855,639	152	17,991
A 40-76	37	570	493	1.15	867,805	152	24,334
A 40-77	38	578	495	1.17	879,468	153	29,156
A 40-78	39	586	497	1.18	890,687	153	37,397
A 40-81	42	607	500	1.21	920,902	153	50,359
A 40-84	45	623	502	1.24	945,578	154	82,252

†Per 1000 women aged 40 years followed over their lifetime, compared to no screening ‡False-positive findings/life-year gained §Discounted at 3.5% per year. N.B. the optimal strategy is marked in grey; Due to rounding of numbers used in the analysis, it may seem that outcomes represented in the Table do not add up to the total number. *ICER: incremental cost-effectiveness ratio; RR: relative risk; T: triennial (3-year interval); B: biennial (2-year interval); A: annual (1-year interval)*

Table 3. Optimal screening scenario and corresponding outcomes and ratios per risk group

	Current screening RR = 1	**Low risk** RR = 0.75	**High risk** RR = 1.8
(Optimal) scenario	B 50-74	T 50-71	B 40-74
Screening rounds	13	8	18
Screening outcomes*			
False positives	187	102	371
Overdiagnosis	5	3	7
BC deaths averted	16	10	26
Life-years gained	206	134	380
Harm-benefit ratios			
False-positives/ deaths averted	11.8	10.1	14.5
False-positives/ life-years gained	0.90	0.76	0.98
Overdiagnosis/ deaths averted	0.34	0.31	0.29
Overdiagnosis/ life-years gained	0.03	0.02	0.02

*Screening outcomes are presented per 1000 women, aged 40 years followed over their lifetime invited for screening. N.B. Due to rounding of numbers used in the analysis, it may seem that outcomes represented in the Table do not add up to the total number.*T: triennial (3-year interval); B: biennial (2-year interval)*

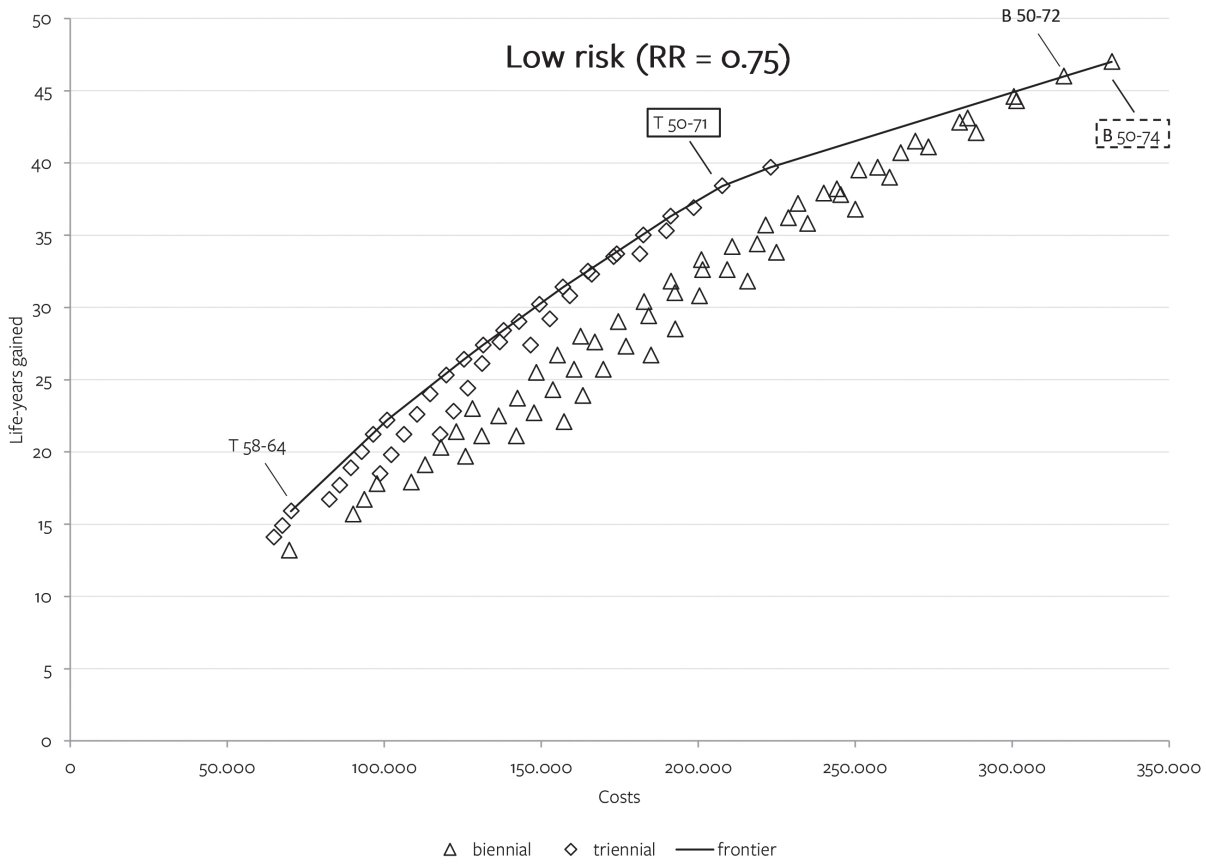


Figure 1. Efficiency frontier of biennial and triennial strategies in low-risk women
The square marks the optimal strategy and the dashed square current uniform screening

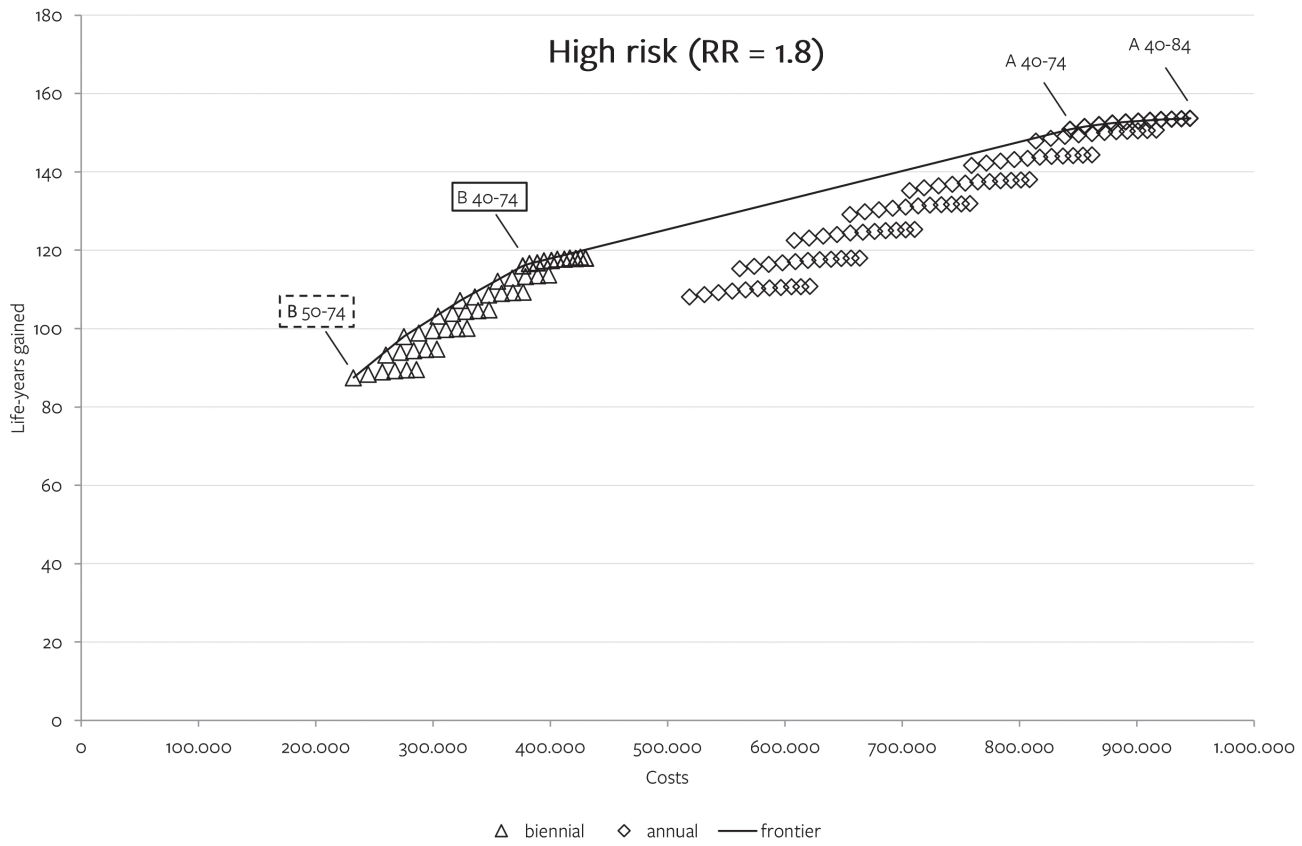


Figure 2. Efficiency frontier of annual and biennial strategies in high-risk women
The square marks the optimal strategy and the dashed square current uniform screening

Scenario analyses

When modelling observed attendance (80%), costs and effects decreased (Supplementary material 3). Modelling realistic attendance did not change the optimal strategy for women at lower than average risk of breast cancer. For women at higher risk there was a slight change in the optimal strategy, which shifted from B40-74 to B42-74 with realistic attendance, as the ICER of B40-74 just exceeded the cost-effectiveness threshold of current screening in this case (Supplementary material 3). When assuming an attenuating RR with increasing age, both benefits and harms including deaths averted, LYG, false-positive findings and overdiagnosis decreased with similar percentages for high risk women and increased with similar percentages for low risk women. For high risk women, lifetime costs increased considerably compared to high risk women with a constant RR, which – in combination with the decreased screening effects - negatively affected the cost-effectiveness of screening strategies (data not shown).

DISCUSSION

We provided estimates of optimal risk-based mammography screening strategies, with ICERs similar to that of uniform screening, for women with a relatively low and high risk of breast cancer, compared to the average population. Offering triennial screening to women at lower risk of breast cancer, would limit the harms associated with biennial screening in this group, by resulting in less overdiagnosis and fewer false-positive findings and additional diagnostic tests, while maintaining most of the benefit of biennial screening. Additional screening before age 50 for high-risk women would lead to greater benefit but also to a relatively high increase in FP findings.

In this study, optimal screening scenarios for both low and high risk were assessed by quantifying screening outcomes for a large number of strategies. We used a validated model, which was recently updated with the newest available screening data and improvements in adjuvant therapy. Our results are likely to be relevant to other (European) countries, in particular those with comparable breast cancer incidence, screening indicators and programme characteristics. However, before

implementing risk-based screening, there are still many practical and ethical issues that need to be considered.

This study had several limitations. First, we did not take into account that a woman's relative risk can change over time in our base case analysis. Mammographic breast density, for example, generally decreases with increasing age altering test sensitivity[22]. Our scenario analysis showed that modelling a decreasing RR with increasing age in high-risk women, leads to considerably higher costs than modelling a constant 'high' RR. A decreasing RR with increasing age will therefore result in a less intensive strategy being considered as optimal. Risk prediction models usually generate absolute risk levels[23]. We used RR levels however, to make our results generalisable to other countries. Second, we assumed that breast cancer risk only affects the incidence of breast cancer and no other important parameters including tumour growth rate, sojourn time and stage distribution. Assuming a higher growth rate for high-risk women, for example, could cause the optimal interval to shift from biennial to annual screening. Further, we could not calculate the ICER of current uniform screening as B50-74 was close to but not on the frontier and thus we used B48-72 as the cost-effectiveness threshold.

Finally, we had no data on the distribution of risk groups among the Dutch female population or the distribution of breast density among risk groups. We could therefore not assess the impact of a risk-based screening programme for the population as a whole. Considering the effects of risk-based screening for low, average and high-risk women all together, with tailored scenarios for each risk group, is an important next step for future research. Our study is however an important first step as it shows that both the screening programme itself and women at lower and higher than average risk of breast cancer could benefit from tailored screening, while the cost-effectiveness is not affected. Preliminary data of the Dutch PRISMA study (approximately 5000 women) show that the risk factors 'age at first child 20-24 years' and 'first degree family member with breast cancer' with RRs corresponding to our low- and high-risk group (0.7 and 1.8 respectively) are both present in approximately 19% of the study population[6]. However, it is still

unknown how large the low-risk and high-risk groups are when taking into account more single risk factors and combinations of risk factors.

We found that low-risk women could benefit, in terms of reduced harms, from less frequent, triennial screening, which is consistent with other work[8, 10]. However, one of these studies, in which optimal strategies for risk groups were determined by selecting the least expensive strategy with benefit similar to that of current uniform screening[10], found that annual screening is optimal for high-risk women, which contradicts our results. This can be (partially) explained by the fact that in our study there was a greater restriction in terms of costs with the selection of optimal strategies, since we only allowed for strategies with a lower ICER than current screening to be optimal. In general, studies show that risk-based screening leads to improved harm-benefit ratios and is more efficient than uniform screening, which is in agreement with our results[10, 24].

We used full attendance to screening and full adherence to treatment (100%) as the goal of our study was to select optimal strategies for women who fully adhere to screening and treatment recommendations. In reality, not all women attend to every screening round. When modelling realistic attendance, screening strategies with shorter intervals could turn out to be more cost-effective than with full attendance. Shorter screening intervals compensate for lower than full attendance by ensuring that women, on average, are still screened the optimal number of screening rounds.[25] For women who do show up for every screening round this could thus result in too intensive screening being recommended as optimal. Our scenario analysis showed that using observed attendance did not influence the conclusion with respect to women at lower than average risk of breast cancer. The optimal strategy for women at higher risk reduced with one screening round – i.e. a two year higher starting age. The conclusion that women at higher risk are optimally screened before the age of 50, in their early 40s, however still holds.

Weighing harms and benefits is subject to personal preferences and it is therefore important to consider different ratios of benefits and harms and perspectives.

Women attending screening may have a higher tolerance for accepting increasing harms with increasing benefits than policymakers. Research on attitudes to FP mammography results among 479 women without a history of breast cancer showed that 63% of women would accept up to 500 FP findings per life saved[26]. Our findings suggest that optimal B40-74 screening in high-risk women would result in approximately 15 FP findings per death averted, which is much lower. Harm-benefit ratios worsen when switching to annual screening for high-risk women. In addition, the net costs of screening double when switching from biennial to annual screening.

Stratified screening will lead to differences in screening approaches between women. This may be difficult to accept, particularly for women with the lowest risk, for whom risk-based screening will probably result in less frequent screening. Less screening is beneficial for women at lower risk as it results in a reduction in screening harms. However, although the ratio of costs to effects with the optimal strategy for low risk women identified in our study is similar to current uniform screening, the absolute screening benefit decreases for these women, compared to uniform screening. If women will not agree with these differences, the attendance rate of the screening programme could drop. It is therefore important to inform women and screening professionals on breast cancer risk and the harms and benefits associated with screening, related to different risk levels. Women at lower risk of breast cancer could then make an informed decision on whether a reduction in screening harms outweigh a slightly reduced number of LYG.

Conclusion

Our findings show that risk-based screening has the potential to respectively reduce absolute harms and increase absolute benefits in low- and high-risk women, while maintaining current cost-effectiveness. The optimal low-risk scenario T50-71 improves the ratio of false-positive findings to LYG for low-risk women, thus resulting in a greater reduction in harms than in benefit, compared to current screening. With the optimal high-risk scenario B40-74 however, the increase in screening benefit is smaller than the increase in false-positive findings. Future

research is needed to link the RRs to risk factors and their prevalence in order to estimate the impact of risk-based screening at the population level.

REFERENCES

1. Giordano L, von Karsa L, Tomatis M, Majek O, de Wolf C, Lancucki L, Hofvind S, Nystrom L, Segnan N, Ponti A, Eunice Working G, Van Hal G, et al. Mammographic screening programmes in Europe: organization, coverage and participation. *J Med Screen* 2012;19 Suppl 1: 72-82.
2. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition--summary document. *Ann Oncol* 2008;19: 614-22.
3. Nelson HD, Zakher B, Cantor A, Fu R, Griffin J, O'Meara ES, Buist DS, Kerlikowske K, van Ravesteyn NT, Trentham-Dietz A, Mandelblatt JS, Miglioretti DL. Risk factors for breast cancer for women aged 40 to 49 years: a systematic review and meta-analysis. *Ann Intern Med* 2012;156: 635-48.
4. Tice JA, Miglioretti DL, Li CS, Vachon CM, Gard CC, Kerlikowske K. Breast Density and Benign Breast Disease: Risk Assessment to Identify Women at High Risk of Breast Cancer. *J Clin Oncol* 2015;33: 3137-43.
5. Gabrielson M, Eriksson M, Hammarstrom M, Borgquist S, Leifland K, Czene K, Hall P. Cohort Profile: The Karolinska Mammography Project for Risk Prediction of Breast Cancer (KARMA). *Int J Epidemiol* 2017;46: 1740-1g.
6. https://www.prisma-studie.nl. Accessed: March 2020.
7. Giordano L, Gallo F, Petracci E, Chiorino G, Segnan N, Andromeda working g. The ANDROMEDA prospective cohort study: predictive value of combined criteria to tailor breast cancer screening and new opportunities from circulating markers: study protocol. *BMC Cancer* 2017;17: 785.
8. 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.
9. van Ravesteyn NT, Miglioretti DL, Stout NK, Lee SJ, Schechter CB, Buist DS, Huang H, Heijnsdijk EA, Trentham-Dietz A, Alagoz O, Near AM, Kerlikowske K, et al. Tipping the balance of benefits and harms to favor screening mammography starting at age 40 years: a comparative modeling study of risk. *Ann Intern Med* 2012;156: 609-17.
10. Vilaprinyo E, Forne C, Carles M, Sala M, Pla R, Castells X, Domingo L, Rue M, Interval Cancer Study G. Cost-effectiveness and harm-benefit analyses of risk-based screening strategies for breast cancer. *PLoS One* 2014;9: e86858.
11. Gray E, Donten A, Karssemeijer N, van Gils C, Evans DG, Astley S, Payne K. Evaluation of a Stratified National Breast Screening Program in the United Kingdom: An Early Model-Based Cost-Effectiveness Analysis. *Value Health* 2017;20: 1100-9.

12. Sankatsing VD, Heijnsdijk EA, van Luijt PA, van Ravesteyn NT, Fracheboud J, de Koning HJ. Cost-effectiveness of digital mammography screening before the age of 50 in The Netherlands. *Int J Cancer* 2015;137: 1990-9.

13. Segnan N, Minozzi S, Armaroli P, Cinquini M, Bellisario C, Gonzalez-Lorenzo M, Gianola S, Ponti A. Epidemiologic evidence of slow growing, nonprogressive or regressive breast cancer: A systematic review. *Int J Cancer* 2016;139: 554-73.

14. https://www.iknl.nl. Accessed: May 2019.

15. National Evaluation Team for Breast cancer screening (NETB). National evaluation of breast cancer screening in the Netherlands 1990-2011/2012. Thirteenth evaluation report. Erasmus MC University Medical Center Rotterdam.

16. Saadatmand S, Tilanus-Linthorst MM, Rutgers EJ, Hoogerbrugge N, Oosterwijk JC, Tollenaar RA, Hooning M, Loo CE, Obdeijn IM, Heijnsdijk EA, de Koning HJ. Cost-effectiveness of screening women with familial risk for breast cancer with magnetic resonance imaging. *J Natl Cancer Inst* 2013;105: 1314-21.

17. http://www.nice.org.uk/. Accessed: March 2020.

18. van Luijt PA, Fracheboud J, Heijnsdijk EA, den Heeten GJ, de Koning HJ, National Evaluation Team for Breast Cancer Screening in Netherlands Study G. Nation-wide data on screening performance during the transition to digital mammography: observations in 6 million screens. *Eur J Cancer* 2013;49: 3517-25.

19. 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.

20. Perry NM, Patani N, Milner SE, Pinker K, Mokbel K, Allgood PC, Duffy SW. The impact of digital mammography on screening a young cohort of women for breast cancer in an urban specialist breast unit. *Eur Radiol* 2011;21: 676-82.

21. Domingo L, Hofvind S, Hubbard RA, Roman M, Benkeser D, Sala M, Castells X. Cross-national comparison of screening mammography accuracy measures in U.S., Norway, and Spain. *Eur Radiol* 2016;26: 2520-8.

22. Sprague BL, Gangnon RE, Burt V, Trentham-Dietz A, Hampton JM, Wellman RD, Kerlikowske K, Miglioretti DL. Prevalence of mammographically dense breasts in the United States. *J Natl Cancer Inst* 2014;106.

23. Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Ann Intern Med* 2008;148: 337-47.

24. Pashayan N, Morris S, Gilbert FJ, Pharoah PDP. Cost-effectiveness and Benefit-to-Harm Ratio of Risk-Stratified Screening for Breast Cancer: A Life-Table Model. *JAMA Oncol* 2018;4: 1504-10.

25. Wilschut JA, Hol L, Dekker E, Jansen JB, Van Leerdam ME, Lansdorp-Vogelaar I, Kuipers EJ, Habbema JD, Van Ballegooijen M. Cost-effectiveness analysis of a quantitative immunochemical test for colorectal cancer screening. *Gastroenterology* 2011;141: 1648-55 e1.

26. Schwartz LM, Woloshin S, Sox HC, Fischhoff B, Welch HG. US women's attitudes to false positive mammography results and detection of ductal carcinoma in situ: cross sectional survey. *BMJ* 2000;320: 1635-40.

SUPPLEMENTARY MATERIAL

1. Justification of Relative Risks (RRs)

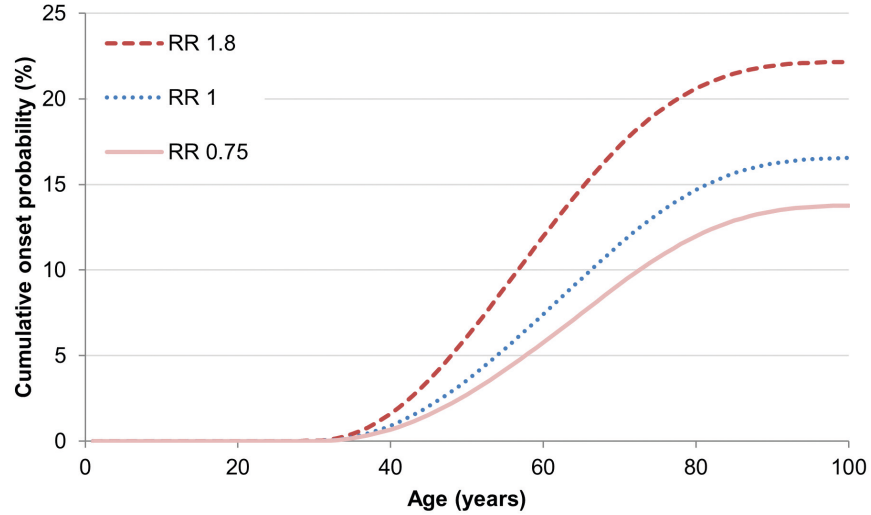
Low risk

Risk factor	Category	RR range	RR	source
Combination 1st child and no. children	25-29 vs. <20 y >5 vs. 1-2		0.5	[1]
Physical activity (5h/w)	per hour per week		0.7	[2]
Ever breastfeed	vs. never		0.78	[3]
Age at menopause 45-49 y	vs. 50-54 y		0.86	[4]
Age at menarche >= 15 y	vs. 13 y	0.87-0.92	0.9	[4,5]
Combined low risk estimate			**0.75**	

High risk

Risk factor	Category	RR range	RR	bron
History of benign breast disease, not otherwise specified	1 vs. 0	1.44-2.07	1.76	[6,7]
1st degree family member with breast cancer	1 vs. 0		1.8	[2, 8]
History of proliferative disease without atypia	1 vs. 0	1.66-2.02	1.84	[6,7]
Combination family member with breast cancer and smoking	1 vs. 0: 1.8 ever vs. never:1.11		2.0 (1.8*1.11)	[2, 5, 8]
Combined high risk estimate			**1.8**	

2. Simulated cumulative probability† of onset of breast cancer in women at low, average or high-risk



†The maximum value for the y-axis is 25%

3. Cost-effectiveness outcomes† with observed attendance (80%) for efficient strategies

	Net costs (€) ‡	LYG ‡	ICER
Low risk (RR=0.75)			
T 58-64	50,614	11	4,686
T 55-64	72,230	15	4,804
T 52-64	98,240	20	5,911
T 51-66	117,304	23	6,355
T_50_6	136,560	25	7,406
T 50-71	148,889	27	8,807
T 50-74	160,396	28	12,785
B 50-72	228,106	32	15,389
B 50-74	239,395	33	16,127
High risk (RR=1.8)			
B 50-74	170,827	61	2,819
B 48-74	200,631	68	3,922
B 46-74	234,199	75	4,937
B 44-74	271,600	81	5,937
B 42-74	313,068	88	6,582
B 40-74	360,280	93	9,635
A 41-74	709,858	120	12,620
A 40-74	758,332	124	14,689
A 40-75	767,543	124	18,423
A 40-77	784,556	125	28,354
A 40-80	806,899	125	37,239
A 40-81	813,276	125	63,768
A 40-84	829,316	126	80,204

†Per 1000 women aged 40 years followed over their lifetime, compared to no screening ‡ Discounted at 3.5% per year. N.B. the optimal strategy is marked in grey. *RR: relative risk; LYG: life years gained; ICER: incremental cost-effectiveness ratio; T: triennial (3-year interval); B: biennial (2-year interval); A: annual (1-year interval)*

References

1 Ewertz M, Duffy SW, Adami H-O, et al. Age at first birth, parity and risk of breast cancer: A meta-analysis of 8 studies from the nordic countries. Int J Cancer (1990) 46(4): 597-603.

2 Rookus MA, Mooij TM, van Leeuwen FE. Risicofactoren van mammacarcinoom. In: Wobbes T, Nortier JWR, Koning CCE (eds.). Handboek Mammacarcinoom. Utrecht: De Tijdstroom, 2007.

3 Bernier MO, Plu-Bureau G, Bossard N, Ayzac L, Thalabard JC. Breastfeeding and risk of breast cancer: a metaanalysis of published studies. Hum Reprod Update 2000; 6:374-86.

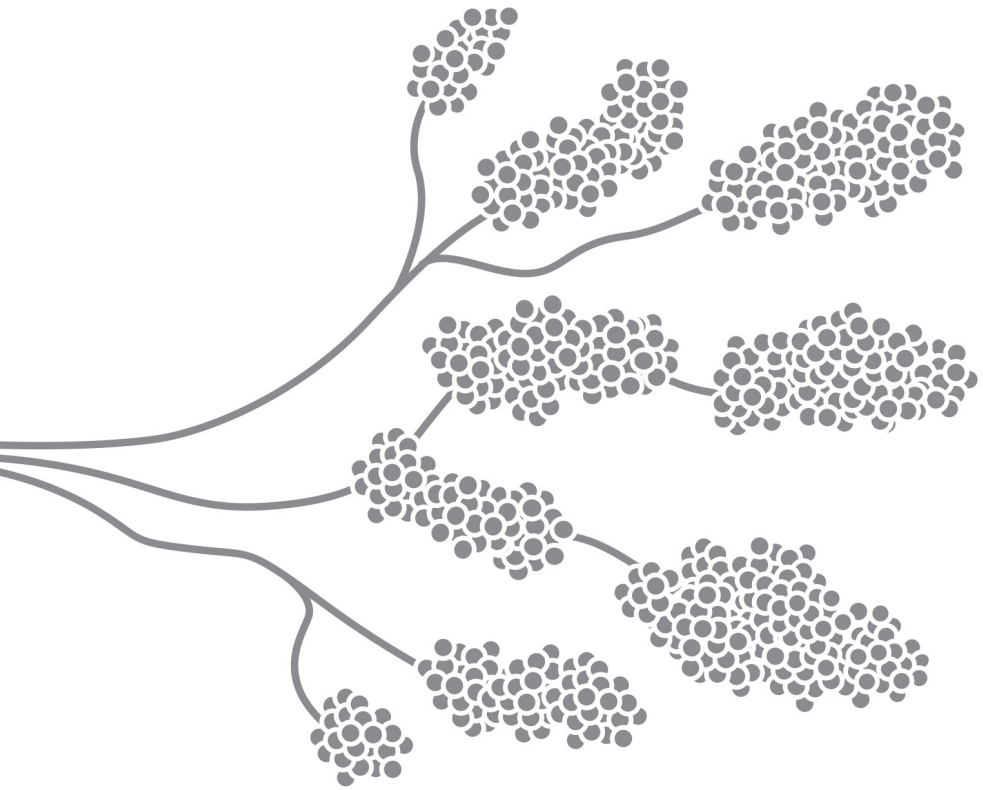
4 Collaborative Group on Hormonal Factors in Breast Cancer. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. Lancet Oncol. 2012; 13:1141-51.

5 Nelson HD, Zakher B, Cantor A, et al. Risk factors for breast cancer for women aged 40 to 49 years: a systematic review and meta-analysis. Annals of Internal Medicine (2012) 156(9): 635-648.

6 Tice JA, Miglioretti DL, Li CS, Vachon CM, Gard CC, Kerlikowske K. Breast density and benign breast disease: risk assessment to identify women at high risk of breast cancer. J Clin Oncol. 2015;33:3137-43.

7 Dyrstad SW, Yan Y, Fowler AM, Colditz GA. Breast cancer risk associated with benign breast disease: systematic review and metaanalysis. Breast Cancer Res Treat. 2015;149:569-75.

8 Collaborative Group on Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58 209 women with breast cancer and 101 986 women without the disease. The Lancet (2001) 358: 1389-1399.



Chapter 6

Cost-effectiveness of digital breast tomosynthesis in the Dutch breast cancer screening program: a probabilistic sensitivity analysis

Valérie D.V. Sankatsing, Nicolien T. van Ravesteyn, Karolina Juraniec, Sabine E. Grimm, Manuela A. Joore, Harry J. de Koning

Radiology 2020;297(1):40-48.

ABSTRACT

Background: Digital breast tomosynthesis (DBT) is a promising screening test, but its outcomes and cost-effectiveness remain uncertain. To determine if biennial DBT is cost-effective in a screening setting when compared with digital mammography (DM) in the Netherlands and to quantify the uncertainty.

Methods: In this study, performed from March 2018-February 2019, the MIcrosimulation SCreening ANalysis (MISCAN) model was used to conduct a probabilistic sensitivity analysis (PSA), consisting of 10,000 model runs with 1,000,000 women simulated per run. The Bayesian Cost-Effectiveness Analysis (BCEA) package and the Sheffield Accelerated Value of Information (SAVI) tool were used to process PSA outcomes. Two simulated cohorts born in 1970 were invited to biennial screening between age50-74; one to DM and one to DBT. DM input parameters were based on data from the Dutch breast cancer screening program. DBT parameters were based on literature and expert opinion. Willingness-to-pay-thresholds of €20,000 ($22,000) and €35,000 ($38,500) per life-year gained (LYG) were considered. Effects and costs were discounted at 3.5% per year.

Results: DBT gained 13 additional life-years per 1000 women invited to screening (7% increase;13/193), followed over lifetime, compared with DM and led to 2% (4/159) fewer false-positives . DBT screening led to incremental discounted lifetime effects of 5.09 LYG (95% confidence interval:-0.80-9.70) and an increase in lifetime costs of €137,555 ($151,311) per 1000 women (95% confidence interval:€31,093($34,202)-€263,537($289,891)) compared with DM, resulting in a mean incremental cost-effectiveness ratio (ICER) of €27,023 ($29,725) per LYG. The probability of DBT being more cost-effective was 0.36 at €20,000 and 0.66 at €35,000 per LYG.

Conclusion: Switching to biennial digital breast tomosynthesis (DBT) from digital mammography (DM) is not cost-effective at a willingness-to-pay-threshold of €20,000 per life years gained (LYG), but DBT has a higher probability of being more cost-effective than DM at €35,000 per LYG.

INTRODUCTION

Several European trials and observational studies have demonstrated that DBT screening, as a stand-alone modality, combined with DM or with synthetic 2D images reconstructed from the 3D DBT acquisitions, improves the breast cancer detection rate compared with DM alone[1-6]. Although many studies show improved breast cancer detection rate with DBT, only a few report interval cancers, which are required to calculate the sensitivity of DBT[1,7]. Also, the few reported sensitivity estimates were based on interval breast cancers of the first screening round only as there are no long term trial results available. Estimates for the positive predictive value (PPV) for DBT in a screening setting vary from similar to that of DM to a twofold higher PPV[1,3,5,6].

When considering a new screening modality, it is important to assess how its implementation will affect screening outcomes and cost-effectiveness. A model study from the United States (US) showed that adding annual DBT screening to DM is cost-effective compared with DM screening only, whereas a more recent study concluded that DBT screening is not cost-effective at current reimbursement rates in the United States[8,9]. These conflicting results may occur due to the unavailability of long term DBT screening data, leading to uncertainty in model input parameters. Microsimulation models can be used to predict long-term screening outcomes and account for all uncertainty in input parameters through probabilistic sensitivity analysis (PSA). These models recalculate the output multiple times by incorporating probability distributions for input parameters and sampling from those distributions using a large number of runs. From PSA outcomes, the generated cost-effectiveness acceptability curves show the probability of a strategy being cost-effective, given a particular willingness-to-pay-threshold (WTP)[10].

Currently, to our knowledge, no studies have conducted a PSA to assess the uncertainty associated with the cost-effectiveness of biennial DBT screening in a population-based breast cancer screening program. We used the MIcrosimulation SCreening ANalysis (MISCAN) model to predict long-term benefit, harm and costs with DBT versus DM in the Netherlands. Performing a PSA, we estimated

the mean incremental cost-effectiveness of DBT versus DM and the probability of DBT being the more cost-effective option, for different WTPs.

METHODS

Model overview

We used the MISCAN model, described previously[11]. MISCAN simulates individual life histories and the natural history of breast cancer as progression through preclinical screen-detectable (either preceded by ductal carcinoma in situ (DCIS) or not) T1A, T1B, T1C and T2+ tumors. When screening is modeled, tumors can be detected in the preclinical phase. MISCAN models actual DM screening participation and therapy use[11]. An overview of the model parameters is shown in the Supplementary material (Supplementary material 1). MISCAN is exempt from institutional review board approval.

This study was performed from March 2018 to February 2019. Given the scarcity of DBT data at this time, we used expert elicitation to obtain estimates for three input parameters and their distributions: test sensitivity, the PPV and screening costs with DBT, following guidelines[12, 13]. Experts were asked to give estimates for a situation with stand-alone DBT. Expert elicitation is valid to use early in the development of new techniques, when input parameters are not available from literature, scarce or when published values are inconsistent[12, 14]. Invited experts were researchers with extensive experience in breast cancer screening, 23 years on average, who were involved in DBT trials or affiliated with the Dutch Expert Centre for Screening. One expert, a Professor of radiology with 20 years of experience in the field, is an author of this study. Elicitation sessions were conducted using the EXPLICIT tool[14]. Nine of the thirteen experts invited responded.

During the expert elicitation, two studies published sensitivity with DBT, using interval cancers from the first screening round: the Oslo Tomosynthesis Screening Trial (OTST) and the Screening with Tomosynthesis Or standard Mammography (STORM) trial[1, 7]. Estimates for the increase in sensitivity with DBT were 6% and 11%, respectively, although the difference between DBT and DM was not significant

for either study. We decided not to dismiss the expert estimates but to pool them with the trial results. We did not only use trial results because: 1. The lack of long term data (only one screening round); 2. One trial[7] had a very low number of interval cancers with DBT, making the estimate less reliable; 3. The use of combined DM and DBT was compared to DM alone. Combined DM and DBT considerably increases the radiation dose, acquisition- and reading time and it may thus not be ideal to implement this combination in a screening program[15-19]. Reconstructing a two-dimensional image from DBT data seems a good alternative[4-6, 20].

The PPV of recall and screening costs with DBT were based on expert opinion only as synthesis of the literature was difficult. Published PPV estimates vary widely, making it difficult to yield an estimate[1, 3, 5, 6]. Screening costs with DBT are also uncertain. Increased reading time (reading the stack of images and synthetic 2D) will probably increase costs the most, compared with DM, but studies on plausible DBT reading strategies are scarce.

Assumptions

We assumed no difference in treatment rates between DM and DBT. For DM, we assumed that all women undergo diagnostic mammography after recall. As this is less likely after DBT, only 30% of the women would undergo mammography after recall from DBT[21]. Some studies show that the biopsy rate over multiple screening rounds is similar for DM and DBT[22], so we used the same frequencies and costs for other diagnostic tests. Finally, we assumed that switching from DM to DBT increases sensitivity.

Costs

Costs are reported in Euros and United States Dollars (USD) in parentheses (1 Euro=1.10 USD; May 19, 2020). Screening costs are based on data from the Dutch Screening Organizations[23] (DM) or expert opinion (DBT). Costs of breast cancer diagnosis after recall and treatment are from a published Dutch study (Supplementary material 2)[24]. Total costs with screening are presented as net costs, compared with a situation without screening. All costs and effects are discounted at 3.5% per year.

Analysis

Two cohorts, each consisting of 500,000 women born in 1970, were simulated and followed over their lifetime, starting at 50 years. Cohort 1 was invited to biennial DM screening between 50-74 years (current practice). Cohort 2 was invited to biennial DBT between 50-74 years, assuming DBT would fully replace DM. For both scenarios, we assumed an attendance rate of 80%, reflecting actual participation over 2013-2018 in the Netherlands[23].

Benefits and harms

Undiscounted benefits and harms were based on a single run and included LYG, breast cancer deaths averted, and false-positives - i.e. screening examinations resulting in further diagnostics but not leading to breast cancer diagnosis, calculated using the PPV of recall. The PPV was defined as the percentage of breast cancers detected, among all recalls.

Probabilistic sensitivity analysis (PSA)

Input parameters were varied simultaneously in 10,000 runs, sampling from fixed probability distributions, generating 10,000 incremental cost-effect pairs for DBT versus DM (Table 1a-b). In each run, one million women were simulated. In MISCAN, model parameters are calibrated with the Nelder-Mead algorithm[25], which does not provide distributions around calibrated values. But we could not obtain uncertainty measures from the literature for most calibrated parameters. To account for uncertainty, we used a previously described approach (Supplementary material 3)[26, 27].

For parameters not calibrated, distributions were chosen based on: 1) unobservable parameters have a uniform distribution and vary 5% around the mode (fixed value); 2) parameters representing costs or PPVs have a beta-PERT (Program Evaluation and Review Technique) distribution and vary10% ; and 3) bounds for DBT parameters, based on expert opinion, have a beta-PERT distribution[14]. The beta-PERT distribution is a version of the beta distribution (interval [0, 1]) defined by the minimum, most likely and maximum value of a parameter and was used because expert estimates were defined by the lowest, most likely and highest

parameter value. Individual expert estimates were combined through unweighted linear pooling by averaging the lowest, highest and most likely values, generating an average most likely value and distribution[13]. We assumed that sensitivity of DBT is higher than that of DM[1, 7]. Therefore, we induced correlation between sensitivity of DBT and DM, after sampling, using the method of Goldhaber-Fiebert and Jalal[28]. We adopted a tolerance threshold of 5%, to allow for some ordering violation to occur.

PSA outcomes were mean incremental costs and effects of DBT versus DM, the mean incremental cost-effectiveness ratio (ICER) and mean incremental net benefit (INB). The ICER was calculated by dividing the difference in costs between DBT and DM by the difference in LYG. A screening scenario was considered cost-effective if the ICER was below the willingness-to-pay (WTP) threshold. The common Dutch WTP is €20,000 ($22,000) per (quality adjusted) LYG[29], generally used with 4% annual discounting of costs and 1.5% of effects. But we used 3.5% discounting of both costs and effects because this is recommended by the National Institute of health and Care Excellence[30], which uses a higher threshold range of £20,000-£30,000 per LYG; thus a maximum of around €35,000 ($38,500) per LYG. We also used this threshold to account for a more internationally representative WTP and because the WTP of €20,000 per LYG has been argued to be low.

The mean INB of DBT vs. DM was calculated as the difference between the expected net benefit (ENB) of DBT and DM, with:

ENB:WTP*LYG – Cost

INB:$ENB_{DBT} - ENB_{DM}$

Cost-effectiveness acceptability curves were calculated selecting the screening modality with the highest ENB in each of the 10,000 runs. Subsequently, we determined the proportion of runs for which the modality has the highest net benefit. The curves for DBT and DM show the probability of each being cost effective, given a particular WTP[10].

Table 1a. Input parameter values, distributions types and boundaries for probabilistic sensitivity analysis (PSA)

Calibrated input parameters	Distribution	Calibrated value	Minimum	Maximum
Stage-specific sensitivity of digital mammography DCIS	lognormal	0.87	0.69	1
Stage-specific sensitivity of digital mammography T1A	lognormal	0.55	0.43	0.70
Stage-specific sensitivity of digital mammography T1B	lognormal	0.48	0.38	0.61
Stage-specific sensitivity of digital mammography T1C	lognormal	0.86	0.66	1
Stage-specific sensitivity of digital mammography T2+	lognormal	1	0.76	1
Breast cancer onset	beta	0.291	0.279	0.305
Onset hazard age 30 years	lognormal	7.049E-05	3.222E-05	1.348E-04
Onset hazard age 50 years	lognormal	0.010	0.007	0.014
Onset hazard age 70 years	lognormal	0.019	0.014	0.026
Onset hazard age 100 years	lognormal	0.024	0.018	0.032
Stage-specific duration† (years) screen-detectable preclinical stage DCIS	lognormal	3.5	2.85	4.37
Stage-specific duration† (years) screen-detectable preclinical stage T1A	lognormal	0.1	0.08	0.12
Stage-specific duration† (years) screen-detectable preclinical stage T1B	lognormal	0.5	0.39	0.63
Stage-specific duration† (years) screen-detectable preclinical stage T1C	lognormal	1.5	1.21	1.85
Stage-specific duration† (years) screen-detectable preclinical stage T2+	lognormal	0.9	0.70	1.16
Probability of DCIS progressing into next stage (T1A) immediately	beta	0.842	0.415	0.996
Probability of DCIS progressing into next stage (T1A) slowly	beta	0.038	0.01	0.09

Table 1a. (Continued)

Non-calibrated input parameters	Distribution	Mode (fixed value)	Minimum	Maximum
Relative sensitivity‡ of younger ages digital mammography	beta-PERT	0.75	0.60	0.90
Relative sensitivity‡ of younger ages tomosynthesis	beta-PERT	0.85	0.70	1
Factor to vary period-specific increase in onset of breast cancer	uniform	original* value	10% lower	10% higher
Factor to vary stage- and age-specific probability of transition node negative to positive	uniform	original value	5% lower	5% higher
Factor to vary stage- and age-specific probability of diagnosis	uniform	original value	5% lower	5% higher
Relative improvement in prognosis screen- vs. clinical detection <70 years	uniform	original value	10% lower	10% higher
Relative improvement in prognosis screen- vs. clinical detection >70 years	uniform	original value	10% lower	10% higher
Costs late stage breast cancer	beta-PERT	€18,000 ($19,800)	€16,200 ($17,820)	€19,800 ($21,780)
Stage-specific treatment costs DCIS	beta-PERT	€4,569 ($5,026)	€4,112 ($4,523)	€5,026 ($5,529)
Stage-specific treatment costs T1A, N-	beta-PERT	€4,333 ($4,766)	€3,900 ($4,290)	€4,766 ($5,243)
Stage-specific treatment costs T1B, N-	beta-PERT	€5,057 ($5,563)	€4,551 ($5,006)	€5,563 ($6,119)
Stage-specific treatment costs T1C, N-	beta-PERT	€11,146 ($12,261)	€10,031 ($11,034)	€12,261 ($13,487)

Table 1a. (Continued)

Non-calibrated input parameters	Distribution	Mode (fixed value)	Minimum	Maximum
Stage-specific treatment costs T2+, N-	beta-PERT	€10,815 ($11,897)	€9,734 ($10,707)	€11,897 ($13,087)
Stage-specific treatment costs T1A, N+	beta-PERT	€16,103 ($17,713)	€14,493 ($15,942)	€17,713 ($19,484)
Stage-specific treatment costs T1B, N+	beta-PERT	€6,744 ($7,418)	€6,070 ($6,677)	€7,418 ($8,160)
Stage-specific treatment costs T1C, N+	beta-PERT	€20,822 ($22,904)	€18,740 ($20,614)	€22,904 ($25,194)
Stage-specific treatment costs T2+, N+	beta-PERT	€15,063 ($16,569)	€13,557 ($14,913)	€16,569 ($18,226)
Positive predictive value of recall with digital mammography screening	beta-PERT	28%	27%	29%
Positive predictive value of biopsy	beta-PERT	70%	64%	75%
Positive predictive value of recall by a general practitioner	beta-PERT	58.3%	52.5%	64%

†Given values for the stage-specific durations represent durations for women aged 50 years (in the model different durations for women aged 40, 50, 60 and 70 years old are used); ‡The relative sensitivity is 1 for age 55 and older, 0.85 for <50 and interpolated between 0.85 and 1 for 50-55;
*Original value differs by stage and age;
Abbreviations: DCIS (ductal carcinoma in situ); PERT (program evaluation and review technique)

Table 1b. Input parameter values, distributions types and boundaries for probabilistic sensitivity analysis (PSA) for parameters based on expert opinion

Input parameters based on expert opinion	Distribution	Mode (fixed value)	Minimum	Maximum
Stage-specific sensitivity of tomosynthesis DCIS	beta-PERT	1	0.92	1
Stage-specific sensitivity of tomosynthesis T1A	beta-PERT	0.65	0.58	0.72
Stage-specific sensitivity of tomosynthesis T1B	beta-PERT	0.57	0.51	0.63
Stage-specific sensitivity of tomosynthesis T1C§	beta-PERT	1	0.91	1
Positive predictive value of recall with tomosynthesis screening	beta-PERT	30%	23%	37%
Screening costs tomosynthesis	beta-PERT	€91 ($100)	€72 ($79)	€125 ($138)

§T2+ was not varied in the probabilistic sensitivity analysis as the lower limit of the distribution was already 1 (100%)
Abbreviations: DCIS (ductal carcinoma in situ); PERT (program evaluation and review technique)

Scenario Analyses

To account for different DBT scenarios we performed one-way scenario analyses, based on a single run:

-The percentage diagnostic mammography after recall from DBT (30% in base case) was varied between 0%-100%
-Biopsy after recall is 38% higher with DBT than with DM[21, 31].
-Attendance with DBT varies between 70%-90% but was kept constant with DM as this was based on observed rates

-PPV of recall with DBT is 50% higher than with DM[32].
-DBT parameters were varied using the boundaries from the distribution obtained by expert opinion.

Statistical analysis

MISCAN was used for the cost-effectiveness analysis with 10,000 iterations. The Sheffield Accelerated Value of Information (SAVI) tool[33] and the Bayesian cost-effectiveness analysis (BCEA) package[34] were used to post-process the results of the 10,000 runs, to obtain mean incremental costs and effects with confidence intervals and the ICER and to assess uncertainty constructing cost-effectiveness planes and acceptability curves. We used R version 3.4.1 (2017) for the BCEA package and SAVI 2.1.2 (R Shiny Server application)[33, 34].

RESULTS

Expert opinion

Based on the pooled results from the expert elicitation and the trials, test sensitivity was estimated to increase by 18% when switching from DM to DBT screening (distribution range: 6%-31%) (Table 2)[1, 7]. The pooled expert estimates were a PPV with DBT of 30% (23% - 37%) and DBT screening costs of €91 ($100) (€72 ($79) - €125 (138)), respectively.

Undiscounted benefits and harms

Biennial DBT screening between 50-74 years, compared with biennial DM, was expected to result in 13 additional LYG (7% increase; 13/193) and 0.9 additional breast cancer deaths averted (6% increase; 0.9/14.5) per 1000 women invited to screening followed over their lifetime (Table 3). There were 4 fewer false-positive findings predicted per 1000 women invited to screening with DBT (2.3% decrease; 4/159) compared with DM.

Table 2. Individual estimates of nine experts and two trials and unweighted linearly pooled value

	Percentage increase sensitivity			**PPV DBT** (%)			**Screening costs DBT** (€)		
	most likely	lowest	highest	most likely	lowest	highest	most likely	lowest	highest
Expert 1	50	5	80	30	23	45	95	70	130
Expert 2	10	5	25	32	22	38	95	75	105
Expert 3	15	5	25	30	25	35	130	70	200
Expert 4	15	5	30	35	30	45	90	80	150
Expert 5	8	4	15	30	26	35	80	70	90
Expert 6	13	4	23	24	15	28	85	75	150
Expert 7	30	10	40	28	20	35	75	67	90
Expert 8	20	0	50	30	25	35	75	70	85
Expert 9	22	15	25	30	25	35	-	-	-
OTST	5	1	11	-	-	-	-	-	-
STORM	11	7	12	-	-	-	-	-	-
Pooled value	**18**	**6**	**31**	**30**	**23**	**37**	**91** ($100)	**72** ($79)	**125** ($138)

Pooled values reflect the mean of all estimates; Empty cells indicate that an estimate was not given or, in case of the trials, that these data were not used.
N.B. Values for 'percentage increase sensitivity' are based on estimates of nine experts and the estimates of the Oslo Tomosynthesis Screening Trial (OTST)[15] ref and Screening with Tomosynthesis OR standard Mammography STORM[16] trial
Abbreviations: PPV (positive predictive value); DBT (digital breast tomosynthesis)

Probabilistic sensitivity analysis (PSA)

Based on 10,000 model runs, the mean lifetime incremental costs of DBT versus DM were €137,555 ($151,311) per 1000 women invited (95% confidence interval:€31,093 ($34,202)-€263,537 ($289,891)) and the mean incremental discounted lifetime effects 5.09 LYG per 1000 women invited (95% confidence interval:-0.80-9.70). The mean ICER of DBT versus DM was estimated at €27,023 ($29,725) per LYG (Table 3). The uncertainty around the ICER estimate is represented by the cost-

effectiveness plane (Figure 1a and b). The majority of the incremental cost-effect pairs of DBT, relative to DM, fell within the northeast quadrant of the cost-effectiveness plane. But a small part occurred within the northwest quadrant. This indicates that DBT is more costly than DM and for a high proportion of the iterations more effective (94.7%).

Table 3. Long-term screening outcomes and results probabilistic sensitivity analysis per 1000 women invited to screening

Screening outcomes (**single run**)	DM	DBT	Incremental‡
Life-years gained†	179.9	192.5	12.6
Breast cancer deaths averted†	13.6	14.5	0.9
False-positive findings	158.8	155.2	-3.6
PSA outcomes (**3.5% discounted**)			
Mean costs†	€305,829 ($336,412)	€443,385 ($487,724)	€137,555 ($151,311)
Mean effect (life-years gained)†	67	73	5
Mean ICER (DBT vs DM)	-	-	€27,023 ($29,725)
ENB (WTP = €20,000)	1,043,980	1,008,231	-35,749
ENB (WTP = €35,000)	2,056,337	2,096,943	40,606

†Relative to no screening; ‡DBT relative to DM;
N.B. Women were invited to biennial screening between 50-74 years and followed over their lifetimes
Abbreviations: PSA (probabilistic sensitivity analysis); DM (digital mammography); DBT (digital breast tomosynthesis); ICER (incremental cost-effectiveness ratio); ENB (expected net benefit); WTP (willingness-to-pay-threshold); INB (incremental net benefit)

Figure 2 shows the cost-effectiveness acceptability curves for DBT and DM (Figure 2). At a WTP of €20,000 ($22,000) per LYG, DBT had a low probability of being (most) cost effective (0.36) and a negative INB of -35,749 compared with DM (Table 3). At €35,000 ($38,500) per LYG, the INB was positive (40,606) and the

probability of cost-effectives 0.66, which means that DBT had the highest ENB in 66% of the 10,000 iterations (Table 4; Figure 2).

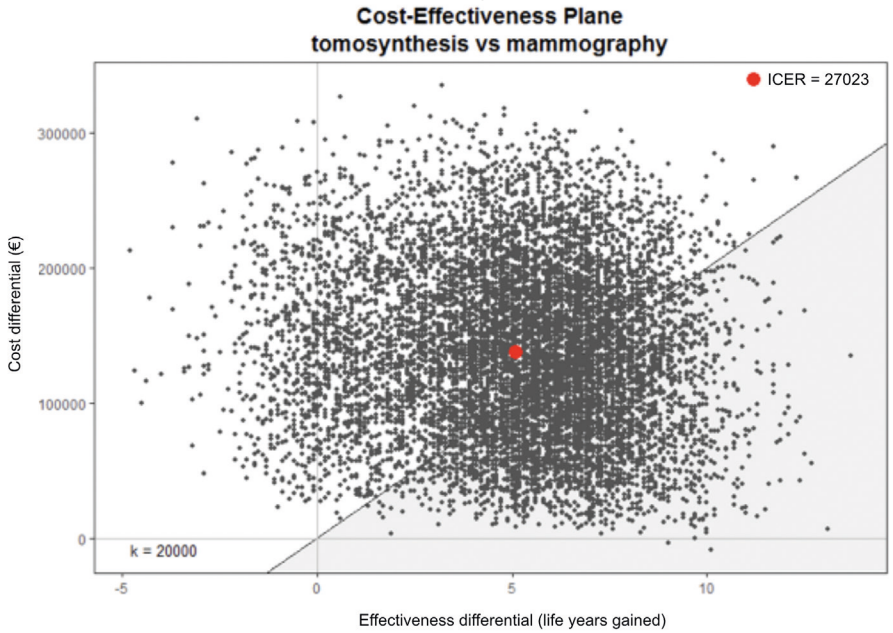
One-way scenario analyses

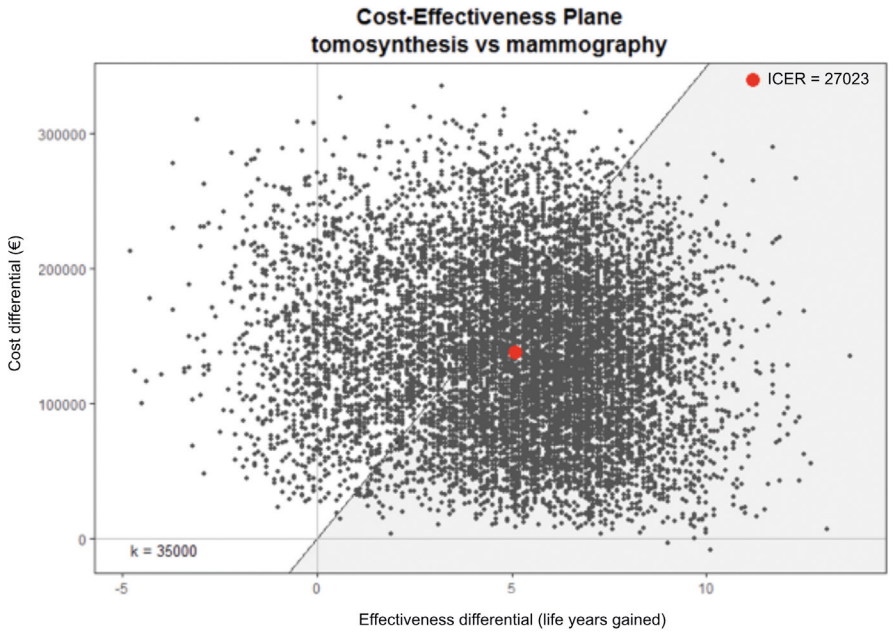
The ICER only ranged between €27,281 and €30,710 per LYG ($30,009-$33,782) when varying the percentage of diagnostic mammography after recall from DBT, the biopsy rate with DBT or the PPV of DBT (Supplementary material 4). Ranging attendance with DBT, resulted in considerable changes in the ICER (€51,898 ($57,088) per LYG with 70% attendance;€19,722 ($21,694) with 90% attendance). Substantial ranges in the ICER of €1,276-€76,685 ($1,404-$84,353) and €22,841-€96,745 ($25,125-$106,419) were found when varying screening costs between €72 and €125 ($79-$138) and the increase in sensitivity with DBT between 31% and 6%, respectively (S4).

DISCUSSION

As only short term outcomes of DBT screening have been published, we predicted long term costs and effects of DBT in a population-based screening setting, performing a PSA to account for the uncertainty in model parameters. Using current data, biennial DBT screening between 50-74 years is expected not to be cost effective, considering the Dutch WTP of €20,000 ($22,000) per LYG and 3.5% annual discounting. But Dutch discounting recommendations (1.5% annually for effects;4% for costs) would likely change this conclusion[12]. Internationally, a threshold of €20.000 per LYG is rather low, also putting the conclusion in a different perspective. Switching to biennial DBT in the Netherlands is likely to be cost-effective when using a higher threshold of €35,000 ($38,500) per LYG, as DBT was most cost effective in 66% of the iterations with this threshold. When converting our findings to U.S. dollars and using a WTP of $100,000, biennial screening with DBT between age 50 and 74 would be cost effective.

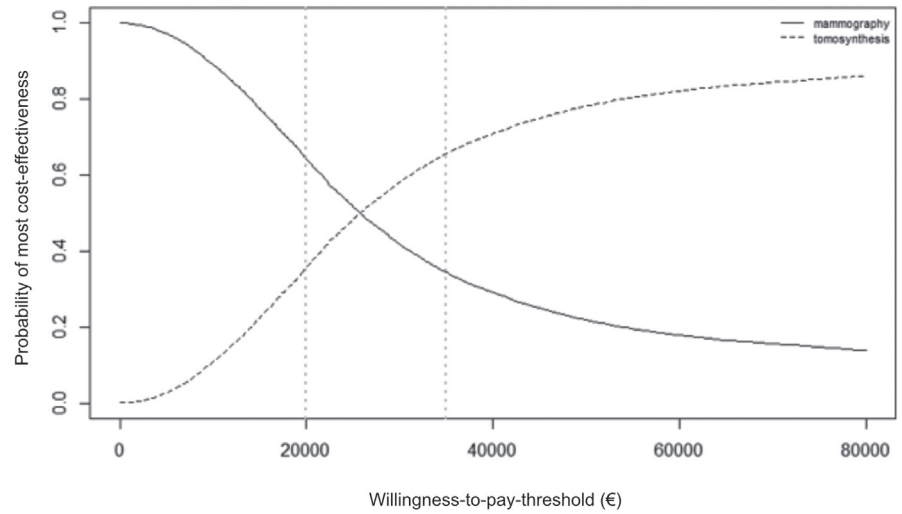
Figure 1a and b. Cost-effectiveness plane of tomosynthesis vs. mammography at a willingness-to-pay-threshold of €20,000 (a; $22,000) and €35,000 (b; $38,500) per LYG (costs and effects per 1000 women invited to screening)





The grey dots show the simulated values for the joint distribution of the effectiveness and cost differentials; The red dot indicates the average of the distribution of the outcomes (ICER); The grey area reflects the part of the plane for which the simulated values are below the willingness-to-pay-threshold; k: willingness-to-pay-threshold
Abbreviations: LYG (life years gained); ICER (incremental cost-effectiveness ratio)

Figure 2. Cost-effectiveness acceptability curves for DBT and DM



Dashed vertical lines represent the WTPs of €20,000 ($22,000) and €35,000 ($38,500) per life year gained
Abbreviations: DBT (digital breast tomosynthesis); DM (digital mammography)

In agreement with a study of Lowry et al., we found that ICERs were most sensitive to costs of DBT[9]. Our results suggest that, at a WTP of €35,000 per LYG, biennial DBT between 50-74 years is more cost-effective than DM in the majority of runs. Using a higher WTP, of 100,000 USD, Kalra et al. showed that in the US annual DBT starting at age 40 years is also more cost-effective than DM in the majority of runs[8].

Some DBT trials and observational studies show lower false-positive rates with use of DBT, compared with DM, whereas others suggest higher rates[4, 15, 35, 36]. In contrast to a study from the United States, which demonstrated a 24-28% reduction

in false-positives, we only found a slight decrease as we assumed a minor increase in the PPV with DBT, relative to DM[9]. Recall- and false-positive rates with DM in the Netherlands are relatively low compared with other countries, thus a reduction when switching to DBT may be less likely than in other settings. Correspondingly, a meta-analysis evaluating recall rates with DBT compared with DM showed that reductions in recall rates with DBT were predominantly found in studies from the United States that reported high baseline recall[37]. The decrease in recall rates and false-positives with DBT thus depends on the initial rate with DM.

Prospective studies show that DBT increases breast cancer detection, compared with DM[1-3, 5], which could contribute to overdiagnosis of breast cancer. Recent literature shows that DBT detects a higher proportion of invasive cancers than DM but it may be too soon to assess the effect on interval cancer rates[38, 39].

Our findings show the impact of DBT, compared to DM, in a European screening setting, supplementing earlier published outcomes for the screening setting in the United States[8, 9].

Our study had some limitations. First, as trial estimates of DBT sensitivity were based on interval cancers of the first screening round only[1, 7], and the performance of combined DBT and DM was compared to DM alone[1, 2], the actual effect of DBT alone could be smaller than reported. We did not account for increased radiation dose with DBT compared with DM. We also did not specify our results for different breast density levels as there was not sufficient evidence from the trials that DBT performs better for dense breasts than for non-dense breasts compared with DM[2]. Furthermore, we assumed no differences in frequency and costs of diagnostics after recall between DBT and DM. But our scenario analyses showed that the diagnostic mammography and biopsy rate hardly influenced the ICER. Finally, we relied on expert opinion for some DBT inputs, which may have induced bias. Still, our inputs fairly represent the latest literature. The OTST recently reported a 28% increase in sensitivity with combined DBT and synthetic 2D, compared with DM[32]. In addition, an Italian screening pilot published an increase of 11%[31]. When averaging these new estimates with the earlier published 11% increase from the

STORM trial[7], the average increase in sensitivity with DBT, compared with DM, is 17%. This is similar to the 18% increase we used for the PSA, based on the expert estimates and trial results at the time of our study. Our estimate for the PPV also falls within the range of published estimates[1, 5, 6, 35]. Furthermore, although DBT costs in the Dutch screening setting are unclear, costs of diagnostic DBT in university medical centers in the Netherlands (€94.58 ($104))[40] are similar to the €91 (($100) used in our analysis.

In conclusion, when analyzing currently available preliminary data, based on one screening round with DBT, switching from biennial DM to DBT screening in the Netherlands is not cost-effective using a WTP of €20,000 ($22,000) per LYG. Switching to DBT screening could be cost-effective at a threshold of €35,000 ($38,500) per LYG.

REFERENCES

1. Skaane P, Sebuodegard S, Bandos AI et al. Performance of breast cancer screening using digital breast tomosynthesis: results from the prospective population-based Oslo Tomosynthesis Screening Trial. Breast Cancer Res Treat 2018; 169: 489-496.
2. Ciatto S, Houssami N, Bernardi D et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. Lancet Oncol 2013; 14: 583-589.
3. Lang K, Andersson I, Rosso A et al. Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmo Breast Tomosynthesis Screening Trial, a population-based study. Eur Radiol 2016; 26: 184-190.
4. Bernardi D, Macaskill P, Pellegrini M et al. Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study. Lancet Oncol 2016; 17: 1105-1113.
5. Hofvind S, Hovda T, Holen AS et al. Digital Breast Tomosynthesis and Synthetic 2D Mammography versus Digital Mammography: Evaluation in a Population-based Screening Program. Radiology 2018; 287: 787-794.
6. Caumo F, Zorzi M, Brunelli S et al. Digital Breast Tomosynthesis with Synthesized Two-Dimensional Images versus Full-Field Digital Mammography for Population Screening: Outcomes from the Verona Screening Program. Radiology 2018; 287: 37-46.
7. Houssami N, Bernardi D, Caumo F et al. Interval breast cancers in the 'screening with tomosynthesis or standard mammography' (STORM) population-based trial. Breast 2018; 38: 150-153.
8. Kalra VB, Wu X, Haas BM et al. Cost-Effectiveness of Tomosynthesis in Annual Screening Mammography. AJR Am J Roentgenol 2016; 207: 1152-1155.
9. Lowry KP, Trentham-Dietz A, Schechter CB et al. Long-term Outcomes and Cost-effectiveness of Breast Cancer Screening with Digital Breast Tomosynthesis in the United States. J Natl Cancer Inst 2019.
10. Barton GR, Briggs AH, Fenwick EA. Optimal cost-effectiveness decisions: the role of the cost-effectiveness acceptability curve (CEAC), the cost-effectiveness acceptability frontier (CEAF), and the expected value of perfection information (EVPI). Value Health 2008; 11: 886-897.
11. Sankatsing VD, Heijnsdijk EA, van Luijt PA et al. Cost-effectiveness of digital mammography screening before the age of 50 in The Netherlands. Int J Cancer 2015; 137: 1990-1999.

12. Bojke L, Grigore B, Jankovic D et al. Informing Reimbursement Decisions Using Cost-Effectiveness Modelling: A Guide to the Process of Generating Elicited Priors to Capture Model Uncertainties. Pharmacoeconomics 2017; 35: 867-877.
13. Iglesias CP, Thompson A, Rogowski WH, Payne K. Reporting Guidelines for the Use of Expert Judgement in Model-Based Economic Evaluations. Pharmacoeconomics 2016; 34: 1161-1172.
14. Grigore B, Peters J, Hyde C, Stein K. EXPLICIT: a feasibility study of remote expert elicitation in health technology assessment. BMC Med Inform Decis Mak 2017; 17: 131.
15. Skaane P, Bandos AI, Gullien R et al. Prospective trial comparing full-field digital mammography (FFDM) versus combined FFDM and tomosynthesis in a population-based screening programme using independent double reading with arbitration. Eur Radiol 2013; 23: 2061-2071.
16. Skaane P, Bandos AI, Gullien R et al. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. Radiology 2013; 267: 47-56.
17. Svahn TM, Houssami N, Sechopoulos I, Mattsson S. Review of radiation dose estimates in digital breast tomosynthesis relative to those in two-view full-field digital mammography. Breast 2015; 24: 93-99.
18. Tagliafico AS, Calabrese M, Bignotti B et al. Accuracy and reading time for six strategies using digital breast tomosynthesis in women with mammographically negative dense breasts. Eur Radiol 2017; 27: 5179-5184.
19. Bernardi D, Ciatto S, Pellegrini M et al. Application of breast tomosynthesis in screening: incremental effect on mammography acquisition and reading time. Br J Radiol 2012; 85: e1174-1178.
20. Skaane P, Bandos AI, Eben EB et al. Two-view digital breast tomosynthesis screening with synthetically reconstructed projection images: comparison with digital breast tomosynthesis with full-field digital mammographic images. Radiology 2014; 271: 655-663.
21. Alsheik NH, Dabbous F, Pohlman SK et al. Comparison of Resource Utilization and Clinical Outcomes Following Screening with Digital Breast Tomosynthesis Versus Digital Mammography: Findings From a Learning Health System. Acad Radiol 2018.
22. McDonald ES, Oustimov A, Weinstein SP et al. Effectiveness of Digital Breast Tomosynthesis Compared With Digital Mammography: Outcomes Analysis From 3 Years of Breast Cancer Screening. JAMA Oncol 2016; 2: 737-743.
23. National Evaluation Team Breast cancer screening. National evaluation of breast cancer screening in the Netherlands 2004-2014. Fourteenth evaluation report. Erasmus MC University Medical Center Rotterdam, Radboud University Medical Center Nijmegen 2019.

24. Saadatmand S, Tilanus-Linthorst MM, Rutgers EJ et al. Cost-effectiveness of screening women with familial risk for breast cancer with magnetic resonance imaging. J Natl Cancer Inst 2013; 105: 1314-1321.

25. Nelder JA, and Mead, R. A simplex method for function minimization. The computer journal 1965; 7(4): 308-313.

26. Wever EM, Draisma G, Heijnsdijk EA et al. Prostate-specific antigen screening in the United States vs in the European Randomized Study of Screening for Prostate Cancer-Rotterdam. J Natl Cancer Inst 2010; 102: 352-355.

27. Ten Haaf K, van Rosmalen J, de Koning HJ. Lung cancer detectability by test, histology, stage, and gender: estimates from the NLST and the PLCO trials. Cancer Epidemiol Biomarkers Prev 2015; 24: 154-161.

28. Goldhaber-Fiebert JD, Jalal HJ. Some Health States Are Better Than Others: Using Health State Rank Order to Improve Probabilistic Analyses. Med Decis Making 2016; 36: 927-940.

29. Zwaap J. KS, van der Meijden C., Staal P., van der Heiden L. Rapport kosteneffectiviteit in de praktijk. Zorginstituut Nederland 2015.

30. Appleby J, Devlin N, Parkin D. NICE's cost effectiveness threshold. BMJ 2007; 335: 358-359.

31. Bernardi D, Gentilini MA, De Nisi M et al. Effect of implementing digital breast tomosynthesis (DBT) instead of mammography on population screening outcomes including interval cancer rates: Results of the Trento DBT pilot evaluation. Breast 2019.

32. Skaane P, Bandos AI, Niklason LT et al. Digital Mammography versus Digital Mammography Plus Tomosynthesis in Breast Cancer Screening: The Oslo Tomosynthesis Screening Trial. Radiology 2019; 291: 23-30.

33. Strong M, Oakley JE, Brennan A. Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: a nonparametric regression approach. Med Decis Making 2014; 34: 311-326.

34. Baio G. BA, Heath A. Bayesian Cost-Effectiveness Analysis with the R package BCEA. Springer International Publishing AG 2017.

35. Lang K, Nergarden M, Andersson I et al. False positives in breast cancer screening with one-view breast tomosynthesis: An analysis of findings leading to recall, work-up and biopsy rates in the Malmo Breast Tomosynthesis Screening Trial. Eur Radiol 2016; 26: 3899-3907.

36. Friedewald SM, Rafferty EA, Rose SL et al. Breast cancer screening using tomosynthesis in combination with digital mammography. JAMA 2014; 311: 2499-2507.

37. Marinovich ML, Hunter KE, Macaskill P, Houssami N. Breast Cancer Screening Using Tomosynthesis or Mammography: A Meta-analysis of Cancer Detection and Recall. J Natl Cancer Inst 2018; 110: 942-949.

38. Bahl M, Gaffney S, McCarthy AM et al. Breast Cancer Characteristics Associated with 2D Digital Mammography versus Digital Breast Tomosynthesis for Screening-detected and Interval Cancers. Radiology 2018; 287: 49-57.

39. Bernardi D, Gentilini MA, De Nisi M et al. Effect of implementing digital breast tomosynthesis (DBT) instead of mammography on population screening outcomes including interval cancer rates: Results of the Trento DBT pilot evaluation. Breast 2020; 50: 135-140.

40. https://cdn.fbto.nl/-/media/documenten/zorgverzekering-ziekenhuiskosten-ziekenhuis-2018.pdf. Accessed: May 2020

SUPPLEMENTARY MATERIAL

1. Model parameters on natural history of breast cancer, performance and survival after adjuvant treatment and screening

Annual increase in background breast cancer incidence (without screening) [a]	%
2008-2012	1.4
2013-2017	1.0
2018-2022	0.6
2023-2027	0.2
From 2028	0

Mean duration (years) of screen-detectable preclinical stage by age and stage[1, 2] [b]	Stage								
Age	DCIS	T1AN–	T1AN+	T1BN–	T1BN+	T1CN–	T1CN+	T2+N–	T2+N+
40	3.5	0.1	0.1	0.4	0.4	1.1	1.1	0.7	0.7
50	3.5	0.1	0.1	0.5	0.5	1.5	1.5	0.9	0.9
60	3.5	0.1	0.1	0.8	0.8	2.2	2.2	1.3	1.3
70	3.5	0.1	0.1	1.0	1.0	2.7	2.7	1.6	1.6

Test sensitivity of digital mammography by age and preclinical stage[1-3] [b]	Stage								
Age	DCIS	T1AN–	T1AN+	T1BN–	T1BN+	T1CN–	T1CN+	T2+N–	T2+N+
<50	65%	41%	41%	36%	36%	64%	64%	90%	90%
≥55	87%	55%	55%	48%	48%	86%	86%	100%	100%

Test sensitivity of digital breast tomosynthesis by age and preclinical stage[15-16] [a,c]	Stage								
Age	DCIS	T1AN–	T1AN+	T1BN–	T1BN+	T1CN–	T1CN+	T2+N–	T2+N+
<50	85%	55%	55%	48%	48%	85%	85%	85%	85%
≥55	100	65%	65%	57%	57%	100%	100%	100%	100%

Long-term relative survival by clinical stage and age, without adjuvant treatment[4-9] [c]	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[2, 10, 11] [b,c]	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[2, 10, 11 b,c]	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[2, 10, 11 b,c]	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[6, 12-14 c]	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+).
[a] based on assumptions, [b] data used for model fit (calibration), [c] data used directly as input

2. Costs of diagnostics and treatment and positive predictive values (not varied in probabilistic sensitivity analysis)

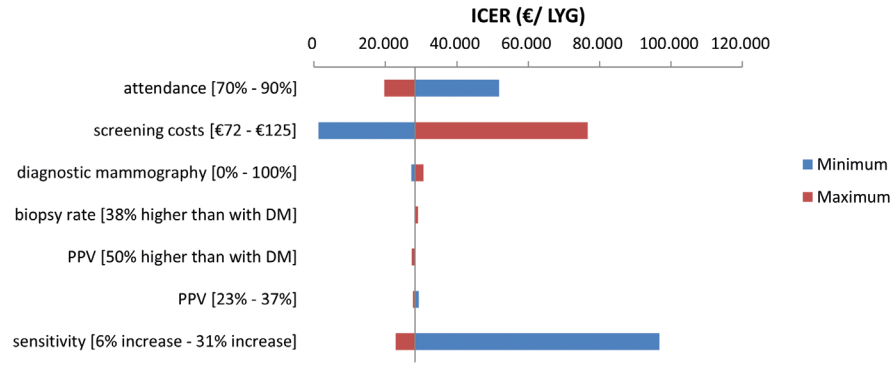
Diagnostics	**Costs**
invitation	€2
DM screening	€66.37
MRI	€367.59
palpation	€69.05
ultrasound	€77.08
FNA	€141.73
biopsy	€175.86
mean FNA/biopsy	€158.80
additional DM/DBT in hospital	€103.23

Costs of diagnostics and treatment were obtained from Saadatmand et al.[17]

3. Methods to obtain parameter bounds and distributions

In MISCAN, model parameters are calibrated with the Nelder-Mead algorithm[18,19] to minimize the likelihood-based deviance between modeled and observed values. As this algorithm does not provide measures of uncertainty around calibrated values, the sampling distributions needed to be determined empirically. For most model parameters, however, we could not obtain measures of uncertainty from the literature. In order to still be able to account for the uncertainty around input parameters, we used an approach similar to previously described methods.[20-22] With this approach, blocks of parameters were varied within given bounds (percentages), yielding new deviances, while the remaining parameters were kept constant at their calibrated values. The deviance increases as the parameters become more distant from their maximum likelihood estimates. As the deviance points form a U-shape, the approximation of the parameter bounds are obtained by fitting a second degree polynomial. Parameter values that yielded a 5% higher deviance than the calibrated value were chosen as lower and upper bounds. By varying parameters in a block simultaneously, the maximum allowed percentage change was the same for all parameters in a particular block. The parameter bounds were used to create appropriate probability distributions, which were chosen according to a previously described method by Briggs et al.[23] Beta distributions were assumed for parameters expressing probabilities and lognormal distributions for non-negative parameters. Distribution parameters were fitted by method of moments, using the calibrated parameter value as the mean and the standard deviation of the estimated uncertainty level.

4. Tornado plot of one-way scenario analyses

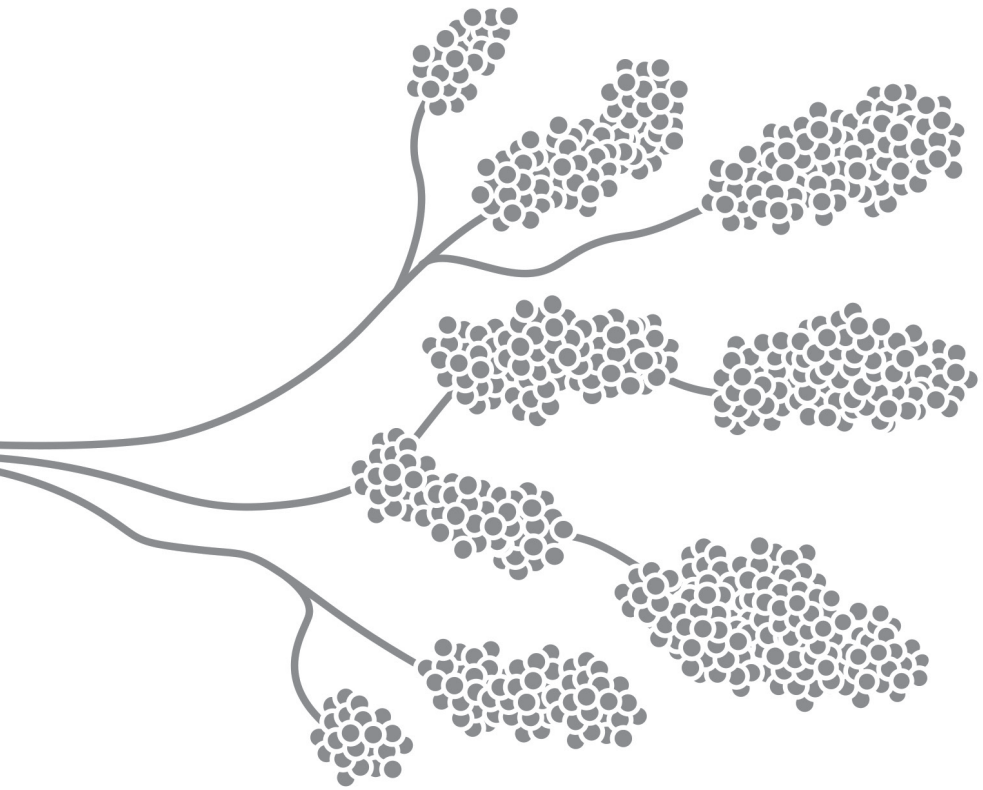


N.B. The outcomes for the different scenarios were based on a single run, because it would be too time consuming to run the scenarios for all 10,000 iterations. The ICER of 28,310 per LYG was the outcome of this single run; the mean ICER reported in the results section of the main text was based on the 10,000 runs.

References

1. National Evaluation Team for Breast cancer screening (NETB). National evaluation of breast cancer screening in the Netherlands 1990-2011/2012. Thirteenth evaluation report. Erasmus MC University Medical Center Rotterdam, Radboud University Nijmegen Medical Centre, 2014.
2. Nederlandse Kankerregistratie. http://www.cijfersoverkanker.nl/. (Accessed: December, 2015)
3. van Luijt PA, Fracheboud J, Heijnsdijk EA, den Heeten GJ, de Koning HJ, National Evaluation Team for Breast Cancer Screening in Netherlands Study G. Nation-wide data on screening performance during the transition to digital mammography: observations in 6 million screens. Eur J Cancer 2013;49:3517-25.
4. 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.
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. 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.
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. 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. Sant M, Allemani C, Capocaccia R, Hakulinen T, Aareleid T, Coebergh JW, Coleman MP, Grosclaude P, Martinez C, Bell J, Youngson J, Berrino F, et al. Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. Int J Cancer 2003;106:416-22.
10. 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.
11. Early Breast Cancer Trialists' Collaborative G. 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.

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. 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.
14. 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.
15. Skaane P, Sebuodegard S, Bandos AI et al. Performance of breast cancer screening using digital breast tomosynthesis: results from the prospective population-based Oslo Tomosynthesis Screening Trial. Breast Cancer Res Treat 2018; 169: 489-496.
16. Houssami N, Bernardi D, Caumo F et al. Interval breast cancers in the 'screening with tomosynthesis or standard mammography' (STORM) population-based trial. Breast 2018;38:150-153.
17. Saadatmand S, Tilanus-Linthorst MM, Rutgers EJ, Hoogerbrugge N, Oosterwijk JC, Tollenaar RA, Hooning M, Loo CE, Obdeijn IM, Heijnsdijk EA, de Koning HJ. Cost-effectiveness of screening women with familial risk for breast cancer with magnetic resonance imaging. J Natl Cancer Inst 2013;105:1314-21
18. Nelder JA, and Mead, R. A simplex method for function minimization. The computer journal 1965; 7(4): 308-313.
19. Tan SY, Van Oortmarssen, G. J. and Piersma, N. Estimating parameters of a microsimulation model for breast cancer screening using the score function method. Annals of Operations Research 2003; 119 (111-114): 143-161.
20. Wever EM, Draisma G, Heijnsdijk EA et al. Prostate-specific antigen screening in the United States vs in the European Randomized Study of Screening for Prostate Cancer-Rotterdam. J Natl Cancer Inst 2010; 102: 352-355.
21. Ten Haaf K, van Rosmalen J, de Koning HJ. Lung cancer detectability by test, histology, stage, and gender: estimates from the NLST and the PLCO trials. Cancer Epidemiol Biomarkers Prev 2015; 24: 154-161.
22. De Carvalho T.M. HEAM, Coffeng L., de Koning H.J. Evaluating parameter uncertainty in a simulation model of cancer using emulators. In press.
23. Briggs A, Sculpher, M. and Claxton, K. . Decision modelling for health economic evaluation. OUP Oxford 2006.



Chapter 7

General Discussion

MAIN FINDINGS

Part One: Evaluation of current breast cancer screening in the Netherlands

Screening is a form of prevention and is therefore offered to a 'healthy' population. Screening for breast cancer in asymptomatic women can lead to early diagnosis of tumours and, therefore, to earlier treatment of breast cancer than in a clinical setting. Early detection and treatment can result in improved survival and may avert cancer deaths, but screening can also yield adverse effects and cause harm. The majority of women attending screening will not benefit from it, whereas they are exposed to screening harms. It is therefore important to frequently evaluate the impact of screening programmes.

The Dutch breast cancer screening programme was initiated in 1990. We estimated important performance indicators for cancer screening over the period in which the transition from screen-film to digital mammography took place and we evaluated the effect of 20 years of mammography screening on breast cancer mortality in the Netherlands.

Performance indicators during the transition from screen-film to digital mammography screening

Analysing screening data from the years 2004 to 2011, we found that the breast cancer detection rate in the Netherlands significantly increased by more than 20%, for both in situ (DCIS) and invasive cancers. The interval cancer rate remained stable during this period. In line with the increased detection, the programme sensitivity, recall rate and false-positive rate also increased significantly. From 2008, digital mammography gradually replaced screen-film mammography in the Netherlands. Although the detection rate was significantly higher with digital mammography screening - a 15% increase compared to screen-film mammography - there was only a small non-significant reduction in interval cancer rates in a subgroup of younger women. As a result, the programme sensitivity was the highest for this subgroup.

We also found that the false-positive rate was higher with digital than with screen-film mammography, which is in line with results from other observational studies.[1] Although we found that breast cancer detection rates were higher with digital mammography than with screen-film mammography, there was no reduction in the interval cancer rate in the first years after digital mammography use. This may suggest that digital mammography adds to overdiagnosis of breast cancer. Other studies also showed similar interval cancer rates with digital mammography[1, 2], although fewer interval cancers with digital mammography were found to have microcalcifications, relative to screen-film mammography.[3] However, at the time of our study, we had limited follow-up data after the full transition to digital mammography, which made it difficult to assess the long-term effect of digital mammography on interval cancer rates.

A relatively new screening technique which may replace digital mammography in the future is digital breast tomosynthesis, which yields a pseudo 3D image by acquiring multiple images of the breast from different angles. At this stage, no statistically significant decreases in interval cancer rates have been reported for tomosynthesis compared to digital mammography.[4-6] These results were, however, based on 2-year follow-up after only the first screening round with tomosynthesis and the number of interval cancers in some of the studies were rather small.

Breast cancer mortality reduction in the Netherlands

Mammography screening was implemented gradually throughout the Netherlands between 1989 and 1997. Taking the time of implementation in municipalities as starting point, we analysed breast cancer mortality trends between 1980 and 2010. We found that, after the introduction of screening in the Netherlands, breast cancer mortality decreased by 30%. A similar reduction had been shown previously in other European studies.[7-10] Although near the end of the implementation period and for certain age groups adjuvant treatment was better and more available, we found that a statistically significant decline in breast cancer mortality was present for all subgroups, irrespective of age and time of implementation.

It is difficult to disentangle the separate contributions of screening and (adjuvant) treatment in reducing breast cancer mortality. While multiple trend studies suggest a significant role for screening in reducing breast cancer mortality[11], some argue that the effect of screening is negligible.[12] In contrast, a recent study analysing Dutch data showed that a single invitation to participate in mammography screening, in addition to current biennial screening, already leads to a reduction in breast cancer mortality.[13]

Models are able to predict the separate effects of screening and treatment on breast cancer mortality. A study that estimated the proportions of breast cancer mortality reduction attributable to screening and treatment in 2008 in the Netherlands using the MISCAN microsimulation model, showed that 47% of the reduction was associated with screening and 53% with treatment.[14] Another study found that for the United States in 2012, these contributions were 37% and 63% for screening and treatment respectively.[15] Inevitably, the improvement of adjuvant treatment affects the effectiveness of screening. If late stage treatment will improve sufficiently in terms of effectiveness and treatment burden, early detection of breast cancer by screening may become less essential.

Part Two: Quantifying the cost-effectiveness, harms and benefits of different screening strategies and screening modalities in the Netherlands using microsimulation modelling

There is wide consistency between breast cancer screening programmes in Europe with regard to the screening interval, the age group invited to screening and the screening test.[16] Most countries offer biennial mammography screening to women between age 50 and 69 or age 50 and 74 years. The evidence for the benefit of these screening strategies is most clear as the majority of trials targeted these age groups.[17]

In the Netherlands, biennial screening is offered to women aged 50 to 74 years. We explored the cost-effectiveness and the benefits and harms of extending screening to younger women, risk-based screening strategies and screening with a different modality (digital breast tomosynthesis).

Extending screening to women aged 40 to 49 years

Whether or not women should be screened for breast cancer before age 50 years has been a topic of debate for a long time. At the same time, the incidence of breast cancer in women younger than 50 years increases.[18] The impact of screening on breast cancer mortality may be different for women in their forties because of several factors associated with younger age, including a lower breast cancer incidence than in older women and a lower sensitivity of mammography due to denser breasts and a higher tumour growth rate.[19-21] Incorporating these differences in our model, we found that extending screening to include women aged 40-49 years is cost-effective, considering a cost-effectiveness threshold of €20,000 per QALY gained. This has been confirmed by later work.[22] We simulated biennial strategies and hybrid strategies, which consisted of annual screening before age 50 years and biennial screening between age 50 and 74 years. In particular biennial strategies were efficient and thus had greater gains in life years per unit of cost than annual alternatives. In general, the literature shows that annual screening, for women at average risk of breast cancer, is not cost effective and considerably more harmful than biennial screening.[23-25] The European Commission Initiative on Breast Cancer (ECIBC) recommends against annual screening for women aged 50-74 years as they consider the evidence for annual screening to be of very low certainty.[26]

Overall, we found that screening before age 50 years increased the number life years gained but only moderately lowered breast cancer mortality. When screening biennially from age 40 instead of age 50 years, the breast cancer mortality reduction was 7% higher. Extending screening will inevitably lead to a higher absolute number of harms. We found that the number of false-positive findings increases considerably with increasing the number of screens before age 50 years. When starting biennial screening at age 40 instead of age 50 years, false-positive findings were predicted to increase by 60%. Screening in younger women leads to relatively more false-positive findings than in older women as the positive predictive value of recall is lower.[27, 28] Overdiagnosis – i.e. screen-detection of tumours that would not have been detected during a woman's lifetime in the absence of screening -

only slightly increased. Overdiagnosis is higher among older women than younger women, due to more competing causes of death at older age.

Our results indicate that starting screening earlier, between age 40 and 49 years, is cost-effective and leads to a modest increase in benefits at the expense of a substantial increase in false-positive findings. Different people may have different viewpoints on whether the modest increase in benefits outweigh the increase in harms, that come with additional screening before age 50 years.

Today, only a few European screening programmes offer screening to women who are younger than 50 years. The evidence for the effectiveness of screening in women between age 40 and 49 years is still debated. Trials that aimed to determine the effectiveness of mammography screening in younger women were methodologically limited because it is not ethically justified to withhold these women from being screened after age 50 years.[29-31] The International Agency for Research on Cancer therefore concluded that the evidence that mammography screening in women aged 40-49 years reduces breast cancer mortality is limited.[17] Although the benefit of mammography screening for women aged 40-49 is less conclusive than for older women, a small non-significant effect for this age group is supported by several studies.[30] In addition, an observational study, performed in a period in which adjuvant therapy was available (1986-2005), compared a screened cohort to women who were not screened and demonstrated a 26% statistically significant breast cancer mortality reduction for women aged 40-49.[32] We, therefore, think it was reasonable to assume some benefit of screening for women age 40-49 years in our analyses, even in the presence of adjuvant therapy. The European Commission Initiative on Breast Cancer (ECIBC) now recommends triennial or biennial screening for women aged 45-49 years.[26]

Tailoring breast cancer screening to a woman's individual risk

A woman's individual risk for breast cancer may influence the balance between benefits and harms of mammography screening. Well-known risk factors for breast cancer are a family history of breast cancer, a previous biopsy, high breast density and genetic risk factors including Single Nucleotide Polymorphisms (SNPs).[33-36]

With risk-stratified screening, women are offered a tailored strategy based on their level of breast cancer risk. At this moment, there are no European screening programmes that offer risk stratified screening. There are however multiple initiatives with regard to risk-based projects.[37-40]

A small part of the female population has a relatively high risk of breast cancer, whereas a larger group has a relatively low risk.[41] Using a model with three different risk groups (low relative risk; relative risk of 1; high relative risk), we identified optimal screening strategies for women with a low (relative risk of 0.75) and high (relative risk of 1.8) breast cancer risk. Optimal strategies were defined as strategies with the highest possible benefit, given that the ICER was similar to and did not exceed the ICER of current screening. Optimal strategies were thus not the ones with the highest absolute benefit. We found that optimal screening for women with a low relative risk consisted of triennial screening between age 50 and 71 years, which would lead to a 30% reduction in false-positive findings and costs and would improve the ratio of benefits and harms.

Biennial screening starting at age 40 instead of age 50 years and ceasing at age 74 was identified as the optimal screening strategy for women with a high relative risk. This strategy would lead to an increase in screening benefit, but at the same time to a relatively higher increase in false-positive findings, leading to a worsened harm-benefit ratio. In this study, risk-based strategies with an ICER higher than that of current screening were not a candidate for optimal screening. Considering a cost-effectiveness threshold of €20,000 per (quality adjusted) life year gained, as recommended by Dutch guidelines[42], would allow for more intensive strategies to be identified for risk-based screening.

To determine the overall impact of a risk-based screening programme - using the identified optimal strategies for women eligible for screening with a low and high relative risk and the current strategy for women at average risk – data on the distribution of risk groups among the Dutch female population are required.

A woman's risk of breast cancer can be (partially) identified using a questionnaire to collect data on risk factors, including a family history of breast cancer, previous biopsy and age at first childbirth, age at menarche or menopause. A study using questionnaires to assess women's preferences with regard to knowing their breast cancer risk showed that among women attending the breast cancer screening programme in the United Kingdom, the desire of women to know their risk of breast cancer is high, around 95%.[43] However, stratified screening will result in less intensive screening being offered to women at the lowest risk, which women may find difficult to understand and accept. Woman at the highest risk should be alerted to their high risk and informed about the fact that, even with more intense screening, there is still a risk of developing an interval cancer. Particularly for these women, it is important not to miss a screening round. The attendance rate of screening could decrease if there is low acceptance of risk-based screening, due to low compliance with the recommended strategy. It is therefore important that women are well informed on the harms and benefits associated with screening, related to different risk levels, before risk stratified screening is implemented.

Screening with digital breast tomosynthesis

Digital breast tomosynthesis is a fairly new imaging technique which generates multiple images of the breast at different angles, as opposed to a single image with digital mammography, creating a pseudo 3D image. There are several European studies that show an increased breast cancer detection rate with tomosynthesis screening, compared to digital mammography screening.[6] Tomosynthesis is therefore considered as a replacement for digital mammography in breast cancer screening programmes. In order to evaluate the impact of the use of tomosynthesis in the Dutch screening setting on the cost-effectiveness, harms and benefits of breast cancer screening, we performed a cost-effectiveness analysis including a probabilistic sensitivity analysis (PSA). The MISCAN model has a large number of parameters, which are informed based on the long running Dutch breast cancer screening programme and published studies, including randomised trials. Since tomosynthesis is not yet used as a screening tool in the Netherlands, there were no direct data from the Dutch screening setting to inform the model. Information from

the literature to inform model input parameters was also scarce. Although there are various (ongoing) tomosynthesis screening trials, only a few have reported the effect of tomosynthesis on interval cancer rates, which are required to calculate estimates for the test sensitivity of tomosynthesis. In addition, several studies published estimates for the PPV with tomosynthesis, but these vary widely.[4, 6, 44-47] Furthermore, all tomosynthesis trials are international and findings from other countries might not be directly applicable to the Dutch screening setting. As it was therefore difficult to obtain estimates for tomosynthesis parameters from the literature, we made use of expert opinion. The aim of the expert elicitation was not to get consensus on values for input parameters but rather to capture the full distribution, including the uncertainty, around the estimate. Using this for the PSA, we found that biennial screening with tomosynthesis between age 50 and 74 years leads to an incremental cost-effectiveness ratio of €27,023 per life year gained, compared to digital mammography screening. The probability of tomosynthesis being more cost effective than digital mammography in the Netherlands was 0.36 at a threshold of €20,000 and 0.66 at €35,000 per life year gained. For this analysis, we assumed an increased sensitivity for tomosynthesis compared to digital mammography, based on the estimates obtained from the expert elicitation. A significant increase in sensitivity has recently been confirmed by the latest publication of the Oslo Tomosynthesis Screening Trial[48] and other new studies.[49]

With respect to screening benefits and harms, tomosynthesis screening would lead to a 7% lifetime increase in the number of life years gained, compared to screening with digital mammography, a 6% increase in the number of breast cancer deaths averted and a 2% decrease in false-positive findings. Although expectations were raised that screening with tomosynthesis may decrease false-positive rates considerably compared to digital mammography, this was thus not what we found. This can probably be explained by the fact that the baseline false positive rate with digital mammography in the Netherlands is already relatively low in Europe and considerably lower than recall rates in the United States. A substantial decrease in false positive findings, when switching to tomosynthesis screening, may therefore not be likely. Supporting this, a recent meta-analysis on

cancer detection and recall with tomosynthesis showed that tomosynthesis recall rates are lower than recall rates with digital mammography in pooled US studies but not in European studies.[50] This suggests that initial recall rate with digital mammography is an important determinant for the extent of the reduction in recall rates with tomosynthesis.

Another expectation of tomosynthesis was that it would particularly perform better in women with dense breasts than digital mammography. This has, however, not (yet) been confirmed sufficiently by the outcomes of the prospective trials.[4, 5, 51] We did therefore not specify our input for different levels of breast density.

An important remaining question is how tomosynthesis use will add to overdiagnosis due to breast cancer screening. A recent publication on a tomosynthesis screening pilot in Italy showed that interval cancer rates did not significantly decline following the increase in cancer detection rate.[6] This may indicate that at least part of the additionally detected cancers by tomosynthesis reflect overdiagnosed cases. The authors noted however that the size of the pilot may not have been adequate to assess an effect on interval cancer rates.[6]

LIMITATIONS / METHODOLOGICAL CONSIDERATIONS

The findings in this thesis are based on either the analysis of trends or the predictions of a microsimulation model. Although these are often used methods, both have their limitations.

Analyses of trends in population breast cancer mortality can provide an insight into possible changes in these trends after the implementation of breast cancer screening, which may be an important step in the evaluation of population based screening. Quantification of the screening effect using trend studies, however, comes with serious limitations.[11] On the one hand, the effect of screening may be diluted when analysing trends due to the inclusion of deaths occurring before invitation to screening, the time lag between first screening and the visibility of the screening effect and the inclusion of the period directly after the implementation

of screening.[11] On the other hand, trends in breast cancer mortality are not solely a reflection of the effect of screening over time but also of (improvements in) other diagnostic measures and (adjuvant) therapy. In the same period of time that screening was implemented in the Netherlands, the quality and administration of adjuvant therapy improved considerably.[52, 53] Opportunistic screening - i.e. screening outside the population based programme - may also contribute to the breast cancer mortality reduction.[54] The annual percentage change in breast cancer mortality is therefore not a quantification of the screening effect alone, but of the overall effect of (early) breast cancer detection and treatment. These drawbacks are important to take into account when interpreting the findings of trend studies.

Microsimulation models can be a proper tool to obtain insight in the effect of screening regimens when prospective studies are not feasible. In the present, performing screening trials is not always ethically justified as women in the age range eligible for screening cannot be withheld from current screening to function as a control group. In addition, to accurately evaluate the effect of screening on breast cancer mortality, large sample sizes and long follow-up are essential. Although models reflect a simplification of reality, they are able to extrapolate findings of RCTs and easily allow for the evaluation of multiple strategies, different attendance rates, different sub groups of women and lifelong follow-up. For our studies, we used the MISCAN model which is based on tumour progression through successive tumour stages, in terms of the tumour diameter. Probabilities of screen-detection and the prognosis after early detection by screening and treatment are different for these different tumour stages.[55, 56] MISCAN simulates the natural history of breast cancer at the individual level and, subsequently, the effect of screening for individual women. In this way, outcomes for a situation with and without screening can be compared at the individual level, as in a trial with perfect randomisation. There are, however, also several limitations to using microsimulation models for the evaluation of screening. An important limitation of breast cancer screening models in general is that there is a lack of proper external validation to demonstrate the generalizability of model outcomes to other populations.[57] However, for MISCAN there are several examples of external validation.[58]

Another drawback is that the quality of the model output is directly determined by the quality of the data used as model input. To incorporate the effect of screening on breast cancer mortality and the age- and stage-specific survival in MISCAN, data from the Swedish randomised trials on mammography screening were used.[55, 59, 60] The Independent UK Panel on Breast Cancer Screening concluded in 2013 that best evidence for the effectiveness of breast cancer screening in reducing breast cancer mortality reduction still comes from the randomised controlled trials of mammography screening.[61] However, these trials were conducted decades ago and several factors may have changed including for example the sensitivity of mammography – digital mammography screening is used now instead of screen-film mammography. In addition, even randomised controlled trials have their flaws and trials outcomes may thus be biased to a certain extent.[61] Furthermore, not all data needed to inform model input parameters are available from randomised trials (or observational studies). A significant part of the parameters associated with tumour progression in the natural history of disease part of the MISCAN model cannot be directly obtained from available data as they are unobservable, for example the duration of the preclinical screen-detectable stages. Parameters that are unobservable from data are estimated through model calibration and often depend on various assumptions. For certain other parameters, the literature may be scarce. Direct data on the natural history - the progression and possible regression - of DCIS, for example, are limited, as women diagnosed with DCIS are almost always treated directly. Recently trials in which DCIS is monitored instead of treated directly have started, including the LORD[62], LORIS[63] and COMET[64] trials. Data from these studies will provide more insight into the natural history of DCIS in the future.

Since the development of MISCAN, the model has been frequently recalibrated by updating parameters associated with breast cancer incidence, rates of screen-detected and interval cancers and adjuvant treatment probabilities, using the most recent data from the Dutch Screening Organisations and the Dutch National Cancer Registry.[65, 66] Although this concerns national data from reliable institutions, in some cases the sample sizes can still be rather small. For example

the probabilities of adjuvant treatment regimens for less common breast cancer stages as T1aN+ were based on small numbers of women with breast cancer.

As described above, a drawback of using models to predict the cost-effectiveness and benefits and harms of screening is that model parameters can be uncertain. Parameters with the most uncertainty in the MISCAN model are without doubt the unobservable parameters in the natural history of disease part of the model. Parameters estimated using empirical data can, however, also be uncertain. Inevitably, estimated quantities are uncertain to some extent. In addition, as mentioned before, the quality of the data from which parameters are estimated also contributes to parameter uncertainty. These data may come from screening trials that itself are limited in terms of validity, there may be trials that report conflicting results, or trial results may not be translatable to a population based screening programme.[67] Important to note is, however, that varying MISCAN parameters does usually not lead to altered conclusions, indicating that parameter uncertainty only affects model output to a certain extent. Studies that compare the outcomes of the MISCAN model for the United States to other models, with different model structure and assumptions, show consistency in conclusions based on model output.[23, 68] Furthermore, despite the uncertainty in model parameters MISCAN reproduces breast cancer incidence and mortality in the Netherlands quite accurately.[56]

To determine the impact of single parameters on model outcomes, one-way sensitivity analyses, in which single parameters are varied to a certain extent, can be conducted. However, the uncertainty around the single parameters is not taken into account in these analyses, which can lead to biased findings. It could be that uncertain parameters to which model outcomes are only slightly sensitive, actually have more impact on model outcomes because of their uncertainty than parameters to which model outcomes are very sensitive but that are estimated with little uncertainty.[67] In a PSA, the uncertainty around model parameters is taken into account by repeatedly sampling from a probability distribution around a parameter over a large number of model runs. However, PSAs are often not

conducted for MISCAN studies as these are extremely time consuming and methodologically complex.

A limitation specifically for Chapter 4 and Chapter 5 is that the effectiveness of mammography, which is incorporated in MISCAN through an improved survival after screen-detection of breast cancer, compared to clinical detection, is not specified by age. The modelled effect of screening is thus the same for the age group 40 to 49 years and for 50 years and older. This is not in line with findings from randomised trials that included women aged 40 to 49 years, which showed a lower reduction in breast cancer mortality for this age group.[29, 30] However, MISCAN does predict a lower reduction in breast cancer mortality as a result of screening because of several aspects: the duration of the preclinical tumour stages is shorter for women aged 40 to 49 years than for older women, which results in lower screening benefit; the sensitivity of digital mammography is lower for younger women; the survival after clinical detection of a tumour is higher for younger than for older women, which means that the improvement from detection by screening is relatively smaller than for older women.[55, 56]

FUTURE DIRECTIONS

Some of the topics addressed in this thesis will develop further in the future, when more advanced study outcomes will be available. For all future developments in breast cancer screening, it will be important to assess how they will affect the cost-effectiveness and the balance between the benefits and harms of screening.

Tomosynthesis

Digital breast tomosynthesis is already used regularly in clinical settings and will most likely gradually replace digital mammography in breast cancer screening programmes. To estimate the effectiveness and costs with tomosynthesis more accurately, additional information on subsequent screening rounds, optimal reading strategies and associated screening costs is required.

Risk stratification

As the knowledge on risk factors for breast cancer has improved substantially over the years, breast cancer screening stratified for different risk levels has become an interesting alternative to uniform screening. Important factors with regard to risk-based approaches are: assessing individual risk, identifying strategies for different risk levels and the acceptance among women eligible for screening. There are multiple ongoing studies that evaluate risk-based screening approaches, of which many are European initiatives. The MyPeBS (My Personal Breast Screening) initiative is a European project in which personalised screening based on a woman's risk of breast cancer for women aged 40-70 years is compared to current screening practice for five European countries.[37] In the Italian TBST (Tailored Breast Screening Trial) study, women aged 45-49 years with dense breasts in the intervention group are screened annually instead of biennially.[69] From age 50 years, all women are screened according to current screening practice. In the Dutch PRISMA study, risk factors for breast cancer are identified and data on the distributions of these risk factors among the Dutch female population are collected.[38] The findings of this study will be used to evaluate the effect of tailored screening strategies for women at relatively high and low risks of breast cancer in the Netherlands. A publication associated with the PRISMA study showed that women are receptive to individual risk assessment but that there are differences in preferences and needs of women between countries.[70]

Screening for women with dense breasts

A form of risk-based screening that receives a lot of attention lately is tailored screening of women with dense breasts. High breast density is a risk factor for breast cancer and is associated with lower sensitivity of mammography.[19, 71, 72] The value of MRI in addition to standard digital mammography screening practice was evaluated in the Dutch randomised controlled DENSE trial.[73] Women aged 50-75 years, eligible for breast cancer screening, with extremely dense breasts (American College of Radiology category 4 breast density) and a negative screening mammogram were either assigned to the control arm (standard practice) or to the intervention arm, in which they received an invitation for additional MRI. A recent publication of the results showed that the interval cancer rate in the group

that was invited to additional MRI was significantly lower than in the control group, emphasizing the additional value of MRI for women with extremely dense breasts.[74] Currently, the MISCAN model is used to estimate the impact of these outcomes, in terms of cost-effectiveness and the balance between benefits and harms, on the population level.

Recommendations for future research

- Estimating the impact, in terms of cost-effectiveness and the balance between benefits and harms, of a risk-based screening strategy addressing the different risk groups, for the population as a whole
- Assessing the feasibility and the acceptance among women of a risk-based screening approach in the Netherlands
- Determining how to best inform women, in particular women with a low relative risk, on the advantages and disadvantages of risk based screening
- Quantifying the possible additional overdiagnosis with digital breast tomosynthesis use, compared to screening with digital mammography
- Determining the optimal reading strategy with digital breast tomosynthesis in the Netherlands and estimating the corresponding costs of screening
- Determining what research would be most valuable to make a decision on implementing digital breast tomosynthesis
- Incorporating the four different categories of breast density in MISCAN model
- Using individual level risk factors linked to individual tumour progression in MISCAN model

FINAL CONCLUSIONS

- The programme sensitivity with digital mammography screening is higher than with screen-film mammography due to increased breast cancer detection with digital mammography.
- Despite the increase in referral rates with the transition to digital mammography screening, programme specificity in the Netherlands is still quite high compared to other European countries.

- Breast cancer mortality rates in the Netherlands declined significantly over the last decades as a result of the implementation of breast cancer screening and (improved) adjuvant treatment.
- Starting screening before age 50 years is a cost-effective approach to increase the screening benefit but comes at the expense of higher false-positive rates
- If generally accepted, screening stratified by risk has the potential to improve screening benefit for women with a high relative risk of breast cancer and to limit harms for women with a low relative risk, while not substantially affecting the cost-effectiveness of the screening programme.
- Analysing currently available data, screening with digital breast tomosynthesis is not expected to be a cost effective alternative in the Netherlands to digital mammography screening, but this conclusion may change in favour of tomosynthesis when using a somewhat higher, internationally accepted cost-effectiveness threshold or Dutch rates for discounting or when tomosynthesis becomes cheaper.

REFERENCES

1. de Munck L, de Bock GH, Otter R, Reiding D, Broeders MJ, Willemse PH, Siesling S. Digital vs screen-film mammography in population-based breast cancer screening: performance indicators and tumour characteristics of screen-detected and interval cancers. *Br J Cancer* 2016;115: 517-24.
2. van Bommel RMG, Weber R, Voogd AC, Nederend J, Louwman MWJ, Venderink D, Strobbe LJA, Rutten MJC, Plaisier ML, Lohle PN, Hooijen MJH, Tjan-Heijnen VCG, et al. Interval breast cancer characteristics before, during and after the transition from screen-film to full-field digital screening mammography. *BMC Cancer* 2017;17: 315.
3. Knox M, O'Brien A, Szabo E, Smith CS, Fenlon HM, McNicholas MM, Flanagan FL. Impact of full field digital mammography on the classification and mammographic characteristics of interval breast cancers. *Eur J Radiol* 2015;84: 1056-61.
4. Skaane P, Sebuodegard S, Bandos AI, Gur D, Osteras BH, Gullien R, Hofvind S. Performance of breast cancer screening using digital breast tomosynthesis: results from the prospective population-based Oslo Tomosynthesis Screening Trial. *Breast Cancer Res Treat* 2018;169: 489-96.
5. Houssami N, Bernardi D, Caumo F, Brunelli S, Fanto C, Valentini M, Romanucci G, Gentilini MA, Zorzi M, Macaskill P. Interval breast cancers in the 'screening with tomosynthesis or standard mammography' (STORM) population-based trial. *Breast* 2018;38: 150-3.
6. Bernardi D, Gentilini MA, De Nisi M, Pellegrini M, Fanto C, Valentini M, Sabatino V, Luparia A, Houssami N. Effect of implementing digital breast tomosynthesis (DBT) instead of mammography on population screening outcomes including interval cancer rates: Results of the Trento DBT pilot evaluation. *Breast* 2020;50: 135-40.
7. Gorini G, Zappa M, Miccinesi G, Paci E, Costantini AS. Breast cancer mortality trends in two areas of the province of Florence, Italy, where screening programmes started in the 1970s and 1990s. *Br J Cancer* 2004;90: 1780-3.
8. Ascunce EN, Moreno-Iribas C, Barcos Urtiaga A, Ardanaz E, Ederra Sanz M, Castilla J, Egues N. Changes in breast cancer mortality in Navarre (Spain) after introduction of a screening programme. *J Med Screen* 2007;14: 14-20.
9. Duffy SW, Tabar L, Olsen AH, Vitak B, Allgood PC, Chen TH, Yen AM, Smith RA. 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: 25-30.
10. Weedon-Fekjaer H, Romundstad PR, Vatten LJ. Modern mammography screening and breast cancer mortality: population study. *BMJ* 2014;348: g3701.

11. Moss SM, Nystrom L, Jonsson H, Paci E, Lynge E, Njor S, Broeders M, Euroscreen Working G. The impact of mammographic screening on breast cancer mortality in Europe: a review of trend studies. *J Med Screen* 2012;19 Suppl 1: 26-32.
12. Autier P, Koechlin A, Smans M, Vatten L, Boniol M. Mammography screening and breast cancer mortality in Sweden. *J Natl Cancer Inst* 2012;104: 1080-93.
13. Van Ourti T, O'Donnell O, Koc H, Fracheboud J, de Koning HJ. Effect of screening mammography on breast cancer mortality: Quasi-experimental evidence from rollout of the Dutch population-based program with 17-year follow-up of a cohort. *Int J Cancer* 2020;146: 2201-8.
14. de Gelder R, Heijnsdijk EA, Fracheboud J, Draisma G, de Koning HJ. The effects of population-based mammography screening starting between age 40 and 50 in the presence of adjuvant systemic therapy. *Int J Cancer* 2015;137: 165-72.
15. Plevritis SK, Munoz D, Kurian AW, Stout NK, Alagoz O, Near AM, Lee SJ, van den Broek JJ, Huang X, Schechter CB, Sprague BL, Song J, et al. Association of Screening and Treatment With Breast Cancer Mortality by Molecular Subtype in US Women, 2000-2012. *JAMA* 2018;319: 154-64.
16. Giordano L, von Karsa L, Tomatis M, Majek O, de Wolf C, Lancucki L, Hofvind S, Nystrom L, Segnan N, Ponti A, Eunice Working G, Van Hal G, et al. Mammographic screening programmes in Europe: organization, coverage and participation. *J Med Screen* 2012;19 Suppl 1: 72-82.
17. Lauby-Secretan B, Scoccianti C, Loomis D, Benbrahim-Tallaa L, Bouvard V, Bianchini F, Straif K, International Agency for Research on Cancer Handbook Working G. Breast-cancer screening--viewpoint of the IARC Working Group. *N Engl J Med* 2015;372: 2353-8.
18. 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.
19. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, Geller BM, Abraham LA, Taplin SH, Dignan M, Cutter G, Ballard-Barbash R. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med* 2003;138: 168-75.
20. Buist DS, Porter PL, Lehman C, Taplin SH, White E. Factors contributing to mammography failure in women aged 40-49 years. *J Natl Cancer Inst* 2004;96: 1432-40.
21. Bailey SL, Sigal BM, Plevritis SK. A simulation model investigating the impact of tumor volume doubling time and mammographic tumor detectability on screening outcomes in women aged 40-49 years. *J Natl Cancer Inst* 2010;102: 1263-71.

22. Koleva-Kolarova RG, Daszczuk AM, de Jonge C, Abu Hantash MK, Zhan ZZ, Postema EJ, Feenstra TL, Pijnappel RM, Greuter MJW, de Bock GH. A modelling study to evaluate the costs and effects of lowering the starting age of population breast cancer screening. *Maturitas* 2018;109: 81-8.

23. Mandelblatt JS, Stout NK, Schechter CB, van den Broek JJ, Miglioretti DL, Krapcho M, Trentham-Dietz A, Munoz D, Lee SJ, Berry DA, van Ravesteyn NT, Alagoz O, et al. Collaborative Modeling of the Benefits and Harms Associated With Different U.S. Breast Cancer Screening Strategies. *Ann Intern Med* 2016;164: 215-25.

24. 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.

25. Stout NK, Lee SJ, Schechter CB, Kerlikowske K, Alagoz O, Berry D, Buist DS, Cevik M, Chisholm G, de Koning HJ, Huang H, Hubbard RA, et al. Benefits, harms, and costs for breast cancer screening after US implementation of digital mammography. *J Natl Cancer Inst* 2014;106: dju092.

26. https://healthcare-quality.jrc.ec.europa.eu/european-breast-cancer-guidelines/screening-ages-and-frequencies.

27. 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.

28. Perry NM, Patani N, Milner SE, Pinker K, Mokbel K, Allgood PC, Duffy SW. The impact of digital mammography on screening a young cohort of women for breast cancer in an urban specialist breast unit. *Eur Radiol* 2011;21: 676-82.

29. Moss SM, Wale C, Smith R, Evans A, Cuckle H, Duffy SW. Effect of mammographic screening from age 40 years on breast cancer mortality in the UK Age trial at 17 years' follow-up: a randomised controlled trial. *Lancet Oncol* 2015;16: 1123-32.

30. Nelson HD, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L. Effectiveness of Breast Cancer Screening: Systematic Review and Meta-analysis to Update the 2009 U.S. Preventive Services Task Force Recommendation. *Ann Intern Med* 2016;164: 244-55.

31. Magnus MC, Ping M, Shen MM, Bourgeois J, Magnus JH. Effectiveness of mammography screening in reducing breast cancer mortality in women aged 39-49 years: a meta-analysis. *J Womens Health (Larchmt)* 2011;20: 845-52.

32. 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.

33. Collaborative Group on Hormonal Factors in Breast C. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* 2001;358: 1389-99.

34. Collaborative Group on Hormonal Factors in Breast C. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol* 2012;13: 1141-51.

35. Tice JA, Miglioretti DL, Li CS, Vachon CM, Gard CC, Kerlikowske K. Breast Density and Benign Breast Disease: Risk Assessment to Identify Women at High Risk of Breast Cancer. *J Clin Oncol* 2015;33: 3137-43.

36. Michailidou K, Lindstrom S, Dennis J, Beesley J, Hui S, Kar S, Lemacon A, Soucy P, Glubb D, Rostamianfar A, Bolla MK, Wang Q, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature* 2017;551: 92-4.

37. https://mypebs.eu/the-project/.

38. https://www.prisma-studie.nl/.

39. Giordano L, Gallo F, Petracci E, Chiorino G, Segnan N, Andromeda working g. The ANDROMEDA prospective cohort study: predictive value of combined criteria to tailor breast cancer screening and new opportunities from circulating markers: study protocol. *BMC Cancer* 2017;17: 785.

40. Esserman LJ, Study W, Athena I. The WISDOM Study: breaking the deadlock in the breast cancer screening debate. *NPJ Breast Cancer* 2017;3: 34.

41. Brentnall AR, Evans DG, Cuzick J. Distribution of breast cancer risk from SNPs and classical risk factors in women of routine screening age in the UK. *Br J Cancer* 2014;110: 827-8.

42. Zwaap J. KS vdMC, Staal P., van der Heiden L. Rapport kosteneffectiviteit in de praktijk. Zorginstituut Nederland 2015.

43. Evans DG, Donnelly LS, Harkness EF, Astley SM, Stavrinos P, Dawe S, Watterson D, Fox L, Sergeant JC, Ingham S, Harvie MN, Wilson M, et al. Breast cancer risk feedback to women in the UK NHS breast screening population. *Br J Cancer* 2016;114: 1045-52.

44. Lang K, Andersson I, Rosso A, Tingberg A, Timberg P, Zackrisson S. Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmo Breast Tomosynthesis Screening Trial, a population-based study. *Eur Radiol* 2016;26: 184-90.

45. Hofvind S, Hovda T, Holen AS, Lee CI, Albertsen J, Bjorndal H, Brandal SHB, Gullien R, Lomo J, Park D, Romundstad L, Suhrke P, et al. Digital Breast Tomosynthesis and Synthetic 2D Mammography versus Digital Mammography: Evaluation in a Population-based Screening Program. *Radiology* 2018;287: 787-94.

46. Caumo F, Zorzi M, Brunelli S, Romanucci G, Rella R, Cugola L, Bricolo P, Fedato C, Montemezzi S, Houssami N. Digital Breast Tomosynthesis with Synthesized Two-Dimensional Images versus Full-Field Digital Mammography for Population Screening: Outcomes from the Verona Screening Program. *Radiology* 2018;287: 37-46.

47. Pattacini P, Nitrosi A, Giorgi Rossi P, Iotti V, Ginocchi V, Ravaioli S, Vacondio R, Braglia L, Cavuto S, Campari C, Group REW. Digital Mammography versus Digital Mammography Plus Tomosynthesis for Breast Cancer Screening: The Reggio Emilia Tomosynthesis Randomized Trial. *Radiology* 2018;288: 375-85.

48. Skaane P, Bandos AI, Niklason LT, Sebuodegard S, Osteras BH, Gullien R, Gur D, Hofvind S. Digital Mammography versus Digital Mammography Plus Tomosynthesis in Breast Cancer Screening: The Oslo Tomosynthesis Screening Trial. *Radiology* 2019;291: 23-30.

49. Zackrisson S, Lang K, Rosso A, Johnson K, Dustler M, Fornvik D, Fornvik H, Sartor H, Timberg P, Tingberg A, Andersson I. One-view breast tomosynthesis versus two-view mammography in the Malmo Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study. *Lancet Oncol* 2018;19: 1493-503.

50. Marinovich ML, Hunter KE, Macaskill P, Houssami N. Breast Cancer Screening Using Tomosynthesis or Mammography: A Meta-analysis of Cancer Detection and Recall. *J Natl Cancer Inst* 2018;110: 942-9.

51. Ciatto S, Houssami N, Bernardi D, Caumo F, Pellegrini M, Brunelli S, Tuttobene P, Bricolo P, Fanto C, Valentini M, Montemezzi S, Macaskill P. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. *Lancet Oncol* 2013;14: 583-9.

52. Voogd AC, van Beek MW, Crommelin MA, Kluck HM, Repelaer van Driel OJ, Coebergh JW. Management of early breast cancer in southeast Netherlands since 1984. A population-based study. Regional Breast Cancer Study Group. *Acta Oncol* 1994;33: 753-7.

53. Nab HW, Hop WC, Crommelin MA, Kluck HM, van der Heijden LH, Coebergh JW. Changes in long term prognosis for breast cancer in a Dutch cancer registry. *BMJ* 1994;309: 83-6.

54. Lynge E, Braaten T, Njor SH, Olsen AH, Kumle M, Waaseth M, Lund E. Mammography activity in Norway 1983 to 2008. *Acta Oncol* 2011;50: 1062-7.

55. 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.

56. Sankatsing VD, Heijnsdijk EA, van Luijt PA, van Ravesteyn NT, Fracheboud J, de Koning HJ. Cost-effectiveness of digital mammography screening before the age of 50 in The Netherlands. *Int J Cancer* 2015;137: 1990-9.

57. Koleva-Kolarova RG, Zhan Z, Greuter MJ, Feenstra TL, De Bock GH. Simulation models in population breast cancer screening: A systematic review. *Breast* 2015;24: 354-63.

58. de Koning HJ, Alagoz O, Schechter CB, van Ravesteyn NT. Reply to Koleva-Kolarova et al. *Breast* 2016;27: 182-3.

59. Tabar L FC, Day NE. The results of periodic oneview mammography screening in a randomized, controlled trial in Sweden. Part 2: Evaluation of the results. In: Day NE, Miller AB, eds. Screening for Breast Cancer. Toronto: H. Huber Publishers, 1988; 39-44.

60. Rutqvist LE. On the utility of the lognormal model for analysis of breast cancer survival in Sweden 1961-1973. *Br J Cancer* 1985;52: 875-83.

61. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. *Br J Cancer* 2013;108: 2205-40.

62. Elshof LE, Tryfonidis K, Slaets L, van Leeuwen-Stok AE, Skinner VP, Dif N, Pijnappel RM, Bijker N, Rutgers EJ, Wesseling J. Feasibility of a prospective, randomised, open-label, international multicentre, phase III, non-inferiority trial to assess the safety of active surveillance for low risk ductal carcinoma in situ - The LORD study. *Eur J Cancer* 2015;51: 1497-510.

63. Francis A, Thomas J, Fallowfield L, Wallis M, Bartlett JM, Brookes C, Roberts T, Pirrie S, Gaunt C, Young J, Billingham L, Dodwell D, et al. Addressing overtreatment of screen detected DCIS; the LORIS trial. *Eur J Cancer* 2015;51: 2296-303.

64. Hwang ES, Hyslop T, Lynch T, Frank E, Pinto D, Basila D, Collyar D, Bennett A, Kaplan C, Rosenberg S, Thompson A, Weiss A, et al. The COMET (Comparison of Operative versus Monitoring and Endocrine Therapy) trial: a phase III randomised controlled clinical trial for low-risk ductal carcinoma in situ (DCIS). *BMJ Open* 2019;9: e026797.

65. National Evaluation Team for Breast cancer screening (NETB). National evaluation of breast cancer screening in the Netherlands 2004-2014. Fourteenth evaluation report. Erasmus MC University Medical Center Rotterdam, Radboud University Nijmegen Medical Centre, 2019.

66. www.iknl.nl.

67. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD, Force I-SMGRPT. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. *Med Decis Making* 2012;32: 722-32.

68. van Ravesteyn NT, Stout NK, Schechter CB, Heijnsdijk EA, Alagoz O, Trentham-Dietz A, Mandelblatt JS, de Koning HJ. Benefits and harms of mammography screening after age 74 years: model estimates of overdiagnosis. *J Natl Cancer Inst* 2015;107.

69. Paci E, Mantellini P, Giorgi Rossi P, Falini P, Puliti D, Group TW. [Tailored Breast Screening Trial (TBST)]Tailored Breast Screening Trial (TBST). *Epidemiol Prev* 2013;37: 317-27.

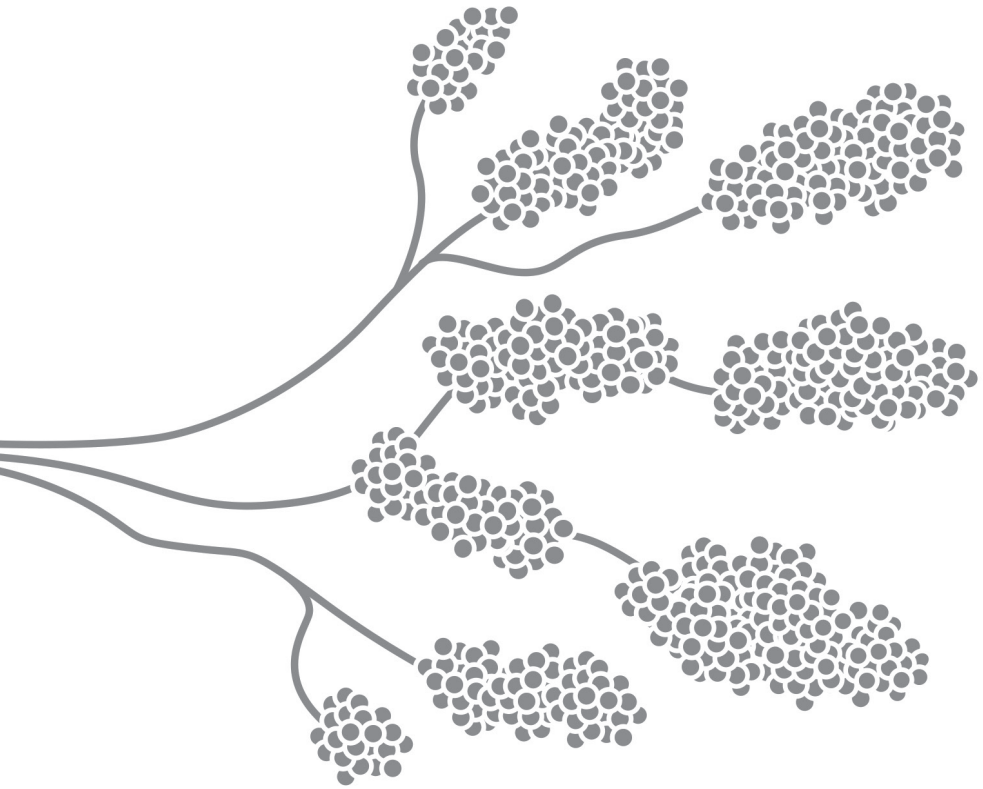
70. Rainey L, van der Waal D, Jervaeus A, Donnelly LS, Evans DG, Hammarstrom M, Hall P, Wengstrom Y, Broeders MJM. European women's perceptions of the implementation and organisation of risk-based breast cancer screening and prevention: a qualitative study. *BMC Cancer* 2020;20: 247.

71. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology* 2002;225: 165-75.

72. Engmann NJ, Golmakani MK, Miglioretti DL, Sprague BL, Kerlikowske K, Breast Cancer Surveillance C. Population-Attributable Risk Proportion of Clinical Risk Factors for Breast Cancer. *JAMA Oncol* 2017;3: 1228-36.

73. Emaus MJ, Bakker MF, Peeters PH, Loo CE, Mann RM, de Jong MD, Bisschops RH, Veltman J, Duvivier KM, Lobbes MB, Pijnappel RM, Karssemeijer N, et al. MR Imaging as an Additional Screening Modality for the Detection of Breast Cancer in Women Aged 50-75 Years with Extremely Dense Breasts: The DENSE Trial Study Design. *Radiology* 2015;277: 527-37.

74. Bakker MF, de Lange SV, Pijnappel RM, Mann RM, Peeters PHM, Monninkhof EM, Emaus MJ, Loo CE, Bisschops RHC, Lobbes MBI, de Jong MDF, Duvivier KM, et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. *N Engl J Med* 2019;381: 2091-102.



Summary

SUMMARY

Breast cancer incidence in the Netherlands is among the highest in Europe. Almost all women with breast cancer are treated, regardless of the breast cancer stage. Mammography screening can result in early detection of breast cancer and, possibly, earlier treatment and better prognosis. In the Netherlands, women aged 50 to 74 are invited biennially to digital mammography screening. The benefits of screening are (quality adjusted) life years gained and prevented breast cancer deaths. However, screening also leads to harms including false-positive screens, overdiagnosis and overtreatment and false reassurance.

Screening programmes are assessed using several performance indicators including the detection rate, the interval cancer rate, programme sensitivity, the referral rate, the false-positive rate and programme specificity.

Several factors that may increase the risk of breast cancer have been identified. Common risk factors are age, a family history of breast cancer, high breast density and a previous biopsy. Reproductive and life style factors have also been associated with higher breast cancer risk. Currently, there are multiple studies that explore risk-based breast cancer screening.

There is no consensus about the optimal screening strategy, in terms of for example the age range. In addition, new developments could improve current strategies. Using simulation models, the long-term effects of different screening strategies can be predicted. The aim of this thesis was to quantify the effects of the current screening programme in the Netherlands and to determine the effects and cost-effectiveness of alternative screening strategies including strategies starting before age 50 years, risk-based strategies and screening with digital breast tomosynthesis.

Part One: Evaluation of current breast cancer screening in the Netherlands

In the study described in **Chapter 2**, data of the Dutch Screening Organisations on all screening examinations performed between 2004 and 2011 were linked to data of the Netherlands Cancer Registry to estimate the detection and interval cancer rates. Besides the trends over time, possible differences in performance indicators between screen-film and digital mammography were addressed. Digital mammography gradually replaced screen-film mammography in the Netherlands from 2008 to 2010. The breast cancer detection rate in the Netherlands, for both in ductal carcinoma in situ (DCIS) and invasive carcinomas, increased by more than 20% over the years 2004-2011. The programme sensitivity, referral rate and false-positive rate also increased. The interval cancer rate remained stable during the study period. Between 2004 and 2011, 7,343,327 screening examinations were performed, of which 64% were screen-film and 36% were digital screens. The detection rate was significantly higher with digital mammography (6.2 per 1000 screens) than with film-screen mammography (5.4 per 1000 screens). The interval cancer rate was similar with both modalities. As a result of the higher detection rate with digital mammography, the programme sensitivity was significantly higher than with film-screen mammography. The referral rate and false-positive rate were also higher, resulting in a lower programme specificity with digital mammography. The programme specificity in the Netherlands is, however, still relatively high compared to other countries.

Analysing the data for 5-year age groups separately, a relatively high increase in programme sensitivity and a slight (non-significant) decrease in the interval cancer rate was observed in women younger than 60 years. These trend changes may be (partly) attributable to the transition to digital mammography screening. Further research is required to assess whether particularly younger women have benefitted from the change to digital mammography screening.

In the next chapter, **Chapter 3**, the trend in breast cancer mortality in the Netherlands between 1980 and 2010 is analysed. Breast cancer screening was gradually implemented in the Netherlands, starting with a few municipalities in 1989/1990 and completing implementation in 1997. Therefore, breast cancer mortality was analysed on a municipality-level, using the municipality specific year of implementation of screening as starting point of the trend analysis. Only overall results– on a national level – are reported. The breast cancer mortality rate

in the Netherlands decreased by 30%, since the introduction of screening until the year 2010, in women aged 55-74 years. The start of this decrease was observed around two years after the introduction of screening. In the age group 75-79 years, a significant decrease of 34% was observed. During the implementation phase of screening, adjuvant therapy in the Netherlands improved, in terms of usage and quality, which caused a decline in breast cancer mortality rates. It is difficult to disentangle the effect of screening, treatment and other factors on breast cancer mortality using trend studies. Analysing breast cancer mortality rates using individual data would provide more accurate and reliable estimates. However, a significant decrease in breast cancer mortality was also found in subgroups of municipalities with an early implementation year of screening, before (the use of) adjuvant therapy improved.

Part two: Quantifying the cost-effectiveness, harms and benefits of different screening strategies and screening modalities in the Netherlands using microsimulation modelling

In **Chapter 4**, the cost-effectiveness of extending digital mammography screening in the Netherlands below age 50 years is assessed. The current Dutch screening programme biennially invites women between age 50 and 74 years to digital mammography screening. However, the breast cancer incidence among women in their forties has increased over the last decades. This study quantified the cost-effectiveness of screening women between age 40 and 50 years in addition to the current screening strategy and estimated the increase in benefits and harms that come with starting screening earlier. Using the MISCAN microsimulation model, different strategies with a starting age between 40 and 50 years were simulated. The results of this study showed that the simulated strategies were cost-effective using the Dutch cost-effectiveness threshold of €20,000 per life year gained. Two strategies were dominated: the strategy with an additional screen at age 49 years and the strategy with additional annual screening between age 45 and 49 years. These strategies were dominated by biennial strategies, which were more effective and less costly per additional life year gained. In terms of increased harms with additional screening, it was predicted that additional screening between age 40 and 50 years only slightly increased overdiagnosis, but substantially increased the number of false-positive findings.

Chapter 5 addresses the cost-effectiveness of another alternative to current screening: screening women based on their risk of breast cancer. Using MISCAN, women with a relative breast cancer risk of 0.75 and 1.8 were simulated. For women with a low relative risk (0.75) screening strategies that were less intensive than the current strategy – biennial screening between age 50 and 74 years - were simulated and for women with a high relative risk (1.8) strategies that were more intensive. The optimal screening strategy was estimated for the two risk-groups by selecting the screening strategy with an incremental cost-efffectiveness ratio just below the cost-effectiveness of current screening, in order that the risk-based strategy does not result in a less favourable ratio of costs to effects at the population level. The model predicted that the optimal strategy for women at lower than average risk was triennial screening between age 50 and 71 years. Compared to current screening, this strategy resulted in a decline in the number of false-positive findings of 33% and an improved harm-benefit ratio, because the decrease in the effect of screening (life years gained) was relatively small. The optimal strategy for women with a high relative risk was biennial screening between age 40 and 74 years. This strategy led to a 25% increase in the number of life years gained, but at the same time to a 44% increase in the number of false-positive findings, resulting in a more unfavourable harm-benefit ratio. The findings of this study show that risk-based screening for women with a lower than average risk can result in a decrease in screening harms while maintaining most of the screening effect. For women with a higher than average risk, risk-based screening can lead to a gain in screening effect, at the expense of a substantial increase in false-positive findings. Future research is required to estimate the impact of risk-based screening at the population level, by linking the relative risks to individual risk factors and their prevalence. It is also important to take breast density into account as a risk factor for breast cancer.

In **Chapter 6**, the use of an alternative screening test, digital breast tomosynthesis, is explored. Digital breast tomosynthesis generates multiple images of the breast

at different angles, creating a pseudo 3D image. Recent trials have shown that screening with tomosynthesis leads to an increased breast cancer detection rate, compared with digital mammography. In this chapter, the cost-effectiveness of tomosynthesis is estimated. To estimate the cost-effectiveness, 10,000 model runs, in which a cohort of women was biennially screened with tomosynthesis between age 50 and 74 years, were performed. Subsequently, the outcomes for tomosynthesis were compared to the outcomes of current screening – biennial digital mammography between 50 and 74 years. The model predicted a 7% increase in the number of life years gained with tomosynthesis, a 6% increase in the number of breast cancer deaths averted and a 2% decrease in the number of false-positive findings, compared to digital mammography, per 1000 women invited to screening with lifelong follow-up. The average incremental costs of tomosynthesis were estimated at €137,555 and the incremental cost-effectiveness ratio (ICER) at €27,023 per life year gained, using annual discounting of 3.5% for both costs and effects. When comparing tomosynthesis and digital mammography, the probability of being most cost-effective was 0.36 for tomosynthesis and 0.64 for digital mammography, using a cost-effectiveness threshold of €20,000 per life year gained. At a threshold of €35,000 per life year gained, tomosynthesis was the strategy with the highest probability of being most cost-effective, with a probability of 0.66.

Using a cost-effectiveness threshold of €20.000 per life-year gained, according to Dutch guidelines, screening with tomosynthesis, with an ICER of €27,023, would not be cost-effective. However, Dutch guidelines recommend annual discounting with 4% for costs and 1.5 for effects, instead of equal discounting with 3.5% for both. Applying Dutch discounting recommendations would result in a more favourable ICER.

The costs and effects of tomosynthesis in a screening setting are still rather uncertain. Additional research taking into account the reading strategy, the exact screening costs with tomosynthesis and the screening effect over multiple screening rounds is therefore necessary.

Conclusions and recommendations

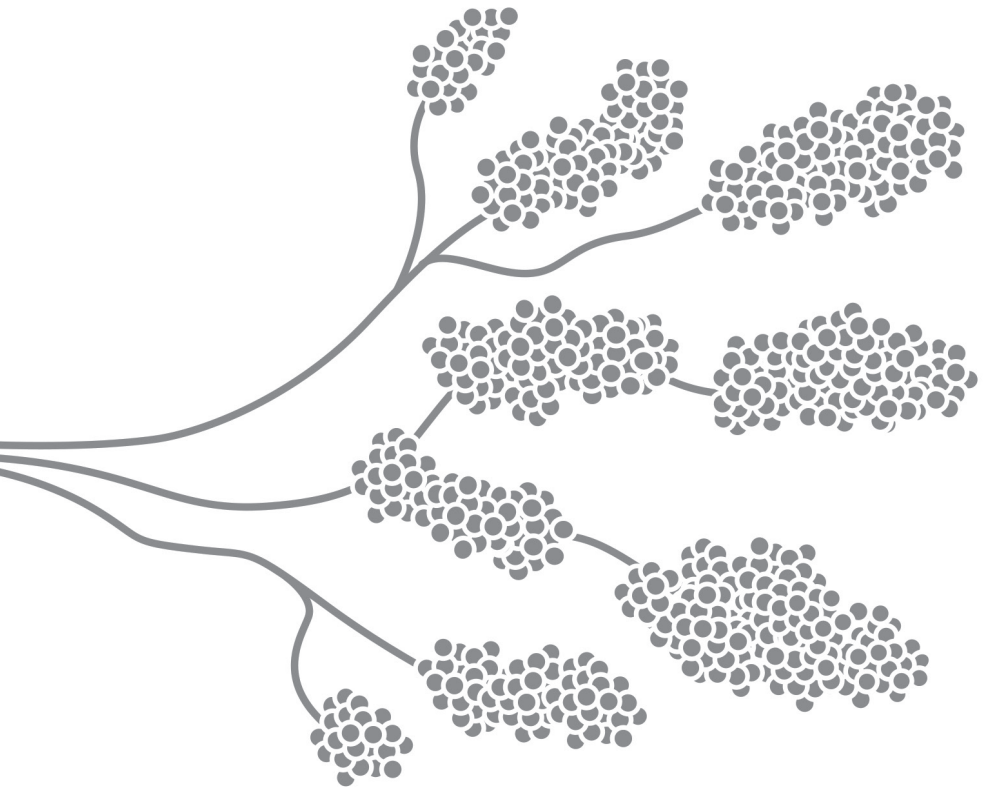
The following conclusions can be drawn from the results of the studies in this thesis:

- The programme sensitivity with digital mammography screening is higher than with screen-film mammography due to increased breast cancer detection with digital mammography.
- Despite the increase in referral rates with the transition to digital mammography screening, programme specificity in the Netherlands is still quite high compared to other European countries.
- Breast cancer mortality rates in the Netherlands declined significantly over the last decades as a result of the implementation of breast cancer screening and (improved) adjuvant treatment.
- Starting screening before age 50 years is a cost-effective approach to increase the screening benefit but comes at the expense of higher false-positive rates
- If generally accepted, screening stratified by risk has the potential to improve screening benefit for women with a high relative risk of breast cancer and to limit harms for women with a low relative risk, while not substantially affecting the cost-effectiveness of the screening programme.
- Analysing currently available data, screening with digital breast tomosynthesis is not expected to be a cost effective alternative in the Netherlands to digital mammography screening, but this conclusion may change in favour of tomosynthesis when using a somewhat higher, internationally accepted cost-effectiveness threshold or Dutch rates for discounting or when tomosynthesis becomes cheaper.

Based on these conclusions, the following recommendations were formulated:

- Estimating the impact, in terms of cost-effectiveness and the balance between benefits and harms, of a risk-based screening strategy addressing the different risk groups, for the population as a whole

- Assessing the feasibility and the acceptance among women of a risk-based screening approach in the Netherlands
- Determining how to best inform women, in particular women with a low relative risk, on the advantages and disadvantages of risk based screening
- Quantifying the possible additional overdiagnosis with digital breast tomosynthesis use, compared to screening with digital mammography
- Determining the optimal reading strategy with digital breast tomosynthesis in the Netherlands and estimating the corresponding costs of screening
- Determining what research would be most valuable to make a decision on implementing digital breast tomosynthesis
- Incorporating the four different categories of breast density in MISCAN model
- Using individual level risk factors linked to individual tumour progression in MISCAN model



Samenvatting

SAMENVATTING

De borstkankerincidentie in Nederland is relatief hoog vergeleken met andere Europese landen. Vrijwel alle vrouwen met borstkanker, ongeacht het stadium, worden behandeld. Screening met mammografie kan leiden tot een vroegere opsporing van borstkanker en daarmee mogelijk tot een eerdere behandeling en betere prognose. In Nederland worden vrouwen tussen de 50 en 74 jaar eens in de twee jaar uitgenodigd om deel te nemen aan screening met digitale mammografie. De belangrijkste voordelen van screening zijn gewonnen (kwalitatief goede) levensjaren en voorkomen borstkankerdoden. Screening resulteert ook in nadelen, waaronder fout-positieve uitslagen, overdiagnose en overbehandeling en onterechte geruststelling. Om het functioneren van een screeningsprogramma te beoordelen zijn verschillende indicatoren van belang, waaronder het detectiecijfer, het intervalkankercijfer, de programmasensitiviteit, het verwijscijfer, het fout-positievencijfer en de programmaspecificiteit.

Factoren waarvan bekend is dat ze het risico op borstkanker kunnen verhogen zijn onder anderen leeftijd, geslacht, familiegeschiedenis van borstkanker, dicht borstweefsel en een voorgaande biopsie. Ook factoren die samenhangen met de voortplanting en levensstijlfactoren kunnen bijdragen aan een verhoogde kans op borstkanker. Er lopen op dit moment verschillende studies naar op risico gebaseerde screeningsstrategieën.

Er is geen consensus over de optimale screeningsstrategie, bijvoorbeeld wat betreft leeftijdsgrenzen. Bovendien zijn er nieuwe ontwikkelingen op het gebied van screening die zouden kunnen leiden tot een verbetering van huidig toegepaste strategieën. Met behulp van simulatiemodellen kunnen lange-termijn voorspellingen worden gedaan over de effecten en de kosteneffectiviteit van verschillende screeningsstrategieën. Het doel van dit proefschrift was het in kaart brengen van de voordelige en nadelige effecten van het huidige screeningsprogramma in Nederland en het onderzoeken van alternatieve manieren van screening, zoals het eerder starten met screenen, het screenen op basis van het risico op borstkanker en het screenen met een andere screeningstechniek

dan digitale mammografie, namelijk tomosynthese. Hierbij is er gekeken naar de kosteneffectiviteit en de voor- en nadelen van alternatieve screening.

Deel 1: Evaluatie van het huidige borstkanker screeningsprogramma in Nederland

In **hoofdstuk 2** wordt data van de Nederlandse screeningsorganisaties van alle screens uitgevoerd in de periode 2004 tot 2011 geanalyseerd om het detectie- en intervalkankercijfer te bepalen. Hiertoe zijn de screeningsdata gelinkt aan data van de Nederlandse Kankerregistratie. Naast de trends over de tijd is er ook gekeken naar eventuele verschillen in screeningsindicatoren tussen analoge en digitale mammografie. Deze laatste heeft analoge mammografie vanaf 2008 geleidelijk vervangen. Over de gehele onderzoeksperiode is het detectiecijfer van borstkanker in Nederland, voor ductal carcinoma in situ (DCIS) en invasieve kankers samen, met meer dan 20% gestegen. Ook de programmasensitiviteit, het verwijscijfer en het fout-positievencijfer zijn toegenomen. Het intervalkankercijfer is stabiel gebleven over de periode 2004 tot 2011. In de onderzoeksperiode bestond het totaal aantal uitgevoerde screeningsonderzoeken (7.343.327) voor 64% uit analoge screens en voor 36% uit digitale screens. Het detectiecijfer van digitale mammografie was met 6,2 per 1000 screens significant hoger dan het detectiecijfer met analoge mammografie (5,4 per 1000 screens). Het intervalkankercijfer was gelijk voor beide screeningsmodaliteiten. Als gevolg van het hogere detectiecijfer was de programmasensitiviteit voor digitale mammografie ook hoger dan voor analoge mammografie. Digitale mammografie had tevens een significant hoger verwijscijfer en fout-positievencijfer, waardoor de programmaspecificiteit lager was dan die van analoge mammografie. De programmaspecificiteit in Nederland is internationaal gezien echter nog steeds relatief hoog.

Wanneer de resultaten werden uitgesplitst naar 5-jaar leeftijdsgroepen viel op dat er in de groep vrouwen onder de 60 jaar een relatief hoge stijging in de programmasensitiviteit heeft plaatsgevonden en een lichte (niet significante) daling in het intervalkankercijfer, die bij oudere vrouwen niet geobserveerd werd. Deze veranderingen zijn mogelijk deels toe te schrijven aan de transitie naar digitale

mammografie. Toekomstige analyses moeten uitwijzen of met name jongere vrouwen hebben geprofiteerd van de overstap naar digitale mammografie.

Het volgende hoofdstuk, **hoofdstuk 3**, behandelt de analyse van de trend in borstkankersterfte in Nederland over de periode 1980 tot 2010. Borstkankerscreening is geleidelijk ingevoerd in Nederland, beginnend met een aantal gemeenten in 1989/1990 tot in 1997, toen de gehele vrouwelijke bevolking in de in aanmerking komende leeftijdsgroep werd uitgenodigd. Om deze reden is de borstkankersterfte geanalyseerd op gemeenteniveau, door het jaartal van invoering van borstkankerscreening in de verschillende Nederlandse gemeenten als startpunt van de analyse te nemen. De resultaten worden alleen op nationaal niveau gerapporteerd. In heel Nederland is de borstkankersterfte vanaf de introductie van screening tot aan het jaar 2010, in vrouwen in de leeftijdscategorie 55-74 jaar, met 30% gedaald. Het begin van de daling in de trend is zichtbaar 2 jaar na implementatie van screening. In de leeftijdsgroep 75-79 jaar werd een significante daling van 34% geobserveerd. Gedurende de implementatieperiode van screening is de adjuvante therapie in Nederland verbeterd in kwaliteit en meer toegepast, wat tot een daling in borstkankersterfte heeft geleid. Het is met trendstudies niet mogelijk om de afzonderlijke effecten van screening, therapie en andere factoren op de borstkankersterfte uit elkaar te halen. Een studie gebaseerd op individuele data zou meer accurate en betrouwbare schattingen leveren. In dit onderzoek werd echter ook in subgroepen van gemeenten met een vroege implementatie van screening, in de periode met minder en minder goede adjuvante therapie, een significante daling in borstkankersterfte gevonden.

Deel 2: Het kwantificeren van de kosteneffectiviteit en de voor- en nadelen van verschillende screeningsstrategieën en screenings-modaliteiten in Nederland met behulp van micro simulatie

In **hoofdstuk 4** wordt de kosteneffectiviteit van het eerder starten met digitaal screenen op borstkanker bij vrouwen in Nederland onderzocht. In het huidige screeningsprogramma worden vrouwen eens in de twee jaar uitgenodigd voor screening, tussen leeftijd 50 en 74 jaar. De incidentie van borstkanker onder vrouwen in de 40 is de afgelopen tijd echter toegenomen. In dit hoofdstuk is gekeken

of het additioneel screenen van vrouwen tussen leeftijd 40 en 50 kosteneffectief zou zijn en wordt de toename in voor- en nadelen bepaald, die gepaard gaat met het eerder starten van screenen. Met behulp van het MISCAN model zijn verschillende strategieën met een startleeftijd tussen de 40 en 50 jaar gesimuleerd. De resultaten van dit onderzoek lieten zien dat de gesimuleerde strategieën kosteneffectief zijn, uitgaande van de vaak gerapporteerde drempelwaarde voor kosteneffectiviteit van €20.000 per gewonnen levensjaar in Nederland. Twee strategieën werden echter gedomineerd, namelijk: een additionele screen op leeftijd 49 als toevoeging op het huidige programma en jaarlijkse screening tussen leeftijd 45 en 49 als toevoeging op de huidige strategie. Deze strategieën werden gedomineerd door strategieën waarbij eens in de twee jaar wordt gescreend, die effectiever waren en minder kostten per additioneel gewonnen levensjaar. Wat betreft de voorspelde nadelen van screening viel het op dat met het toevoegen van screening tussen leeftijd 40 en 50 aan de huidige screeningsstrategie overdiagnose nauwelijks toenam terwijl het aantal fout-positieve bevindingen sterk steeg met de intensiviteit van een screeningsstrategie.

Hoofdstuk 5 behandelt de kosteneffectiviteit van een andere variatie op het huidige screeningsprogramma, namelijk het verschillend screenen van vrouwen met een verlaagd en een verhoogd risico op borstkanker. Met MISCAN zijn vrouwen met een relatief risico van 0.75 en vrouwen met een relatief risico van 1.8 gesimuleerd. Voor vrouwen met een verlaagd relatief risico (0.75) zijn strategieën gesimuleerd die minder intensief zijn dan de huidige strategie (interval van 2 jaar in leeftijd 50-74) en voor vrouwen met een verhoogd relatief risico (1.8) strategieën die intensiever zijn. Voor deze twee groepen is de optimale screeningsstrategie bepaald door van alle gesimuleerde strategieën de strategie te kiezen die een incrementele kosteneffectiviteitsratio heeft die net onder de kosteneffectiviteit van het huidige screeningsprogramma valt, zodat de kosteneffectiviteit niet minder gunstig uitvalt op populatieniveau dan die van de huidige screening. Met het model is geschat dat de optimale strategie voor vrouwen met een laag risico een interval van 3 jaar heeft en een start- en stopleeftijd van respectievelijk 50 en 71 jaar. Vergeleken met de huidige strategie resulteerde deze optimale strategie in een verlaging van het aantal fout-positieve bevinding van 33% en een verbeterde ratio van voor en

nadelen omdat het screeningseffect (gewonnen levensjaren) minder sterk afneemt. Voor vrouwen met een verhoogd risico is geschat dat de optimale strategie bestaat uit een interval van 2 jaar en een leeftijdsrange van 40-74 jaar. Deze strategie leidde tot een 25% toename in het screeningseffect, maar tegelijkertijd ook tot een 44% toename in het aantal fout-positieve bevindingen waardoor de ratio van voor- en nadelen ongunstiger werd. Dit onderzoek laat zien dat het screenen op basis van het risico op borstkanker bij vrouwen met een verlaagd risico kan leiden tot een verlaging van de nadelen van screening terwijl een groot deel van het screeningseffect behouden blijft. Voor vrouwen met een verhoogd risico geldt dat er een winst in screeningseffect te behalen is, maar dat dit gepaard gaat met een sterk verhoogde kans op een fout-positieve bevinding. Toekomstig onderzoek is nodig om de impact van screenen naar risico op populatieniveau te bepalen, door de relatieve risico's te linken aan individuele risicofactoren en de bijbehorende prevalentie. Ook is het belangrijk om de borstdensiteit als risicofactor mee te nemen.

In **hoofdstuk 6** wordt weer een andere variatie op het huidige screeningsprogramma onderzocht, namelijk het gebruik van digitale borst tomosynthese als screeningtest in plaats van de huidig gebruikte digitale mammografie. Tomosynthese genereert meerdere borstfoto's, genomen uit verschillende hoeken, waardoor een pseudo 3D beeld wordt gecreëerd. Uit recente trials blijkt dat tomosynthese screening leidt tot een verhoogde borstkankerdetectie, vergeleken met digitale mammografie. In dit hoofdstuk wordt gekeken of screening met tomosynthese ook kosteneffectief is. Om dit te onderzoeken zijn 10.000 modelruns uitgevoerd voor een cohort vrouwen dat eens in de twee jaar wordt uitgenodigd voor screening met tomosynthese tussen leeftijd 50 en 74. De uitkomsten voor tomosynthese zijn vervolgens vergeleken met die van huidige screening (digitale mammografie tussen leeftijd 50 en 74). Het model voorspelde dat tomosynthese screening leidt tot een verhoging van 7% in het aantal gewonnen levensjaren, een verhoging van 6% in het aantal voorkomen borstkankersterfgevallen en een verlaging in het aantal fout-positieve bevindingen van 2%, per 1000 vrouwen die worden uitgenodigd voor screening en de rest van hun leven gevolgd worden, vergeleken met digitale mammografie. De gemiddelde incrementele kosten van tomosynthese over de 10.000 runs werden

geschat op €137.555 en de incrementele kosteneffectiviteitsratio van tomosynthese vergeleken met digitale mammografie op €27.023 per gewonnen levensjaar, waarbij de kosten en effecten beide met 3,5% per jaar zijn verdisconteerd. Ook lieten de uitkomsten zien dat de kans dat tomosynthese de meer kosteneffectieve optie is, vergeleken met digitale mammografie, 36% is, bij een drempelwaarde voor de kosteneffectiviteit van €20.000 per gewonnen levensjaar. Bij een hogere drempelwaarde van €35.000 per gewonnen levensjaar zou deze kans 66% zijn.

Uitgaande van de Nederlandse richtlijnen voor de drempelwaarde voor kosteneffectiviteit (€20.000 per gewonnen levensjaar) zou tomosynthese screening met een ICER van €27.023 niet als kosteneffectief worden beschouwd met een gelijke verdiscontering voor effecten en kosten van 3,5%. Echter, volgens de Nederlandse richtlijnen dienen effecten en kosten met respectievelijk 1,5% en 4% verdisconteerd te worden. Bij deze manier van verdisconteren zou de incrementele kosteneffectiviteitsratio een stuk gunstiger uitvallen.

Er bestaat nog behoorlijke onzekerheid rondom de kosten en effecten van borstkanker screening met tomosynthese. Aanvullend onderzoek waarbij de kosten van tomosynthese in een screeningssetting beter in kaart worden gebracht, waarin verschillende leesstrategieën worden onderzocht en waarin meerdere screeningsrondes worden meegenomen om de effectiviteit van tomosynthese beter te bepalen is daarom van belang.

Conclusies en aanbevelingen

De volgende conclusies kunnen worden getrokken op basis van de resultaten van de studies in dit proefschrift:

- De programmasensitiviteit met digitale mammografie is hoger dan met analoge mammografie als gevolg van de verhoogde borstkankerdetectie met digitale mammografie.
- Ondanks de verhoging van het verwijscijfer na de overgang naar digitale mammografie is de programmaspecificiteit in Nederland nog steeds relatief hoog vergeleken met andere Europese landen.

- De borstkankersterfte in Nederland is significant gedaald over de laatste decennia als gevolg van de invoering van borstkankerscreening en (verbeterde) adjuvante therapie.
- Het starten met screenen vóór de leeftijd van 50 jaar is een kosteneffectieve aanpak om het effect van screening te verhogen, maar heeft een hoger fout-positievencijfer als keerzijde.
- Indien algemeen geaccepteerd, heeft screening op basis van het risico van een vrouw de potentie om het effect van screening te verhogen voor vrouwen met een hoog relatief risico op borstkanker en om de nadelige effecten van screening te beperken voor vrouwen met een laag relatief risico, terwijl dit geen grote invloed heeft op de kosteneffectiviteit van het screeningsprogramma.
- Op basis van de huidig beschikbare data is niet te verwachten dat digitale borst tomosynthese een kosteneffectief alternatief is voor digitale mammografie in Nederland, maar deze conclusie kan in het voordeel van tomosynthese veranderen wanneer een hogere, internationaal geaccepteerde drempelwaarde voor de kosteneffectiviteit wordt gehanteerd, wanneer Nederlandse richtlijnen voor verdiscontering van kosten en effecten worden toegepast of wanneer tomosynthese goedkoper wordt.

Gebaseerd op bovenstaande conclusies zijn de volgende aanbevelingen opgesteld:

- Het schatten van de impact, wat betreft kosteneffectiviteit en de balans tussen voor- en nadelen, van screeningsstrategieën op basis van risico op borstkanker voor verschillende risicogroepen, voor de screeningspopulatie als geheel.
- Het vaststellen van de haalbaarheid en de acceptatie van screening op basis van het risico op borstkanker in Nederland.
- Bepalen hoe vrouwen het beste geïnformeerd kunnen worden over de voor- en nadelen van screenen op basis van het risico op borstkanker, met name wat betreft vrouwen met een relatief laag risico op borstkanker.

- Het kwantificeren van de mogelijke toename in overdiagnose met digitale borst tomosynthese, ten opzichte van digitale mammografie.
- Het vaststellen van de optimale leesstrategie met digitale borst tomosynthese in Nederland en het schatten van de bijbehorende kosten.
- Bepalen welk (aanvullend) onderzoek het meest van waarde zou zijn om een beslissing over het wel of niet invoeren van digitale borst tomosynthese te kunnen maken.
- Het opnemen van de vier verschillende categorieën van borst densiteit in het MISCAN model.
- Het gebruiken van risicofactoren op individueel niveau, gelinkt aan tumorprogressie, in het MISCAN model.

CURRICULUM VITAE

Valérie Sankatsing was born on the 14th of October 1986, in Nijmegen, the Netherlands. She completed her secondary education at the Stedelijk Gymnasium Nijmegen. In 2006, she started studying Biomedical Sciences at the Universiteit Utrecht. This was followed by a Master of Science in Infectious diseases at the Vrije Universiteit in Amsterdam in 2010. She wrote a Master's thesis about the quantification and effects of rotavirus-related pediatric hospitalisations and possible risk factors for nosocomial rotavirus infections, a research project at the Julius Center in Utrecht. In 2013, she started her PhD at the Department of Public Health at the Erasmus University Medical Center in Rotterdam, focusing on the evaluation of the Dutch breast cancer screening programme and the cost-effectiveness of breast cancer screening strategies. The results of this research are described in this thesis. Currently, Valerie is working as a researcher at Nivel, in Utrecht.

LIST OF PUBLICATIONS

Publications in this thesis

Valérie D.V. Sankatsing, Nicolien T. van Ravesteyn, Karolina Juraniec, Sabine E. Grimm, Manuela A. Joore, Harry J. de Koning.
Cost-effectiveness of digital breast tomosynthesis in the Dutch breast cancer screening program: a probabilistic sensitivity analysis. Radiology 2020;297(1):40-48.

Valérie D.V. Sankatsing, Nicolien T. van Ravesteyn, Eveline A.M. Heijnsdijk, Mireille J.M. Broeders, Harry J. de Koning.
Risk stratification in breast cancer screening: cost-effectiveness and harm-benefit ratios for low-risk and high-risk women. International Journal of Cancer 2020;147(11):3059-3067.

Valérie D.V. Sankatsing, Jacques Fracheboud, Linda de Munck, Mireille J.M. Broeders, Nicolien T. van Ravesteyn, Eveline A.M. Heijnsdijk, André L.M. Verbeek, Johannes D.M. Otten, Ruud M. Pijnappel, Sabine Siesling, Harry J. de Koning.
Detection and interval cancer rates during the transition from screen-film to digital mammography in population-based screening. BMC Cancer 2018;18(1):256.

Valérie D.V. Sankatsing, Nicolien T. van Ravesteyn, Eveline A.M. Heijnsdijk, Caspar W.N. Looman, Paula A. van Luijt, Jacques Fracheboud, Gerard J. den Heeten, Mireille J.M. Broeders, Harry J. de Koning.
The effect of population-based mammography screening in Dutch municipalities on breast cancer mortality: 20 years of follow-up. International Journal of Cancer 2017;141(4):671-677.

Valérie D.V. Sankatsing, Eveline A.M. Heijnsdijk, Paula A. van Luijt, Nicolien T. van Ravesteyn, Jacques Fracheboud, Harry J. de Koning.
Cost-effectiveness of digital mammography screening before the age of 50 in the Netherlands. International Journal of Cancer 2015;137(8):1990-9.

Publications not in this thesis

Sabine E. Grimm, Xavier Pouwels, Bram L.T. Ramaekers, Nicolien T. van Ravesteyn, **Valérie D.V. Sankatsing**, Janneke Grutters, Manuela A. Joore.
Implementation barriers to Value of Information analysis in health technology decision-making: results from a process evaluation. Submitted.

Valérie D.V. Sankatsing, H. Amarens Geuzinge, Jacques Fracheboud, Nicolien T. van Ravesteyn, Eveline A.M. Heijnsdijk, Lindy M. Kregting, Mireille J.M. Broeders, Johannes D.M. Otten, André L.M. Verbeek, Ruud M. Pijnappel, Arry E. de Bruijn, Harry J. de Koning: Landelijk Evaluatie Team voor bevolkingsonderzoek naar Borstkanker (LETB).
Landelijke evaluatie van bevolkingsonderzoek naar borstkanker in Nederland 2004 – 2014. Het veertiende evaluatierapport. LETB XIV 2019.

Fiona J. Gilbert, Fleur Kilburn-Toppin, **Valérie D.V. Sankatsing**, Harry J. de Koning.
Breast Cancer: Population and targeted screening. In: Cancer Prevention & Screening. Concepts, principles and controversies. Wiley Blackwell 2019;129-144.

Valérie D.V. Sankatsing, Harry J. de Koning.
Screening. In: ESMO Handbook of interpreting oncological study publications. ESMO Handbook Series 2018;13-28.

Valérie D.V. Sankatsing, Nicolien T. van Ravesteyn, Eveline A.M. Heijnsdijk, Harry J. de Koning.
Authors' reply to: "Questionable method for estimating the influence of mammography screening on breast cancer mortality in the Netherlands". International Journal of Cancer 2017;141(8):1709-1710.

PORTFOLIO

Seminars, conferences and courses	Year	Workload
Planning and Evaluation of Screening (NIHES course), Rotterdam, the Netherlands	2013	1.4
Breast Cancer Surveillance Consortium Working Group Meeting, San Francisco, United States.	2013	0.7
International Workshop on Breast Cancer Risk Assessment, San Francisco, United States.	2013	1.0
Genetic Risk Prediction for Common Diseases, VUMC, Amsterdam, The Netherlands	2013	0.2
Master class: Advances in Epidemiologic Analysis (NIHES), Erasmus MC, Rotterdam, the Netherlands	2014	0.4
Master class: Advances in Genomics Research (NIHES), Erasmus MC, Rotterdam, the Netherlands	2014	0.4
CPO Course, Erasmus MC, Rotterdam, the Netherlands	2015	0.3
International Cancer Screening Network meeting, Rotterdam, the Netherlands	2015	0.7
Scientific Career Orientation, Erasmus MC, Rotterdam, the Netherlands	2015	0.1
Genetic Screening: Who, Why and When?, Amsterdam, the Netherlands	2015	0.3
Scientific Integrity course, Erasmus MC, Rotterdam, the Netherlands	2015	0.3
Werkgroep 'Wetenschap en Innovatie' RIVM, Utrecht, the Netherlands	2015	0.1
Presenting for Scientists, Erasmus MC, Rotterdam, the Netherlands	2016	0.2
Symposium The Impact of Breast Imaging, Utrecht, the Netherlands	2016	0.2
WEON Conference, Wageningen, the Netherlands	2016	1.0
Werkgroep 'Wetenschap en Innovatie' RIVM, Utrecht, the Netherlands	2017	0.1
Congres 'Bevolkingsonderzoeken op het spoor', Utrecht, the Netherlands	2017	0.3
Advanced Decision Modeling course (NIHES)	2018	1.4

Seminars, conferences and courses	Year	Workload
Symposium Dutch Association for Community Genetics and Public Health Genomics Meeting, Utrecht, the Netherlands	2018	0.2
International Cancer Screening Network Meeting, Rotterdam, the Netherlands	2019	0.7
Research Seminars, Erasmus MC, department of Public Health, Rotterdam, the Netherlands	2013-2019	3.7
National Evaluation Team Breast cancer screening meetings, Utrecht, the Netherlands	2013-2019	1.0
Presentations and posters		
Monthly seminar 'Club Meth', Rotterdam, the Netherlands. Oral presentation.	2013	1.0
European Breast Cancer Conference (EBCC-9), Glasgow, Scotland. Poster presentation.	2014	1.0
Research group meeting 'Early Detection', Rotterdam, the Netherlands. Oral presentation.	2014	1.0
Research group meeting 'Early Detection', Rotterdam, the Netherlands. Oral presentation.	2015	1.0
National Evaluation Team Breast cancer screening meeting, Utrecht, the Netherlands. Oral presentation.	2015	1.0
Research group meeting 'Early Detection', Rotterdam, the Netherlands. Oral presentation.	2016	1.0
Research group meeting 'Journal club', Rotterdam, the Netherlands. Oral presentation.	2016	1.0
WEON Conference, Wageningen, the Netherlands. Oral presentation.	2016	1.0
Working Group meeting RIVM 'Wetenschap en Innovatie', Utrecht, the Netherlands. Oral presentation.	2017	1.0
Meeting RIVM, ZonMW and Dutch Health Council 'VOI tomosynthesis', Utrecht, the Netherlands. Oral presentation.	2019	1.0
International Cancer Screening Network Meeting, Rotterdam, the Netherlands. Oral presentation.	2019	1.0

Teaching		
Community Project Mentor: supervising a group of third year Medicine students in a project on 'Breast cancer and deodorants', Erasmus MC, Rotterdam, the Netherlands.	2017	0.7
Co-supervising Econometrics master student	2018	1.5

Other		
IARC Handbooks of Cancer Prevention Volume 15. Writing of two chapters.	2014	1.5
Contributing to Screening Evaluation part of MGZ website	2016	0.1
Writing report on regional breast cancer mortality in the Netherlands for the RIVM	2016	1.0
Writing report on model calibration for the RIVM	2016	1.0
Organising Scientific Career Orientation day for the Public Health department, Rotterdam, the Netherlands	2017	0.2
Cancer Prevention and screening: concepts, principles and controversies, WILEY Blackwell. Writing of one chapter.	2014-2018	1.5
ESMO Handbook Interpreting Oncological Study Publications, ESMO Handbook Series. Writing of one chapter.	2016-2018	1.5
Fourteenth Evaluation Report National Evaluation Team Breast cancer screening, RIVM. Coordination of report and writing of two chapters.	2018-2019	2.0
Reviewing manuscripts for journals (for example Journal of Medical Screening)	2019	0.9
Total		37.6

DANKWOORD

Ik had nooit gedacht dat ik op mijn werkplek zoveel bijzondere mensen zou leren kennen. Er zijn een hoop collega's die ik dankbaar ben. Een aantal collega's wil ik hieronder in het bijzonder bedanken.

Dit proefschrift was niet tot stand gekomen zonder de goede begeleiding van mijn directe collega's.

Allereerst bedank ik graag mijn promotor, Harry. Bedankt voor het vertrouwen dat je in me had de afgelopen jaren, voor alle kansen die je me hebt geboden en de vrijheid die je me hebt gegeven om het onderzoek uit te voeren. Ik heb tijdens mijn ontwikkeling als onderzoeker veel geprofiteerd van jouw kennis en kritische blik. Daarnaast was het leuk om Japan-tips met elkaar uit te kunnen wisselen!

Nicolien, ik heb veel gehad aan onze wekelijkse meetings en je hebt dan ook een enorme bijdrage geleverd aan dit proefschrift. Ik heb heel veel geleerd van jouw kennis, inzicht en scherpe kijk op dingen, waardoor ik ook meer vertrouwen in mezelf heb gekregen. Ook waardeer ik het heel erg dat je, ondanks je steeds voller rakende agenda, altijd voor me hebt klaargestaan en betrokken bent geweest. Ik vond het een hele prettige samenwerking!

Eveline, ook jouw aandeel in dit proefschrift is heel groot. Je hebt me op een zeer efficiënte manier kennis laten maken met de wondere wereld van MISCAN. Ik heb veel gehad aan jouw ondersteuning, niet alleen bij modelanalyses, maar ook op het gebied van wandeltips! Bedankt voor de fijne begeleiding.

Arry, de screeningssectie kan onmogelijk zonder jou! Ik vond het heel fijn dat ik altijd bij je terecht kon met vragen over van alles en nog wat.

Ook andere collega's met wie ik direct heb samengewerkt wil ik graag bedankten. Jacques, ik vond het heel fijn dat ik altijd bij je terecht kon met mijn vragen over de data van het IKNL, jouw zeer uitgebreide en complexe (!) Excel bestanden en met name ook om gewoon gezellig even bij te kletsen! Karolina, you made performing

a PSA so much easier! I really enjoyed our night out in Rotterdam at the end of your internship. Verder wil ik alle leden van het LETB bedanken voor de goede suggesties en adviezen over meerdere manuscripten in dit proefschrift. Tenslotte wil ik ook alle coauteurs die mee hebben gewerkt aan mijn manuscripten graag bedanken voor alle goede en leerzame feedback.

Graag wil ik ook de leden van mijn promotiecommissie bedanken. De kleine commissie voor het beoordelen van mijn proefschrift en alle leden voor de bereidheid en interesse om hierover met mij van gedachten te wisselen.

Uiteraard ben ik ook al mijn kamergenootjes dankbaar, in het bijzonder die van mijn laatste kamer: Lea en Maryse en flex-kamergenootjes Nikki en Catherine. Ik vond het erg gezellig om mijn werkruimte met jullie delen. We hebben lief en leed gedeeld en vooral ook heel veel gelachen! Bedankt Lea en Catherine voor alle gezelligheid en de Warung Mini avonden (en soms ook middagen). Het Surinaamse eten in Rotterdam is een stuk beter dan in Utrecht! En bedankt Lea en David voor de vele sushi momenten. Ik ben blij dat ik jullie 'echte' Rotterdammers heb mogen leren kennen!

Ook Andrée en Suzanne wil ik bedanken voor (weer) vele sushi etentjes, de geniale lemon daiquiri's bij Ayla en onze fijne gesprekken!

Amarens, Brechtje, Carlijn, Lisanne, Lucie en Sabine, alias 'De eetclub', Arthur, Arne, Niels, Erik en Jeroen. Ik wil jullie graag bedanken voor alle leuke dingen die we de laatste jaren ondernomen hebben. Het waren stuk voor stuk super gezellige etentjes (met als hoogtepunt de High Friet in Breda (nooit meer!)), borrels en avondjes uit (in Rotterdam, maar natuurlijk ook carnaval!).

Mirjam en Nadine, onze koffiemomenten waren altijd erg gezellig! Ik kon heel fijn met jullie praten en heb veel geleerd van hoe jullie in het leven staan.

Timor, bedankt dat je me door de vierdaagse hebt heen gesleept! Na de zoveelste 50 km zat het soms tegen, maar we hebben het toch maar mooi gehaald samen!

Lisanne, toen we elkaar nog maar net kenden voelde het al als een jarenlange vriendschap. We lijken erg op elkaar wat eigenlijk best wel prettig is. Ons weekje weg was dan ook heel gezellig!

Amarens, ook met jou voelde het meteen heel vertrouwd. Ik ben heel blij met onze vriendschap en dat we zo dicht bij elkaar wonen, waardoor we regelmatig gezellig samen kunnen eten. Ook vind ik het heel fijn dat je mijn paranimf bent en me tijdens deze bijzondere (en spannende) periode steunt!

Verder wil ik ook de hele screensectie bedanken en in het bijzonder de borstkankerscreeningsgroep. Bedankt voor de gezelligheid tijdens onze borstkanker lunches en diners!

Ik heb het erg naar mijn zin gehad op de afdeling en wil tenslotte daarom alle overige MGZ-ers bedanken voor de praatjes in de gang of bij de koffieautomaat, de gezellige lunchpauzes, en de vele etentjes en borrels!

Ook buiten mijn collega's zijn er een hoop mensen die ik graag wil bedanken. Lieve vrienden uit Nijmegen, Utrecht en van mijn studie; De Lloret groep, Ca & Su, GEISV, de Gin-Tonic club en anderen. De meeste van jullie ken ik inmiddels al een hele tijd en we hebben dan ook veel samen meegemaakt. Bedankt dat jullie er voor me zijn geweest door samen leuke dingen te doen, goede adviezen te geven of gewoonweg door te vragen hoe het met me ging.

Lieve vrienden/buurtjes, Marcel en Jacqueline, het heeft echt een meerwaarde om naast elkaar te wonen! Ik ben heel blij dat we in deze rare tijd vol met maatregelen één huishouden vormen :)

Lieve familie, bedankt voor alle ondersteuning en gezelligheid. De neevinnen, Amber, Iris, ik beschouw jullie meer als halfzussen dan als nichtjes. We hebben heel veel met elkaar meegemaakt en zijn zó op elkaar ingespeeld dat niemand onze gesprekken begrijpt :) Isabelle, ik vind het heel leuk dat jij nu ook met onze uitjes

mee gaat. Ik heb zelden zo gelachen als toen we met z'n allen in het rariteiten museum waren.

Vandana, bedankt voor alle gezellige lunch- en koffie-uitjes. Wat toevallig dat jij aan de overkant van de straat werkte! De etentjes bij jullie thuis met Raaj, Karan en Shirani vond ik ook altijd erg leuk. Alleen wel jammer dat ik daarna nooit thuis kon komen zonder midden in de nacht op een of ander station te stranden!

Lieve kleine, grote Camille, mijn kleine zusje maar stiekem al een volwassen vrouw. Ik vind onze band heel bijzonder en ik ben erg trots op je! En natuurlijk ben ik ook heel blij dat jij mijn Paranimf bent!

Lieve papa en mama, ik wil jullie graag bedanken voor jullie onvoorwaardelijke steun, jullie staan altijd voor me klaar. Ook geniet ik erg van het lekkere eten dat jullie altijd voor me meebrengen (wat jammer dat ik jullie kookkunsten niet heb geërfd!) en van onze vele etentjes met z'n allen in in De Plak, Merano, Hygge etc.

Lieve Jeffrey, een leven zonder jou kan ik me niet voorstellen. We zijn al 12 jaar samen en hebben het nog steeds iedere dag gezellig met elkaar. Je geeft me heel vaak wijze en heldere adviezen, waar ik veel aan heb, en je helpt me altijd relativeren. Ik kijk erg uit naar ons verdere leven samen!