Health Council of the Netherlands

Population screening for breast cancer: expectations and developments

Gezondheidsraad
Health Council of the Netherlands

To the Minister of Health, Welfare and Sport

Subject : presentation of advisory report *Population screening for breast cancer: expectations and developments*
Your reference : PG/OGZ 3117093
Our reference : I-1272-12/LvR/pm/894-A65
Enclosure(s) : 1
Date : January 22, 2014

Dear minister,

In response to your request for advice dated 15 June 2012, I hereby submit the report Population screening for breast cancer: expectations and developments. The advisory report has been compiled by the Committee on Population Screening and reviewed by the Standing Committee on Medicine, the Standing Committee on Public Health and a number of external experts.

There is continuing controversy about the effectiveness and efficacy of breast cancer screening. Some believe that the efficacy is not as great as had been anticipated, or even marginal, while the harmsare considerable. Others argue that breast cancer screening is in fact very effective and could be more effective still if the programme were intensified.

I therefore conclude that the in-depth, methodological analysis that the Committee performed was very important. Having done so, the Committee has concluded that the Netherlands has a long-term and effective population based screening programme for breast cancer, which – despite changing circumstances – continues to satisfy expectations. According to the Committee, in part the success of the Dutch population screening programme is due to the high quality and organizational efficiency of the programme. For example, in the Netherlands, false positives and overdiagnosis are less common that in some other countries, while the number of false negatives is barely any higher.

In the short term, the Committee sees no reason for major changes to the programme, such as adjusting the age limit of the target group. Nevertheless, the Committee does believe there is room for improvement in the breast cancer screening programme. Its

P.O.Box 16052
NL-2500 BB The Hague
The Netherlands
Telephone +31 (70) 340 67 34
E-mail: l.van.rossum@gr.nl

Visiting Address
Rijnstraat 50
NL-2515 XPThe Hague
The Netherlands
www.healthcouncil.nl

Gezondheidsraad
Health Council of the Netherlands

Subject : presentation of advisory report *Population screening for breast cancer: expectations and developments*
Our reference : I-1272-12/LvR/pm/894-A65
Enclosure(s) : 1
Date : January 22, 2014

recommendations focus primarily on minimizing the harms inherent to such programmes. For example, women who receive false positive results often unnecessarily undergo additional (outpatient) check-ups. The Committee recommends investigating ways of providing better support and ways of avoiding unnecessary extra check-ups for women receiving false positive results. The screening programme could also be improved by developing a separate follow-up pathway for women whose results indicate a small risk of breast cancer. For such women, further mammograms or ultrasound scans are usually sufficient; it is not always necessary to perform a biopsy.

I endorse the Committee's conclusions and recommendations.

Yours sincerely,
(signed)
Professor W.A. van Gool,
President

P.O.Box 16052
NL-2500 BB The Hague
The Netherlands
Telephone +31 (70) 340 67 34
E-mail: l.van.rossum@gr.nl

Visiting Address
Rijnstraat 50
NL-2515 XPThe Hague
The Netherlands
www.healthcouncil.nl

Population screening for breast cancer: expectations and developments

to:

the Minister of Health, Welfare and Sport

No. 2014/01E, The Hague, January 22, 2014

The Health Council of the Netherlands, established in 1902, is an independent scientific advisory body. Its remit is "to advise the government and Parliament on the current level of knowledge with respect to public health issues and health (services) research..." (Section 22, Health Act).

The Health Council receives most requests for advice from the Ministers of Health, Welfare & Sport, Infrastructure & the Environment, Social Affairs & Employment, Economic Affairs, and Education, Culture & Science. The Council can publish advisory reports on its own initiative. It usually does this in order to ask attention for developments or trends that are thought to be relevant to government policy.

Most Health Council reports are prepared by multidisciplinary committees of Dutch or, sometimes, foreign experts, appointed in a personal capacity. The reports are available to the public.

The Health Council of the Netherlands is a member of the European Science Advisory Network for Health (EuSANH), a network of science advisory bodies in Europe.

This report can be downloaded from www.healthcouncil.nl.

Preferred citation:
Health Council of the Netherlands. Population screening for breast cancer: expectations and developments. The Hague: Health Council of the Netherlands, 2014; publication no. 2014/01E.

ISBN: 978-94-6281-011-2

Contents

Executive summary

Randomized controlled trials, conducted in the 1970s and 1980s, have shown that mammographic screening can reduce breast cancer mortality significantly. In the period from 1989 to 1998, the Netherlands introduced a national screening programme to detect early-stage breast cancer. In 1990, the expectation was that ultimately (by around 2015) 700 fewer women in the Netherlands would die of breast cancer each year. Today, twenty-five years on, there is a diversity of opinions about the benefits and harms of breast cancer screening varying from substantial mortality reduction to almost no screening effect. Therefore, the Minister of Health, Welfare and Sport wants to know how effective population screening for breast cancer in the Netherlands really is.

The effectiveness and efficacy of breast cancer screening

The Minister's question is not easy to answer. The sheer complexity of population screening makes it difficult to evaluate. A programme's benefits and harms can be estimated only after a sufficient follow-up time. In addition, population screening for breast cancer was introduced at a time of fundamentel change in the treatment of breast cancer. Since then, the treatment of breast cancer has significantly improved, and is still rapidly developing. Moreover, women have become more aware of breast cancer, and consult their GP quicker than before. The situation today is very different from what it was 25 years ago.

The breast screening debate has not helped to have a clear view of the efficacy of mammographic screening. Some argue that the benefit of screening is overstated and that the overdiagnosis associated with screening is understated. They question if the marginal benefit, really compensates for the drawbacks (harms) involved in such programmes. Others believe that population screening actually yields important health gains. The conclusions vary widely. Why is this? According to the Committee, the explanation originates primarily in the area of research methodology.

Benefits of population screening

Given the continuing controversy about the effectiveness of breast cancer screening, the Committee has delved even deeper than usual into the various types of evaluation methodology. It was concluded that trend studies generally do not allow solid conclusions. However, screening effects can be measured by means of well-designed cohort studies and case-control studies. This approach makes it possible to distinguish a screening effect from other factors that affect breast cancer mortality, such as improved treatment.

Based on evidence from the most reliable cohort studies, a breast cancer mortality reduction of 26 per cent can be expected in women aged 50-69 offered service screening (who will not all participate). Case-control studies determine the relationship between actual participation in screening and breast cancer mortality. The results consistently show that participation in screening is associated with a significant reduction in breast cancer mortality.

In the Netherlands between 1986/1988 and 2012 breast cancer mortality (European standardized rate) declined with 34 per cent to 62 per 100,000 in the 50 to 75 age group. It is difficult to determine what part of this 34 percent decline is attributable to population screening. Based on computer modelling, it is estimated that about half of this decline is due to population screening, and the remainder to improved treatment.

The conclusion is that the effectiveness of population screening for breast cancer in the Netherlands continues to meet the initial expectations, even though circumstances have changed. The question whether the benefits of such population screening outweigh the ever-present harms of screening still remains.

Harms of population screening

For optimal image-quality of the mammogram, the breast must be firmly compressed. Half of the participants find this unpleasant or even painful. For

every thousand participants, there are seventeen women who are told that they may have breast cancer, while this later turns out not to be the case (false positives). This causes – retrospectively – unnecessary anxiety. Screening does not detect all cancers. About 2 in 1,000 participants with a non-suspect screening result are diagnosed with breast cancer in the period of two years between the screening intervention (false negatives). It also happens that cases of breast cancer are detected that would never have been identified clinically in the lifetime of a woman without screening ("overdiagnosis"). For each individual case, it is impossible to know whether or not overdiagnosis is involved, so treatment is routinely given. As a result, the women in question experience all of the harms but none of the benefits of early diagnosis and treatment.

The frequency of overdiagnosis can only be estimated at population level over a period of several years. The estimated percentages vary considerably, ranging from 0 per cent to more than half of all cases of breast cancer diagnosed in the target population. This has generated considerable debate. Confusion arises from the use of different definitions. Another important factor is the length of the observation period, because screening on average allows diagnoses to be made two to three years earlier (the "lead time"). As a result, cases of breast cancer are diagnosed more frequently during the introductory stage of a population screening programme (the "prevalence peak"). Conversely, after the age of 75, when women are no longer invited for screening, fewer cases of breast cancer are diagnosed than are usual invited for. This compensatory drop can only fully be estimated if all women in the target population are no longer invited for screening.

This means that the level of overdiagnosis can only be estimated after a long period of follow-up. Researchers who only take the introductory stage into consideration, without allowing for the subsequent compensatory drop or without correcting for lead time, overestimate the level of overdiagnosis. Suitable studies tend to produce significantly lower figures, ranging from one to ten percent overdiagnosis relative to the expected incidence of breast cancer in the absence of screening. The three percent estimate for the Netherlands is within this range corresponding with eight percent of screen detected breast cancer cases. The conclusion is that overdiagnosis does indeed occur, but not to the extent often suggested.

Risk-benefit ratio

How do the benefits of the Dutch current screening programme weigh up against the harms? Modelling (with MISCAN) shows that population based screening

leads to on average 775 less deaths from breast cancer annually. Around 1,200 women need to be screened to prevent one death from breast cancer. For each prevented death from breast cancer, 23 women will be referred every year. Of these, 16 will have a false positive screening result. Of each seven true positive results, 5 experience no health gains (with the possible exception of having less invasive treatment as a result of early diagnosis). Of these 5 true positives, 0.9 women will die from breast cancer despite having participated in the screening programme 3.7 would have also survived without screening, and 0.5 women up would never have had a diagnosis of "breast cancer".

Cost-effectiveness analysis shows that the costs of the Dutch population screening amounts to EUR 1,600 per life year gained. For each breast cancer death prevented by screening, the woman will be spared the terminal stages of this disease, in question and she will gain an average of 16.5 life years.

The claim that screening generates more overdiagnosis than health benefits does not apply to population screening in the Netherlands. There are no compelling scientific reasons to terminate this screening programme.

Adaptations in the current screening programme

The Minister asked what adaptations and developments might further improve this population screening programme. First an alternative referral pathway might be developed for women with BI-RADS 0 screening results, with a small risk of breast cancer. For 60 per cent of the participants with BI-RADS 0, imaging techniques (mammography and ultrasound) alone are sufficient to exclude breast cancer, without further invasive diagnostic procedures. Rapid diagnostic imaging would help to alleviate much of the fear and anxiety caused by the screening result.

Evidence suggests that women with benign breast lesions after a (false positive) screening, often undergo further outpatient check-ups. Is this due to persisting anxiety? In which case, is the offer of a follow-up appointment an adequate response? Or might this be a task for the GP? Further research could improve the management of such women.

Several studies, currently performed in the Netherlands, may identify ways to perform mammopraphy with less pain, without impairing the image-quality affecting the radiation dose.

Developments with potential promise in the medium-term

The current screening programme offers the same service to all women in the target group. According to the Committee, there are no compelling scientific reasons for modifying either the target age group, the screening interval, or the screening method. One appealing way of improving the risk-benefit ratio of screening would be to adapt this process to the women's individual breast cancer risk. Extensive ongoing research focuses on improving the discrimination between high and low risk groups. There are also questions regarding the logistics of risk stratification in the context of screening, and the acceptance and the effects of providing intensive screening (younger starting age, additional screening method) to the high-risk group and less intensive screening to the low-risk group.

A new technique, tomosynthesis, can supplement regular two-dimensional mammograms by constructing three-dimensional images of the breast. Tomosynthesis offers a great promise, as a method for improving cancer detection and reducing the number of false positives. As yet its drawbacks are: higher and longer radiation exposure, longer compression time and therefore longer lasting pain and longer interpretation time. A number of major issues remain to be resolved before population screening trials can be considered to assess this technique.

Conclusion

The Committee concludes that population screening remains worthwhile. Anually, it prevents on average 775 breast cancer deaths. The main disadvantage of screening is overdiagnosis, which will occur in about 8 per cent of screen detected breast cancer cases. This is a substantial disadvantage, but not to the extent that is often suggested. Dutch population screening stands out in terms of its high participation and low referral combined with a reliable test performance. Further improvements are expected in the near future, involving new techniques that make the procedure less painful and further reduce the radiation dose by half in mammographic screening (which is already low). In the long run, it may be possible to improve the efficiency of the population screening programme by tailoring the screening to the individual's estimated risk of breast cancer.

Chapter 1

Introduction

In 1989, the Netherlands began the introduction of a population based breast cancer screening programme. Initially, only women between the ages of fifty and seventy were invited for screening, but in 1998 the programme was extended to include women aged between seventy and seventy-five. Since then, a great deal has happened. There has been further professionalization of the organization and implementation, links with the health care system have been improved and the mammography process has been digitized. In parallel with the introduction of population screening, there have been various other important developments, the most significant being advances in breast cancer therapy, with the collective outcome that breast cancer mortality has decreased considerably.

Screening often generates high expectations and reinforces the widespread desire for reassurance regarding one's health. The demand for screening is increasing.[1] People are often unaware that screening offers real health benefits for relatively few participants and does have drawbacks (harms), such as false positive results. It is even possible for screening to completely fail in its purpose, as was the case with the national tuberculosis screening programme, in which millions of people underwent radiological examination, without any impact on the risk of 'open' tuberculosis.[2]

As indicated in the request for advice, it is necessary to evaluate the benefits and harms of population based screening with regular intervals. Regular evaluation provides an opportunity to consider whether, in the light of new research findings, a screening programme should be modified or even ended.

There is certainly a need for critical evaluation of population based screening for breast cancer, as (justified or not) it's efficacy has been questioned.[3-6] The balance between the benefits and harms is critical when it comes to deciding whether or not to offer population based screening.

1.1 Request for advice

The initial decision for the Dutch population based breast cancer screening programme was based upon a 1987 Health Council advisory report on the acceptability and desirability of such a programme.[7] The Council subsequently evaluated the programme on relevant issues in 2002 and 2006.[8,9] On 15 June 2012, the Minister of Health, Welfare and Sport asked the Health Council to update its advisory reports (Annex A). Specifically, the Council was asked:

- to comment on the efficacy, or the balance between the benefits and harms of the existing programme in the Netherlands
- to consider whether the programme might be optimized by revising the screening strategy (screening method, age limits of the target group, screening interval) or by applying risk stratification.

At the request of the President of the Health Council, the Council's Committee on Population Screening (Annex B) compiled this report. The text was prepared by W.A. van Veen, a physician and member of the Committee.

1.2 Structure of this report

Chapter 2 contains statistics on breast cancer and population screening in the Netherlands. In Chapter 3, the Committee describes what was initially expected from population based screening for breast cancer in the Netherlands, in reference to the results of the long-term screening trials. Chapter 4 outlines the changes in the background situation that have been taking place since the screening trials, where they may be relevant to assessment of the efficacy of screening. That is followed, in chapter 5 and 6, by discussion of the research into the benefits and harms of population based screening. Chapter 7 is devoted to the efficiency and cost-effectiveness of breast cancer screening and addresses the Minister's question about the overall efficacy of the Dutch existing population based breast cancer screening programme. The section ends with the Committee's conclusion on that question.

Chapter 8 highlights the developments that could improve the screening programme in the short term. That is complemented by Chapter 9, which deals

with developments that may make breast cancer screening more effective and more efficient in the long term.

In Chapter 10, the Committee answers each of the questions contained in the Minister's request for advice. Finally, the Committee's recommendations are presented in Chapter 11.

Annex C lists the experts consulted, in the Dutch report (not in English) Annex D contains a glossary of terminologies used in the report and Annex E summarizes the main results of population screening in the Netherlands in 2012 and in the period 1998 to 2010.

Chapter 2

Breast cancer and population screening

2.1 Incidence and mortality

Breast cancer is a malignant condition originating in the glandular tissue of the breast. If the tumour is confined to the place of origin, within the basal membrane, it is called *in situ* cancer. The place of origin is typically the ducts of the glandular tissue, in which case the cancer is termed a ductal *in situ* carcinoma (DCIS). However, by the time of diagnosis, the tumour has usually become 'invasive'. In 2011, 14,070 women in the Netherlands were diagnosed with invasive breast cancer and a further 2,049 women with *in situ* breast cancer.

The average age at which breast cancer is diagnosed is sixty-one. The five-year survival rate following diagnosis is 82 per cent. In women diagnosed in the recent period (2006 to 2010), the five-year survival rate is 86 per cent (www.cijfersoverkanker.nl, 1 Aug 2013). In 2011, a total of 70,482 women died in the Netherlands, of whom 3,261 (4.6 per cent) died of breast cancer: 39 per 100,000 women. In 2012, the age-standardized breast cancer mortality (*European standardized rate*) in fifty to seventy-four year-olds was 62 per 100,000 women: 34 per cent lower than in the period 1986 to 1988, before the introduction of population based screening. More than half of the mortality reduction is attributable to population based screening, the rest to improved therapy.[10]

2.2 Population screening

History

The aim of population screening for breast cancer is to reduce breast cancer mortality by detecting the disease at an early stage, thus facilitating treatment. The method of early detection (screening) employed in the Dutch programme is mammography: radiography of the breasts. Experience with mammographic screening was first gained in pilot projects in Utrecht and Nijmegen starting in 1974, after which the national rollout of breast cancer screening for women between fifty and seventy years of age began in 1989. Women in the target group are called up for screening every two years. The introduction process was completed at the end of 1997. In 1998, the target group was enlarged to include women aged seventy to seventy-five. The national programme provides screening for women of fifty to seventy-five years of age only. Women outside that age group, who wish to have their breasts checked for cancer, can consult their GPs.

In 1997, it was concluded that the implementation phase of the screening programme had been successful. However, some aspects gave cause for concern. The programme sensitivity (i.e. the sensitivity based on a screening interval of two years and mammography as the method of screening) was lower than expected (given constant background incidence of breast cancer over time); it was found to be 71.5 per cent at a specificity of 99.1 per cent[11]. Furthermore, considerable inter-regional differences were found in the detection and referral rates. The referral rate in the Netherlands (the percentage of participants with abnormal screening results who are referred to a breast clinic for further diagnostic testing) was less than 1 per cent, i.e. the lowest in the world. Typical international referral rates were between 4 and 10 per cent.[12-17] It was therefore decided to carry out research to determine explanations for the low referral rate and to identify possible optimizing strategies.

The resulting ‘Optimization Study’ led to recommendations and clearer arrangements regarding the ‘blind’ reading of mammograms by two screening radiologists (*double reading*), regarding the procedure to be followed in the event of the two radiologists reaching divergent conclusions (decisive review by a third screening radiologist), the referral criteria and the revision of interval carcinomas.[18] The Optimization Study concluded that lowering the referral threshold would increase the breast cancer detection rate.[19] In accordance it was recommended that a lower referral threshold should be adopted.

The recommendations arising from the Optimization Study were communicated during routine inspections and on refresher training courses, resulting in the rate of referral more than doubling (Annex E, Table 2). Moreover, the detection rate also increased, augmented by making two mammograms per breast more often and by the introduction of digital mammography screening. The better image quality of digital mammography enhances the visibility of certain abnormalities, such as microcalcifications. Since June 2010, the Dutch breast cancer screening programme has been fully digitized.[11]

Current situation

There are now 68 screening units, where on average more than 14,000 women are screened each year. Within a few days, the mammograms are assessed independently by two radiologists, specifically trained for screening. If they reach different conclusions, the system alerts the second reader, who may reconsider his or her assessment. If consensus is not reached, the case is referred to a third radiologist for arbitration.

Within two weeks of the mammogram, a participant receives her result in writing. If no suspect abnormality has been detected, she is reminded that screening cannot guarantee the absence of disease and is indicative of health status only at a particular moment in time. The woman is urged to consult her GP if she experiences any problems.

If screening does reveal an abnormality that could indicate breast cancer, additionally the woman's GP is informed of the result and asked to contact the patient. The day after the GP (but not immediately before the weekend) the woman herself receives a letter, informing her that the screening revealed a possible abnormality, which requires further investigation. The woman is advised to consult her GP immediately. She is provided with information about the follow-up diagnostic testing and, again, told that the presence of an abnormality does not necessarily mean that she has breast cancer.

The information sent to the woman draws attention to the *Breast Cancer Care Monitor*. The *Monitor* can be useful in the event of specific referral to a breast clinic (a clinic with a team of specialists in the diagnosis and treatment of breast conditions). Great emphasis is placed on minimizing the intervals between screening and diagnosis, and between diagnosis and treatment. For example, a project called MammoXL is currently in progress with the aim of minimizing the time required to make screening mammograms available in the hospital. The digital interface between the screening centres and the care providers should ensure better, earlier availability, with the added benefit that no unnecessary new

mammograms need to be made at the hospital. Furthermore the project aims to expedite the availability of the hospital's follow-up data to the screening organizations for use in monitoring and quality assurance (www.mammoxl.nl).

To avoid misunderstandings about the nature and seriousness of the referral, the letter to the GP states the result of the screening using the Breast Imaging-Reporting and Data System (BI-RADS) developed by the American College of Radiology.[20,21] BI-RADS has been adapted for use in the screening programme and provides insight into the likelihood of breast cancer being present (the PPV of a referral). This enables GPs to prepare their patients better for what lies ahead, depending on the level of suspicion of cancer. Furthermore, use of this system facilitates the identification of regional variations in referral strategy and may over time provide a basis for adjustments to the general referral pattern, since more than half of the women who are referred for further examination can now be returned to the screening programme on the basis of imaging techniques alone.

In 2012, 24 in every thousand participants were referred to breast clinics by their GPs for further diagnostic testing (Annex E, Table 2).[23] The Breast Cancer Guideline published by NABON (The National Breast Cancer Forum) states that a woman should be seen within five days of referral and that 90 per cent of women seen should receive a result and a treatment plan within seven days of their first outpatient visit (www.oncoline.nl). According to the most recent available information, the latter interval currently averages 7.9 calendar days (NABON Guidelines, evaluation 2007) and the waiting time for surgery usually is four weeks. In the south-eastern part of the Netherlands, the waiting time in 80 per cent of cases was up to five weeks from the date of diagnosis.[22] SONCOS (the Oncology Cooperation Foundation) is preparing a uniform multidisciplinary quality framework for oncological care, aiming to assure high-quality transparent oncological care focussing on the patient's interests. Screening organizations follow the referred women, until further examination has taken place.

In 2012, nearly 1.3 million women received an invitation for screening (Annex E, Table 2).[23] The invitation states a date and time for the woman to attend for screening, but comes with an opportunity to change this pre-set appointment if needed. For a long time, participation in the programme increased every year, reaching more than 82 per cent in 2007. Since then, participation has declined slightly, to approximately 80 per cent in 2012.[23] In comparison the decline is most marked amongst women called up for the first time. Of those who had previously been invited, 75 per cent received a repeat invitation 'on time' (within a period of 22 to 26 months; see also subsection 8.4). However, in 2010

at least, there was considerable regional variation in the percentage of women called up 'on time', ranging from 41 to 91 per cent.[11]

Women who have breast prostheses are able to participate in the screening. Although breast augmentation surgery is now relatively commonplace in countries such as the United States, very few recent incidents have been reported of an implant being ruptured by mammography. Where appropriate, the degree of compression for the mammogram can be adjusted. Proper assessment is nearly always possible because sufficient breast tissue is imaged. If that is not the case, the woman is advised to consult her GP, who is sent a letter containing specific recommendations.

Population screening is coordinated at the national level on behalf of the Ministry of Health, Welfare and Sport by the Centre for Population Screening (CvB) at the National Institute of Public Health and the Environment (RIVM). The Breast Cancer Screening Programme Committee established by the RIVM advises the CvB on national coordination matters. Regional implementation is in the hands of five screening organizations, which, since 2011, have also been responsible for the training (including in-service and refresher training) of the screening personnel. The National Expert and Training Centre for Breast Cancer Screening (LRCB) in Nijmegen monitors the quality of the training and the (para)medical and physical-technical quality of the population screening programme. The National Evaluation Team for Breast Cancer Screening in the Netherlands (LETB), whose secretariat is at the ErasmusMC in Rotterdam, undertakes annual audits and periodic evaluation. Auditing and evaluation are based on annually submitted, (sub)regional aggregated data, supplemented by data from Statistics Netherlands (CBS) and the Comprehensive Cancer Centres (IKNL).

The population screening programme for breast cancer is funded from the national budget. In 2012, the cost was 64.6 million euros, which equates to 64 euros per examination.

Population screening in other countries

In 2003, the Council of the European Union made a number of recommendations regarding screening for breast, cervical and bowel cancer.[24] The Council advised mammographic screening of women aged 50 to 69 for breast cancer, in accordance with European quality assurance guidelines.[24] The situation was reviewed in 2007.[25]

In that year, there were more than 59 million women aged 50 to 69 in the 27 member states. Currently, 22 countries have population based screening for

breast cancer either operational or under development. In 11 member states breast cancer screening takes place on a national level while 7 others have programmes under development. In 2007, roughly 21 million women were invited for screening and 12 million (41 per cent of the annual target group) participated.[25] The quoted figures relate to organised, population based screening programmes, which at least have a defined target group, screening method and screening interval. Most of the programmes also provide for the reporting of results and quality assurance. Population based implies that each new screening round involves the identification and individual invitation of eligible women.

Population based screening is a complex enterprise involving numerous different actors. Considerable commitment is required to make a screening programme successful.[26] Because, in many countries, health care is organized on a non-centralized basis and regional governments decide on screening programmes, these programmes often operate regionally. That can lead to significant inter-regional differences.[27] In Sweden, for example, women in one county are invited for screening once they reach the age of forty, while in another county participation begins at fifty. In Poland, not all parts of the country have a system where mammograms are assessed by two radiologists working independently, as recommended.[24] The annual number of screened women per mammographer varies in Europe from 670 in Flanders to 13,700 (now more than 14,000) in the Netherlands.[27] Non-centralized screening programmes generally achieve lower participation rates and are consequently less effective.[28] In Wallonia and Brussels, less than 10 per cent of the target group participates in the screening programme, while in Flanders the figure is 49 per cent.[29,30]

The Dutch screening programme is population based and is therefore distinct from, for example, the non-programmatic unorganized individualized breast cancer screening in the United States. In the United States' culture of 'defensive medicine', screening is focused on missing as few abnormalities as possible by applying a low referral threshold. Inevitably this strategy results in a high number of false positives. The Netherlands aims at an optimal balance between the detection rate and the referral rate at the population level.

Opportunistic screening

'Opportunistic' screening is aimed to be carried out when the opportunity arises, for example when a woman happens to visit her doctor. Such non-programmatic screening does not take place in the context of a coordinated programmatic system, and is also sometimes referred to as 'wild' screening. Opportunistic screening mainly takes place in countries without a population screening

programme, such as the US, or where population screening was introduced relatively recently, such as in Belgium, Germany, France, Austria, the Czech Republic and Switzerland.[31-33] The number of mammography units in use in such countries is often considerably greater than would be required for programmatic screening.[34]

Opportunistic screening is not very efficient. It does not involve a call-recall system of women in the target group. Consequently, the women who undergo screening are mainly young and well-educated, and are liable to be screened more often than necessary.[35,36] Besides being less efficient, opportunistic screening is at least twice as expensive as programmatic screening,[33] This is caused by higher charges for an opportunistic mammogram and because extra diagnostic tests (imaging, biopsy) are performed more often than necessary.[29,33,35]

With opportunistic screening, referral rates are frequently twice as high.[35] Comparative research (in women aged 50 to 69) also indicates that the test performance is often considerably inferior.[37-39] Performance levels tend to be lower for several reasons: there is no systematic quality monitoring and assurance; the radiologists are not specifically trained for the screening; they see far fewer mammograms per year and the mammograms are assessed by a single reading (whereas in population based screening programmes double reading is the standard). The inferior quality is partially offset by the higher screening frequency and higher referral rate. Opportunistic screening, as undertaken in Switzerland, could under certain circumstances be an alternative to programmatic population based screening, but is generally far more expensive and less efficient.[33]

Although the Netherlands has a population based screening programme, women might sometimes attend as outpatients for ad hoc screening. However, no recent data are available regarding the frequency of opportunistic screening in the Netherlands, or the reasons to consider it. Research into such questions would be necessary.

To sum up: around the world, many breast cancer screening activities take place. Screening is often undertaken in the context of non-centralized programmes or on an opportunistic, non-programmatic basis. There are only a few countries, such as Finland and the Netherlands, that have well-organized national population based screening programmes that meet the relevant quality requirements.[24] The Dutch population based screening programme is characterized by a high participation rate, a low referral rate and a high positive predictive value.

The European Network for Information on Cancer (EUNICE) can potentially reduce disparities between screening programmes. By collecting performance data, programmes may be regularly compared and lessons can be learned with a view to bringing about improvement.[27]

Chapter 3

The anticipated effects of population screening

3.1 Anticipated effect on breast cancer mortality

Since 1963, eight randomized controlled trials (RCTs) have been performed in the US, Canada, the United Kingdom and Sweden, with the aim of establishing whether periodic mammographic screening reduces breast cancer mortality. To support political decision-making regarding the introduction of population based screening in the Netherlands, researchers from Rotterdam, Nijmegen and Utrecht performed a cost-effectiveness analysis.[40] The researchers concluded that the results of two Swedish screening trials were the most reliable and relevant for the Dutch situation, and therefore provided the best basis for estimating the likely effect of national population based screening on breast cancer mortality. One study was conducted in Malmö and the other was the Two-County study in Kopparberg (now Dalarna) and Östergötland. From the results in women above the age of 50, the researchers estimated a mortality reduction of 33 per cent amongst women in the age group 50 to 69 who were offered screening compared to women in a control group who were not offered screening. Naturally, more pronounced reduction figures may be expected amongst women who actually participate in the screening. Consequently it was estimated that, in the Dutch population as a whole (women of all ages) by about 2015 a stable (absolute) mortality reduction of 16 per cent would be achieved. That would equate to roughly 700 fewer women dying of breast cancer in that year than would have been the case without screening.[40] More recent estimates for women in the age

group 50 to 75 have suggested an average of 775 deaths prevented (683 in 2008, rising gradually to 858 in 2018).[41] To achieve that mortality reduction, more than a million mammographic examinations are performed each year (Annex E, Table 2).

3.2 Systematic reviews

Population screening for breast cancer is one of the best documented forms of health care. Nevertheless, it has met with criticism from the outset. In 2000 and 2001, for example, two Danish researchers published the results of a meta-analysis of screening trials.[42] On the basis of their 'Cochrane review', the Danish team argued that six of the eight trials available for the meta-analysis were not of the required standard. They rejected breast cancer mortality as a valid indicator of the effectiveness of population based screening and concluded that there was no evidence that screening had any health benefit in terms of life-years gained. The Danish criticism of the screening trials has since been debated at length by the original RCT teams and by other experts.[8,43-47]

The Health Council of the Netherlands did not concur with the main criticisms levelled at population screening for breast cancer.[8] According to the Council, there were good grounds for disregarding only one of the screening trials rejected by the Danes as sub-standard (the Edinburgh trial). The Committee did agree that it was necessary to consider the harms of screening and to include death from causes other than breast cancer in the evaluation. However, the argument that breast cancer mortality was not a valid indicator of effectiveness was considered to be unfounded. The Committee found no evidence of serious distortion of breast cancer mortality in favour of screening or of over-mortality due to causes other than breast cancer. The reduction in total mortality amongst women who were offered screening was not less than one would expect from the decline in breast cancer mortality. The Committee adhered to its earlier conclusion that the screening trials demonstrated a relative breast cancer mortality reduction of 25 per cent in women above the age of 50. No scientific reason was seen to conclude from the Cochrane review that population screening for breast cancer, as practised in the Netherlands, had no value.[8,48]

This report does not re-examine the details of the randomized screening trials; it considers only the results of meta-analyses of those trials. The Committee sees no reason to doubt the scientific validity of the screening trials. Hence, the following sections of this report focus mainly on the subsequently published observational studies of population screening practice.

Various researchers have combined and analysed the results of the available screening trials. Their meta-analyses show a (relative) reduction in breast cancer mortality of roughly 20 per cent in women above the age of 40 to whom screening was made available.[5,49,50] In the Netherlands, screening is offered only to women above the age of 50. The trial findings indicate that, in that age group, the reduction in breast cancer mortality is 28 per cent after a seven-year follow-up and 23 per cent after thirteen years.[5]

A meta-analysis is a statistical study combining the results of several published (clinical) trials. A more accurate method of combining data is a so-called 'IPD meta-analysis'. An IPD meta-analysis is based on individual patient data from the trials. A committee of the Swedish Cancer Society used the IPD meta-analysis technique to uniformly collate and analyse data from the four Swedish trials (Malmö, Kopparberg (now Dalarna)/Östergötland, Stockholm, Gothenburg).[51] Published in 1993, the resulting 'Swedish overview' concluded that population screening was associated with a relative reduction in breast cancer mortality of 24 per cent in women aged between 40 and 75, and 29 per cent in women aged between 50 and 70.[51]

A second overview of the Swedish randomized trials, published in 2002, analysed data from an observation period of nearly sixteen years, i.e. seven years longer than the previous overview.[52] The conclusion was that, in women over the age of 40, population screening was associated with a breast cancer mortality reduction (relative risk reduction) of 21 per cent: RR=0.79 (95 per cent confidence interval 0.70-0.98). The relative risk reduction was 16 per cent for women between the ages of 50 and 60 and 33 per cent for women aged between 60 and 70.[52]

3.3 Conclusion

In 1990, it was expected that the introduction of population based screening to the Netherlands would reduce breast cancer mortality in women between 50 and 70 by 33 per cent. In due course, that would mean roughly 700 fewer women dying of breast cancer per year. That expectation was based on the results of two population screening trials in Sweden. Subsequent meta-analyses of all the relevant screening trial data then available have indicated that population screening was associated with a relative mortality reduction of approximately 25 per cent in women aged between 50 and 70.

Randomized trials have the highest level of evidence for the expected effect of screening. However, they were carried out mainly in the eighties; their results might therefore not be entirely applicable to the current situation. Since the trials,

the background situation has changed considerably, for example because of the introduction of more effective forms of breast cancer treatment. In the following section, the Committee considers the changes that may be relevant to the assessment of the efficacyof population based screening.

Chapter 4

Changed circumstances

4.1 Trends in incidence and risk factors

Since 1960, research in 28 western countries found that the incidence of breast cancer (after correction for aging and population growth) has risen everywhere.[53] Icelandic data covering an even longer period show that the incidence in the period 1985 to 2002 was four times as high as that in the period 1921 to 1944.[54] In non-western countries too, the incidence of breast cancer is rising.

It is not entirely clear why more women are developing breast cancer. Fewer than one case in ten is primarily hereditary. There has been public and political disquiet about a number of unproven risk factors, such as the use of deodorant and having had an abortion.[55]

Reproductive factors are known to play a significant role. Childlessness is increasingly common: one in five women born between 1960 and 1964 has no children.[56] Those who do have children have fewer than in previous generations, are older when their first child is born, are less likely to breastfeed their babies, and continue breastfeeding for a shorter time. Women nowadays enter puberty younger and reach the menopause later in life.[57,58] It is increasingly common for girls to start their periods while still at primary school; five years earlier than eighty years ago.[59] Taken together, those changes mean that, during the fertile phase of her life, the average woman is estimated to go through three times as many ovulation cycles.[60] That in turn increases the breast cancer risk, which is known to be associated with the cyclical growth of epithelial cells in the

mammary glands during the menstrual cycle. The more often cells divide, the greater the chance of mutation.

Environmental factors may also be driving the rising incidence of breast cancer. A century ago, there were no hormone-disrupters (certain pesticides, fire retardants, plasticisers, antibiotics in livestock, oral contraceptives, or hormone suppletion therapeutics). Such substances are known to be present in drinking water and food in increasing quantities. However, it is unclear whether and, if so, how often exposure to such substances during foetal development, during adolescence and prior to the birth of a woman's first child can lead to breast cancer later in life.

There have also been changes in lifestyle. Lack of exercise and obesity are increasingly common, and both are associated with a higher incidence of breast cancer.[57] For example, the American Nurses' Health Study found that a bodyweight increase of more than twenty kilos before the menopause was linked to a doubling of breast cancer risk.[61] Other risk factors are smoking, alcohol consumption and working nights for a prolonged period (i.e. for more than thirty years).[62-65]

The apparent changes in incidence can differ between countries. In a number of western countries, the incidence of breast cancer, particularly in women aged between 50 and 60, suddenly began to fall in 2002, after previously rising for several decades. The fall coincided strikingly with the rapid decline in the use of hormone suppletion therapy by menopausal and postmenopausal women.[66] The use of hormone suppletion therapy – lauded in the sixties as a panacea for conditions of old age (*feminine for ever*) – was found to increase the risk of breast cancer (and cardiovascular disease).[67]

On the individual level too, it was demonstrated that the elevated risk of breast cancer fell considerably after stopping hormone suppletion therapy. The reduced incidence could not be attributed to less mammographic screening.[68] In the Netherlands, where hormone suppletion therapy was considerably less common in the nineties than, for example, in Belgium, Denmark or France, no such incidence reduction was observed.[69]

Data from the period 1975 to 1989 show that the risk of developing breast cancer was rising in the Netherlands, as it was elsewhere.[70] The higher incidence recorded after 1989 may be explained partly by the introduction of screening, since screening leads to breast cancer being diagnosed earlier than would otherwise have been the case. The age-standardized incidence (European standardized rate, including DCIS) rose from 102 to 129 per 100,000 women between 1989 and 1994. After that, there was a slight drop in the incidence of breast cancer. When in 1998 women between the ages of 70 and 75 were added

to the target group, the incidence rose again, from 136 to 154 per 100,000 women between 1999 and 2011.

At 154 per 100,000 women, the incidence (ESR) is now 50 per cent higher than in 1989 (102 per 100,000). Whereas in 1989 the chance of a woman being diagnosed with breast cancer at some time in her life was one in ten, by 2003 it had increased to one in seven.[71]

When the effect of screening is excluded, there remains an underlying upward trend in the incidence of breast cancer (http://cijfersoverkanker.nl/trends-53.html). In women aged between 35 and 50 and (until 1999) in women aged between 75 and 85, the incidence has continued to rise since 1989.[72] The incidence of breast cancer in women aged 44 is today similar to that in women of 50 when population screening first began.[73]

4.2 Trends in treatment

In nine cases out of ten, the primary treatment for breast cancer is surgery. Only in women above the age of 70 does non-surgical treatment, in the form of hormone therapy, play a significant role (roughly 20 per cent of all primary therapies in that age group). When the screening trials were carried out, complete removal of the breast (mastectomy, breast amputation) was the standard treatment. In the early eighties, however, a global shift began towards more limited surgery (breast-conserving treatment).[74] As women became more alert to the warning signs and following the introduction of population screening, it became increasingly common for breast cancer to be detected at an early stage. The advantage of breast-conserving surgery over amputation is retention of the breast, resulting in a cosmetically more acceptable outcome and improved quality of life.

In the Netherlands, the percentage of women with invasive breast cancer who received breast-conserving surgery rose from 29 per cent in 1990 to 50 per cent in 2004, while in the same period the breast amputation rate fell from 59 to 38 per cent.[75] In 2010, more than 60 per cent of patients in the south-eastern Netherlands received breast-conserving surgery.[22] On the basis of the Dutch quality criteria, three quarters of patients with breast cancer may be treated using breast-conserving surgery in combination with radiotherapy.

Early detection makes breast-conserving surgery possible in more cases. However, that benefit has been partly offset by more restrictive specification of the indications for such surgery. More experience with this procedure highlighted the need to take greater account of the factors influencing the risk of cancer returning after breast-conserving surgery.[76] Consequently the percentage of

patients with local recurrence of cancer decreased considerably. Breast-conserving surgery is almost always followed by radiotherapy.[22,77] Following traditional surgery, radiotherapy is indicated only if there is a high risk of the cancer returning (www.oncoline.nl/mammacarcinoom).

The introduction of adjuvant pharmaceutical therapy (i.e. supplementary treatment following surgery) has significantly improved survival in breast cancer patients.[78] Whereas a third of new patients received adjuvant therapy in 1990, more than half did so in 2004. The use of adjuvant chemotherapy in premenopausal patients with positive axillary lymph nodes rose from 15 per cent in the period 1975 to 1979, to 30 per cent in the period 1980 to 1984.[75] Since the eighties, adjuvant hormone therapy has been given to postmenopausal women with oestrogen receptor-positive breast cancer and positive lymph nodes.

The pharmaceutical treatment of breast cancer is a highly dynamic field. The ability to distinguish molecular subtypes of breast cancer has facilitated the development of targeted therapies.[79] More than ten years ago, with the launch of trastuzumab, a new category of monoclonal antibodies became available, targeted at HER2+ breast cancer. The latter subtype includes 20 to 30 per cent of all invasive breast cancers and is associated with a poor prognosis. A great deal of research is being done to establish the optimal duration of treatment. However, the antibodies are often ineffective. Fortunately, molecular analysis of metastases or molecular imaging makes it possible to identify the patients in whom no effect is likely. Meanwhile new groups of drugs are appearing, such as ADCs (antibody-drug conjugates, i.e. antibodies attached to chemotherapeutics) and CDK (cyclin-dependent kinase) inhibitors. Cyclin-dependent kinases are involved in cell cycle regulation. If a CDK is blocked, DNA replication is not possible. These various developments can contribute to the efficient use of therapy.

Neoadjuvant chemotherapy (chemotherapy given prior to surgery) is an increasingly important element of therapy, both for locally advanced breast cancer and for earlier stage breast cancer (www.oncoline.nl/mammacarcinoom). One reason for giving chemotherapy *before* surgery is to reduce the size of the tumour, thus enabling breast-conservation. In 2010, neoadjuvant chemotherapy was given in nearly 30 per cent of cases where patients with invasive breast cancer received chemotherapy.[22] Following neoadjuvant chemotherapy, surgery and possibly radiotherapy, the patient receives no further adjuvant chemotherapy.

Survival amongst breast cancer patients has improved greatly in recent decades. Over the last twenty years, the five-year survival rate has increased from 77 to 86 per cent (www.ikn.nl). Annual breast cancer mortality (ESR) fell

by 34 per cent after 1986/1988 (partly as a result of screening), to 62 per 100,000 women aged 50 to 74 in 2012.

4.3 Conclusion

Much has changed since the screening trials were undertaken. The treatment of breast cancer has improved considerably. Regardless of the effect of population screening, the risk of breast cancer has been increasing for decades in the Netherlands, as it has in other western countries. Changes in known risk factors for breast cancer, such as women having their first child later and postmenopausal obesity, can explain only part of the increase in breast cancer incidence. On the other hand, survival amongst breast cancer patients has improved greatly in recent decades. Annual breast cancer mortality (ESR) fell by 34 per cent after 1986/1988, to 62 per 100,000 women aged 50 to 74 in 2012.

Chapter 5

Effectiveness of population screening

5.1 Divergent conclusions

By means of experimental research (randomized screening trials), it is possible to determine the efficacy of population screening: what might the effects be under controlled conditions? Once population screening has been introduced to a country or region, there is no longer a control group available for use in experimental research on the basis of randomization. Then one needs other, observational research methods to establish whether the aims of the programme have been met and continue to be realized: what is the effectiveness of the programme?

The academic literature on the effectiveness of population screening for breast cancer presents conflicting outcomes. While some studies indicate that population screening is responsible for a substantial reduction in breast cancer mortality, others suggest that the effect is minimal. It is therefore pertinent to consider whether these discrepancies can be explained.

One important consideration is, that the effect of screening is often measured by reference to the mortality statistics derived from cause-of-death records. The effect of screening cannot be determined directly from such data, because, for a long time, mortality statistics continue to predominantly reflect breast cancer cases diagnosed prior to the introduction of population screening. The presence of cases involving women who had no opportunity to undergo screening, will weaken any effect of screening. This form of bias can be avoided by using

incidence-based mortality (IBM) as effect parameter, where inclusion in the analysis is restricted to mortality from cases with a diagnosis of cancer after invitation for screening. Identical to a screening trial, in an observational study breast cancer cases that were known to exist before the patient had an opportunity to undergo screening, need to be excluded.

Another, related, explanation for the diverging outcomes, is variation in the length of the observation period. This may be illustrated with the results of the Swedish Two-County Trial, begun in 1977.[80,81] All women in whom breast cancer was detected during the screening period (seven years in the randomized trial) were followed up for three decades. After a follow-up of up to 29 years, breast cancer mortality in women aged 40 to 70 to whom screening had been made available was 30 per cent lower than in the control group, to whom no screening had been offered.[81] That finding replicated the finding made in 1985, when the first evaluation was carried out.[82] Over time, the stable relative risk reduction was translated into a rising absolute number of breast cancer mortality cases prevented by screening. That pattern is to be expected, because variations in the speed of breast cancer progression mean that screening can prevent deaths that, without screening, would not have occurred until more than ten years later. By the end of the follow-up period, the absolute number of prevented mortality cases was more than twice the number after a follow-up of ten years. In other words, most of the health benefit becomes manifest more than ten years *after* the introduction of population screening. Consequently, the effect of screening can only be accurately determined over the long term.

Another cause of heterogeneity in the results is, that less effect is observable when population screening is compared with a reference situation characterized by a more or less marked level of opportunistic screening rather than with no screening at all.

Furthermore, there are substantial differences in quality between screening programmes, and a variety of observational methods have been used by researchers, including trend analyses, cohort studies and patient-control studies. In view of the continuing debate regarding the effectiveness of population based screening for breast cancer, the Committee considers the methodology and results from observational studies in more detail below.

5.2 Trend studies

A first step in the evaluation can be studying the changes in breast cancer mortality after the introduction of population screening, or comparing breast cancer mortality between regions with and without population screening. This

methodology is relatively straightforward and therefore widely used. However, mostly interpretation of the results is next to impossible.

Trend research often depends on published demographic data, data from cause-of-death records, from which the effect of screening cannot be directly determined, as was explained in subsection 5.1. Moreover, when comparing data between different countries, the findings depend greatly on the choice of countries. Circumstances that can influence mortality of breast cancer, unrelated to population screening, can considerably bias the results. Comparison can only be meaningful if adequate correction is made for differences, in breast cancer risk factors, in therapy for breast cancer, in hormone suppletion therapy for menopausal or postmenopausal women and in the use of opportunistic screening.

Results of trend studies

A study of breast cancer mortality between 1989 and 2006 in thirty European countries revealed major trend differences, ranging from a 45 per cent decrease in Iceland to a 17 per cent increase in Romania.[83] Significant differences were also found between countries with population screening programmes: in the United Kingdom the decrease was more than twice as high as in Finland or Sweden. Breaking down the figures by age, the decrease was most pronounced in women under the age of fifty, even in countries where women of that age group were not targeted for screening. The researchers had expected breast cancer mortality to decrease most in the countries with the greatest screening capacity, such as France and Sweden. They concluded that more detailed data than national mortality statistics were required in order to establish a link between population screening and reduced breast cancer mortality.

Another study compared trends in three countries that were considered to be comparable, including Belgium and the Netherlands.[84] As in the Netherlands, breast cancer mortality across all age groups fell in Flanders by 25 per cent between 1989 and 2006, even though population screening was not introduced to Flanders until 2001. The researchers' conclusion was that population screening had very little influence on breast cancer mortality.

However, the supposed comparability of the selected countries is not well documented. For example, there are major differences in terms of the level of opportunistic screening and the use of hormone suppletion therapy.[69] Already well before the introduction of population screening, 33 per cent of Flemish women already underwent 'diagnostic' mammography each year, nearly always accompanied by ultrasonography or other forms of diagnostic testing as well.[29] The comparison was somewhat superficial for still other reasons as well. For the

study period (1989-2008), age-specific mortality statistics were available for Flanders only from 2005,[6] while on a national level breast cancer mortality in women aged 50 to 69 hardly decreased at all.[85]

A United States' trend analysis led to the conclusion that screening often leads to overdiagnosis, while the effect on breast cancer mortality is marginal at best.[86] However, this study did not utilize mortality data and concerned opportunistic screening, not population based screening. Research comparing population screening in Norway and opportunistic screening in the US state of North Carolina (in women aged 50 to 69) found that the test performance was considerably better in the Norwegian population based screening.[38]

In Denmark, comparative research was possible because, for seventeen years, only a fifth of the population was offered population based screening: the screening began in Copenhagen in 1991 and in Funen two years later. Breast cancer mortality was studied between 1971 and 2006, using the rest of Denmark as a non-screened control group.[87] The conclusion was that in women aged 55 to 74 breast cancer mortality decreased by 1 per cent per year in Copenhagen and Funen in the period 1997 to 2006, compared with 2 per cent in the rest of Denmark. The researchers concluded that the Danish screening programme had no effect on breast cancer mortality.[87]

However, the researchers' conclusion is not supported by the research. By restricting themselves to the period 1997 to 2006, the researchers failed to consider the bigger picture. In the period 1982 to 1991, before population screening began, breast cancer mortality was higher in the regions with population screening than in the rest of Denmark (121 per 100,000 women aged 55 to 74 compared with 109 per 100,000). Yet, in the period 1997 to 2006, the figure was slightly lower in regions with population screening compared with the rest of Denmark (102 per 100,000 versus 106 per 100,000). In other words, on the contrary, mortality decreased considerably more in the regions with population screening than in the rest of Denmark.

In 1988, the UK began introducing population screening based on a three-year cycle for women aged between 50 and 65. On the basis of data from the years 1971 to 1989, the levels of breast cancer mortality to be expected in the period 1990 to 1998 without and with screening were compared.[88] The estimated relative mortality reduction attributable to screening in 1998 was 6 per cent in women who were expected to have the greatest health benefit, i.e. women in the 55 to 69 age group. The researchers suggested that the calculated effect was relatively small because the evaluation followed quite soon after the introduction of population screening. A quite reasonable assumption, because the programme was fully operational only from 1995. Furthermore, the analysis method used did

not take account of the 'diluting effect' of breast cancer mortality in women who were diagnosed before they had the opportunity to undergo screening. A later study, with a follow-up period of at least ten years from the year of full implementation, found that breast cancer mortality in women between 50 and 70 was 28 per cent lower than in the period before 1989 taking trends in other age groups into account.[89]

Similar findings were made in Italy and Spain. In Barcelona, only a modest mortality reduction was observed in the year that introduction of the screening programme was completed,[90] but a considerably larger effect was apparent in Navarra after a longer follow-up.[91]

In a trend study in the Netherlands, researchers could account for the month that population screening was introduced in each individual municipality.[92] As the gradual introduction of screening generally takes years, in the Netherlands the effect on breast cancer mortality could be measured earlier than without this information, because the starting times could be aligned. In women aged 55 to 75, the mortality rate was rising 0.3 percentage points per year before the introduction, but fell by 1.7 percentage points per year following introduction. The turning point was consistently close to the year that population screening began in the given municipality. The decline in mortality became statistically significant after roughly five years.

A later trend analysis in the Netherlands involved a longer period of screening and follow-up.[72] It was found that, in women aged 55 to 75, after a period in which breast cancer mortality rose, a decline of about 2.5 percentage points a year was observable in the period 1994 to 2006. In the age group 75 to 84, mortality began to fall in 2001, four years after the target group for screening was enlarged to include women aged 70 to 75. In women aged between 45 and 55, mortality rose in the period 1950 to 1971, then began falling in the seventies; that trend continued after 1992. There was a clear correlation between the observed age-specific trends and the introduction and expansion of population screening.[72]

Observation of an effect starting just a few years after women were first invited for screening is consistent with the results of screening trials. In the Swedish trials, an effect on breast cancer mortality became apparent four years after randomization.[52] Most trend studies start follow-up at the first introduction of screening, even though introduction may take several years to complete. Then the time lag between the start of the study and the first observable effects would take longer, as was the case in the British, Italian and Spanish studies referred to earlier.[88-91]

The Committee concluded that trend studies do not generally provide a sound basis for conclusions.[50] The few trend studies that satisfy the minimum requirements suggest that screening has a favourable effect on breast cancer mortality at least as great as that predicted by the original screening trials.

5.3 Analysis of trends in the incidence of advanced cancer

An alternative to analysing trends in breast cancer mortality is to analyse trends in the incidence of advanced breast cancer. One advantage of using this 'surrogate' outcome measure is that, while advanced stages of cancer and mortality are closely related, a fall in the incidence of advanced stages of disease is likely to be observable sooner than a fall in mortality. Another advantage is that (unlike mortality) the stage of the disease at the time of diagnosis is not influenced by therapy.

Nevertheless, like the trend in mortality, the trend in advanced stages is subject to 'underlying' changes in incidence: if the incidence is rising for reasons unrelated to screening, a similar rise in the incidence of advanced stages can be expected. One could account for this form of bias in the effect of screening by adjusting the effect on the basis of information about underlying trends.

Ideally, the introduction of a screening programme should lead to a rise in the incidence of early-stage breast cancer, followed by a fall in the incidence of advanced stages of breast cancer.[93] A fall in advanced stages should precede a fall in breast cancer mortality and serves as an early indication that screening is effective.

Research results

In the Netherlands, population screening for women aged 50 to 70 was introduced between 1989 and 1997. In that period, the incidence of advanced breast cancer (defined as invasive tumours with a diameter of more than 20 millimetres, combined with positive lymph nodes and/or remote metastases (T2+N+/M1)) fell by 12 per cent, from 72 to 63 per 100,000 women per year (European Standardized Rate).[94] In 2003, the fall relative to 1989 was 13 per cent in women aged 50 to 70 and 11 per cent in women above the age of 70.[75]

The figures cited in the previous paragraph take no account of the upward trend in background risk. That trend is illustrated by the fact that, in women under the age of 50, the incidence of advanced breast cancer gradually rose by 22 per cent in the Netherlands between 1989 and 2003.[75,95]

In 1996, biennial screening for women aged between 50 and 70 was introduced to part of Norway. In women who were called up for screening in the period 1996 to 2004, the incidence of advanced breast cancer was 15 per cent lower than in the preceding years: RR (relative risk) =0.85 (95% CI 0.84-0.87).[96] A follow-up study demonstrated that the incidence of advanced (i.e. stage III or IV) breast cancer in screening participants was considerably lower than in non-participants (16.4 per 100,000 versus 45.1 per 100,000).[97]

In thirteen Swedish regions, the incidence of advanced breast cancer in women aged between 50 and 70 to whom screening had been made available was on average 33 per cent lower than the incidence prior to the introduction of population screening.[98] In seventeen regions in central and northern Italy where population screening was organized, the incidence of advanced breast cancer fell by 19 per cent between 1997 and 2001.[99] A later Italian study found a 20 to 30 per cent lower incidence of advanced breast cancer after the second round of screening.[100]

The Committee concludes that research into trends in the incidence of advanced breast cancer generally indicates an effect that is less pronounced than the effect on breast cancer mortality observed in the original screening trials.

5.4 Cohort studies

In a cohort study, researchers generally compare the incidence of a disease in a population that *is* exposed to a factor that may be linked to the development of the disease, with the incidence in a population that is *not* exposed to that factor. In this case, comparison is made between breast cancer mortality in a population that *was* invited for screening (the study group) and mortality in a population that was *not or not yet* invited (the control group).

In other words, the aim is to compare the effect of screening with non-screening. However, the research cohorts are defined based on women being invited for screening or not, similar to the screening trials. Because the selection in the cohorts does not depend on actual *participation* in the screening programme, self-selection does not bias the results.

A cohort study has a deeper research level than a trend study and does not rely merely on mortality statistics from demographic data. Obviously screening can only affect mortality in breast cancer cases diagnosed after invitation for screening, i.e. incidence-based mortality (IBM). The use of IBM avoids a negative bias of the screening effect by the inclusion of mortality data relating to patients who had no opportunity to undergo screening before being diagnosed with breast cancer.

In a cohort study, more than in a trend study, researchers aim to correct for factors other than population screening, which might influence breast cancer mortality, such as opportunistic screening, improved therapy and changes in lifestyle. The control cohort has to be selected very carefully. The best correction is possible by either selecting a control group comprising women from the same time period and region, who were not invited for screening, or by using historical data from the index region plus historical and contemporary data from a region without population screening.

Cohort study results

A Finnish cohort study found that breast cancer mortality was 28 per cent lower in women aged 50 to 69 who had been called up for screening: RR=0.72 (95% CI 0.51-0.97).[101] However, with only a historical control group, the effects of screening and the effects of other temporal changes, such as improved breast cancer therapy could not be adequately distinguished in this study. Consequently, screening might be less effective than the data might suggest.

A Norwegian study, concluding that the effect of a screening was only 10 per cent, caused considerable debate.[102] The study made use of incidence-based mortality and control groups, but took no account of the fact that, in Norway, as in Belgium and elsewhere, population screening was introduced fairly late, by which time opportunistic screening was already practised on a fairly large scale.[103] Analysis of the five available data sources on opportunistic screening subsequently indicated that the effect of population screening would have been at least twice as great if it had been introduced to a situation where there was no screening at all.[103]

To estimate what the level of breast cancer mortality would have been without screening, a (different) Finnish study and an Italian study used a control group of women who, during the study period, were not yet invited for screening.[104,105] When Finland introduced population based screening in 1987, cohorts were preselected on the basis of birth year and invited for screening before other birth-year cohorts, enabling the subsequent comparison of IBM statistics.[105] All women were followed up at the individual level. The (relative) reduction in breast cancer mortality associated with the availability of screening was calculated to be 24 per cent. This effect of screening was slightly too small to be statistically significant, because the period without screening for the controls was too short. Already four years after the first cohorts, the control cohorts were invited for screening, which is why only the early effect of

screening could be analysed. After a follow-up of three to four years, the mortality reduction rose to 31 per cent: RR=0.69 (95% CI 0.35-0.99).[105]

Because population screening was introduced to Florence gradually (from 1990), the Italian city served as a 'natural experiment', in the context of which breast cancer mortality in women invited for screening could be compared with uninvited women from the same municipality in the same period.[104] Here, the (relative) reduction in breast cancer mortality was found to be 19 per cent.

In other cohort studies, the expected breast cancer mortality without screening was calculated from historical data combined with contemporary data from regional control groups. In a Danish cohort study, individualized data on all women were used to determine the effect of the first five screening rounds in Copenhagen (1991-2001).[106] After correction for temporal changes and regional differences in breast cancer mortality, the relative reduction in breast cancer mortality attributable to the availability of screening was found to be 25 per cent. In women who actually participated in the screening, after correction for self-selection bias, breast cancer mortality was 37 per cent lower than would otherwise have been expected.[106] Like after various other studies, the Danish study gave rise to debate as to whether it was possible for screening to have an observable effect within as little as five years.[87,107,108]

Cohort studies in which expected breast cancer mortality was estimated on the basis of historical control groups (without opportunistic screening) combined with contemporary data on non-participants observed reductions between 18 and 48 per cent.[91,109-112]

Still, taking a critical view, the effectiveness of screening may be disputed. However, in further support of the effectiveness of screening a meta-analysis was performed of methodologically the most reliable cohort studies in Europe: involving adequate control groups enabling correction for differences in background risk factors, with an intake period matching the follow-up period. This meta-analysis concluded that breast cancer mortality was 26 per cent lower in women between 50 and 70 who were invited for screening: RR=0.74 (95% CI 0.64-0.87).[113]

5.5 Case-control studies

While a cohort study estimates the breast cancer mortality reduction amongst women invited for screening, a case-control study considers the reduction amongst screening participants. In a case-control study (or case referent study), the odds of screened and unscreened individuals in a group of women who died from breast cancer (cases) is compared with the odds of screened and unscreened

individuals in a control group, comprising women who, like the case group, were invited for screening. The ratio between the two (the odds ratio, OR), i.e. the factor by which the odds differ, forms the outcome parameter. If there is no correlation between screening and breast cancer mortality, the OR should be 1. If there is a positive correction (an adverse effect), the OR will be >1, and if there is a negative correction (a favourable effect), the OR will be <1. In most case-control studies, the odds ratio may be interpreted as relative risk (or 'risk ratio').[114]

Like a cohort study, a case-control study avoids the negative bias of the screening effect that occurs in trend studies; cases of breast cancer mortality without opportunity to undergo screening prior to diagnosis are excluded. In a case-control study only women who were diagnosed during the period when screening was available and who were in the relevant age group at the time, are selected. Cases and controls must have had equal opportunity to undergo screening and must be drawn from the same source population. By linking individual screening participation data to individual breast cancer mortality data, the mortality reduction directly attributable to screening can be quantified.

Case-control studies do not involve randomization and therefore the level of evidence is lower than experimental research; because participants and non-participants in screening may differ in terms of risk factors for breast cancer mortality. Consequently, this confounding bias may lead to inaccurate estimates of the screening effect.

Confounding bias can largely be avoided by a study design and analytical methods that account for age – the primary risk factor for breast cancer mortality. In other words, the age distributions of the case and control groups need to correspond well. Residual confounding by other risk factors and by self-selection may nevertheless remain.

Residual confounding is a major focus of debate regarding the effectiveness of population based screening for breast cancer. It has been argued that favourable outcomes observed in case-control studies are not attributable to screening, but primarily to the fact that screening participants were at lower risk for unrelated reasons. Participants and non-participants are claimed to differ in terms of background breast cancer mortality risk to such an extent that correction is impossible. Therefore, case-control studies would not be adequate to evaluate screening.[115]

Is it correct to assume that the background risk for non-participants is considerably higher than that for participants? Breast cancer is one of the few forms of cancer that is more common amongst wealthy, well-educated women.[116,117] Women of lower socioeconomic status (SES) are less inclined to

participate in population screening and have poorer survival prospects. However, in the Netherlands the differences are not high. The participation rate is not much lower in low-SES women (79 per cent) than in high-SES women (87 per cent).[118] This difference in participation is relevant, but considerably smaller than found with opportunistic screening.[36] In one study, all women diagnosed with breast cancer in the Netherlands between 1995 and 2005 were divided into five equally large SES groups.[119] The (breast cancer-specific) ten-year survival rate in the lowest SES group (74 per cent) was not much lower than the rate in the highest SES group (79 per cent).

Non-participants do not necessarily have a higher background risk. They may also have a lower background risk.[120,121] In the Netherlands, it is often women from ethnic minority backgrounds who are more reluctant to participate. However, breast cancer mortality in the ethnic minorities is lower than in the ethnic majority.[122-125] This is also true for women of low SES within the ethnical-majority.[126] Calculated on the basis of incidence-based mortality, in Limburg non-participants were found to have a lower background risk than participants.[127] The difference in background risk can vary by region and by period.

Analysis of realistic scenarios for the frequency of risk factors and the relative risk of those factors in screened and non-screened groups indicates that in the Netherlands residual confounding will not substantially bias the findings of case-control studies.[128] Correction factors were calculated for five screening regions in the Netherlands.[129] For three regions, correction for self-selection did not affect the odds ratio, while for the two other regions correction *increased* the screening effect. A similar study in Italy also found that residual confounding could not substantially bias the results of case-control studies.[130]

Has self-selection not increased over time? That is unclear. Data about the characteristics of non-participants are scarce and information regarding changes in those characteristics over time is completely unavailable. It is known that the level of compliance with participation is high. Over 90 per cent of the women who participated after the previous invitation, participated again after the next invitation.[11,75] That pattern has persisted for twenty years.

Results of case-control studies

In the last decade, the results of more than ten case-control studies have been published. A large US study detected very little screening effect.[131] However, the screening consisted mainly of clinical breast examination (palpation), as was usual in the US in the eighties, whose effectiveness has never been demonstrated. Furthermore, information regarding the quality of the mammographic screening

was unavailable. As the researchers themselves concluded, their findings are not valid outside their particular setting.

Case-control studies often do not correct for selection bias. Studies that did and were carried out in countries with population screening programmes have consistently found that mammographic screening has a favourable effect.[121,127,130,132-136]

Nevertheless, the observed relative mortality reductions vary considerably between 38 and 70 per cent[127,136]. Six patient-control studies were analysed to establish if the study methods that were used, could explain these outcome differences.[137] Although several minor design differences were found in the study designs, the methods were broadly similar. The differences in outcome were attributed mainly to other factors, such as the screening interval and factors concerning the implementation. In the Netherlands, after correction for self-selection bias, studies indicate a relative reduction in breast cancer mortality of at least 50 per cent for screening *participants*.[127,133,134] That implies, that by participating in screening, for example, a woman of fifty reduces the risk of dying from breast cancer before the age of eighty from 2.6 per cent[138] to 1.3 per cent.

The Committee concludes that the results of well-designed case-control studies consistently indicate a significant reduction in breast cancer mortality due to screening.

5.6 Effect of screening on general mortality

If screening reduces breast cancer-specific mortality, the next question that arises is whether general mortality (mortality from all causes) is proportionately reduced. If, on the basis of the trials, it is assumed that the provision of screening reduces breast cancer mortality by 20 per cent (or 25 per cent in women above the age of fifty), it follows that general mortality should decline by about 1 per cent. Statistics indicate that this is indeed the case. The Swedish overview found that, after nearly sixteen years' follow-up, general mortality was 2 per cent lower in women offered screening: RR=0.98 (95% CI 0.96-1.00).[52]

Mammographic screening should only have a favourable effect on mortality in women in whom breast cancer is diagnosed as a result of screening. The mortality pattern in such women therefore provides a more accurate picture than the mortality pattern in all women to whom screening is made available.[47] When restricted to women diagnosed with cancer, analyses of the Two-County Trial results revealed that, after a twenty-year follow-up, the general mortality in

women called up for screening was 13 per cent lower than in women with breast cancer in the control group.[139]

5.7 Effect of screening and improved therapy

Although it is not a universal trend, breast cancer mortality decreases in many countries, including countries without population based screening. In addition to the introduction of population based screening, there have been various significant developments, including the increase of opportunistic screening and the introduction of adjuvant therapy.[34] Adjuvant therapy has made the treatment of breast cancer more effective. Furthermore, women are quicker to act on signs that might point to breast cancer. Does the new landscape, in which breast cancer is typically diagnosed sooner and remains treatable even at relatively advanced stages, mean that screening is now less effective and less necessary? Or are the improved treatment outcomes partly due to early detection?

The cohort studies and case-control studies referred to above were performed at a time when current treatment methods were already in use. Yet their findings show a screening effect at least as great as suggested by the original screening trials. It is the combination of early detection and adequate treatment at an early stage that has been the key to reducing mortality.[140] Yet it tends to be assumed that the effects of population screening and improved treatment are mutually independent.[50]

A computer model has been used to calculate the respective contributions of population screening and adjuvant therapy to the reduction in breast cancer mortality in the Netherlands between 1975 and 2008.[10] It was estimated that the application of adjuvant therapy brought down breast cancer mortality by 14 per cent in 2008, while the current population screening programme was responsible for a further 16 per cent reduction.[10] The latter percentage equates to 683 breast cancer deaths prevented by screening in 2008.

In the decision making about the introduction of population screening in the Netherlands, modelling indicated that a stable relative mortality reduction of 16 per cent (in women of all ages) would be achieved by about 2015, and that then roughly 700 fewer women would die from breast cancer in that year than would otherwise have been the case.[40] The most recent estimates – 683 fewer cases of breast cancer mortality due to screening in 2008 and 858 fewer in 2018 – are in line with the original forecast.[10]

Modelling in other countries has also indicated that adjuvant therapy and population screening have contributed about equally to the observed fall in breast cancer mortality.[141] By inputting data from the same sources into seven different

mathematical models, researchers estimated that, in the United States, nearly half of the reduction in breast cancer mortality between 1975 and 2000 was attributable to screening.[142]

Studies of population screening in Florence and Turin between 1990 and 2001 pointed to a reduction in breast cancer mortality of 27 per cent in women invited for screening, relative to women not yet invited.[143] Comparison on the basis of individual data on tumour characteristics (stage, malignancy grade) and detection method indicated that the better prognosis for women who had been invited for screening, relative to the prognosis for other women, was attributable to the earlier cancer stages at the time of diagnosis, and not to improved treatment.[143]

It is remarkable that screening programmes have proved no less effective than the trials indicated. Naturally, national population based screening is logistically quite different from a trial. On the other hand, screening mammography has improved considerably; it is now the norm to take two mammograms per breast and to assess the mammograms by two radiologists working independently.[144]

In addition, the organization and implementation of the screening programme has become increasingly professional over time, with the introduction of quality and training requirements. Meanwhile the nationally active reference centre, monitoring and impact evaluation have pushed the optimization of population based screening.[18,145] Finally, improvements have been made in diagnosis following a positive screening result and to the connections between the screening programme and breast clinics.

A case-control study in Nijmegen, analysed by calendar year, indicates that the effectiveness of population screening has increased over time.[134]

5.8 Conclusion

The Committee concludes that observational research into the effectiveness of breast cancer screening has yielded conflicting results, ranging from a substantial reduction in breast cancer mortality due to screening to a minimal screening effect. The divergent findings may be explained primarily by the research methodologies used, such as gauging the effect of screening from mortality statistics derived from cause-of-death records, which for years after the introduction of screening are dominated by cases of breast cancer detected before opportunity for screening was given, thus negatively biasing the screening effect.

Because of this bias, trend studies lack in discriminatory power. The inadequacy of the outcome parameter and the usually limited level of analyses

mean that trend studies cannot sufficiently support substantial conclusions. On the other hand, a well-designed cohort study or case-control study can show the effect of screening. In such studies, the effect of screening can be distinguished from the effects of other factors influencing breast cancer mortality, such as improved therapy and trends in the background risk of breast cancer.

A meta-analysis of methodologically the most reliable cohort studies in Europe found that breast cancer mortality was 26 per cent lower in women aged 50 to 70 who were offered screening (and who did not all participate). Case-control studies provide information on the correlation between participation in screening and breast cancer mortality. The results of such studies consistently indicate a substantial reduction in breast cancer mortality due to participation in screening. According to case-control studies in the Netherlands, Dutch participants can expect to halve their risk. Based on computer modelling, it has been estimated that, between 1975 and 2008, screening reduced breast cancer mortality in the Netherlands by 16 per cent.[10] That equates to 683 fewer breast cancer mortality cases due to screening in 2008.

The Committee concludes that, in the Netherlands, population screening for breast cancer continues to be as effective as was originally expected, even though the landscape has changed. In light of that conclusion, it is relevant to ask whether the benefits of population screening outweigh the inevitable harms of screening. That question is considered in the following section.

Chapter 6

Harms of population screening

6.1 Radiation exposure

Mammographic screening is a form of radiological examination, involving the exposure of participants to ionizing radiation. In the Dutch screening programme, the average glandular dose has been calculated to be 1.3 milligray (mGy) per (analogue) mammogram of both breasts.[146] According to a US study,[147] the glandular dose associated with digital mammography is one fifth lower than with analogue mammography. For comparison: a radiation dose of 1.3 mGy is the equivalent of three and a half weeks' exposure to the natural background radiation in the Netherlands [1.3 mGy x 0.12 (breast tissue weighting factor) (= 0.156 millisievert): 2.4 millisievert (background radiation per year) x 365 days = 24 days].[148]

Assuming that the participation rate is 100 per cent and that two mammograms per breast are taken only on the first occasion that a woman attends for screening, modelling indicates that the cumulative dose is 18.2 mGy (14 x 1.3).[149] Using that figure as input for the BEIR-VII radiation risk model, an estimate has been made of the number of breast cancer mortality cases attributable to screening in a simulated population of 100,000 women in the age range zero to one hundred years in 1989. Again assuming a 100 per cent participation rate, biennial screening of women aged fifty to seventy-five would result in 1.6 fatal cases of breast cancer per 100,000 women in the simulated population in 1989, while 1,121 cases of breast cancer mortality would be

prevented in the same population.[149] If two mammograms per breast were standard screening procedure, the number of radiation-related deaths would be doubled, to 3.2 per 100,000 (at a participation rate of 100 per cent). Other authors have estimated the risk to be lower, i.e. 2.2 fatal cases of breast cancer per 100,000 women, given two mammograms per breast as standard, despite the assumption of a higher glandular dose (1.8 mGy).[150]

6.2 Overdiagnosis

Overdiagnosis is a significant drawback of population screening for breast cancer and a constant topic of debate. The term implies the detection, as a result of screening, of a malignant tumour that would never have been diagnosed during the patient's life is she had not been screened. Overdiagnosis occurs if, for example, cancer is detected early and the patient subsequently dies of some other cause before the tumour would have caused any problems. Another overdiagnosis scenario is the detection of a tumour that does not progress, or so slowly that it would never reach a symptomatic stage. In both scenarios, early diagnosis and treatment have only drawbacks for the woman in question. However, it is not possible to ascertain whether an individual case involves overdiagnosis. The extent of overdiagnosis can be estimated only at the population level over a period of years.

Overdiagnosis is an inevitable consequence of screening (and diagnostic testing) and constitutes a sizeable problem in screening for certain forms of cancer, such as prostate cancer, thyroid cancer and neuroblastoma.[151-153] Screening for breast cancer also leads to overdiagnosis.

Some believe that many of the cases of breast cancer detected as a result of screening, would regress spontaneously without intervention.[154] However, the Committee has been unable to find any scientific evidence to support that supposition.[155] Cases of spontaneous regression following confirmed breast cancer diagnosis are extremely rare.[156,157] Spontaneous regression has been documented in certain types of childhood cancer (neuroblastomas).[158]

Roughly 20 per cent of the tumours detected as a result of screening involve DCIS (Annex E, Table 2). The natural course of this precursor of breast cancer is uncertain: there is little opportunity to study the natural course of the condition, because it is normally treated. However, studies involving a long observation period, in which 'treatment' was restricted to diagnostic biopsy, have found that more than half of the abnormalities, most of them having a low malignancy grade, do not develop into invasive breast cancer.[159,160] Even if a clearly

differentiated form of DCIS becomes invasive, it does not result in a very aggressive tumour.

6.2.1 *Major differences in estimates*

A systematic review has found that estimates of the extent of overdiagnosis associated with breast cancer screening range from zero to 54 per cent.[161] The divergence is to a large extent due to the use of different definitions for the degree of overdiagnosis.[162] Much depends, for example, on whether the calculated percentage relates to women of all ages, women in the target age group for screening or only 'screening carcinoma' cases (cases of breast cancer detected as a result of screening). The smaller the denominator in the fraction, the greater the percentage, given a constant numerator.

A second reason for the range in estimates is, that some are based exclusively on the introductory phase of the screening programme, without taking adequate account of lead time. By screening breast cancer is detected earlier than would be the case without screening. Obviously, the time by which screening expedites diagnosis (the lead time) cannot be directly measured. Estimates of the average lead time vary from less than two years to more than four years.[163,164] Estimates that account for some tumours not to progress, put the figure at about two and a half years.[80,165-169] This implies that, during the introductory phase of a screening programme, breast cancer is detected more often than usual.

A third explanation are the existing differences in the 'intensity' of screening programmes (age limits, screening intervals, referral thresholds). Moreover, some estimates of overdiagnosis relate to opportunistic screening, rather than population based screening.[86] Opportunistic screening increases overdiagnosis, because it often involves annual examination and women under the age of fifty. Those characteristics increase the likelihood of referral and detection of abnormalities.

Finally, there are major differences between the studies estimating overdiagnosis, in terms of methodology and modelling assumptions.[170]

6.2.2 *Methodology*

The degree of overdiagnosis can be estimated accurately by comparing the incidence of breast cancer in a population that is offered screening with the incidence in a population that is not, provided that the background risk of breast cancer in both populations is the same.[171]

Correction for differences in background risk is possible on the basis of information about risk factors for breast cancer, such as age, use of hormone suppletion therapy, age at first childbirth, number of children and postmenopausal obesity. If the incidence in the non-screened population is derived from the period prior to the introduction of screening, correction must also be made for trends in the incidence of breast cancer. If the incidence in the non-screened population is based on a contemporary control group, correction must be made for geographical differences in incidence in the past.

It is also important to take account of the lead time. During a screening programme's introductory phase, the first round of screening (the prevalence screen) causes a 'prevalence peak', which is more prominent if screening is introduced over a shorter timeframe. If the whole of the target group in a given region is first invited for screening within two to three years of the programme's inception, the prevalence peak can be expected to be two to three times higher than the annual incidence of breast cancer without screening.[165] The rise in prevalence is caused primarily by cases of breast cancer being diagnosed sooner than would otherwise have been the case (lead time); it is not a manifestation of overdiagnosis. Screening brings diagnosis forward in time, as demonstrated by the fact that the incidence of interval cancer following the first round of screening is lower than the incidence that would be expected without screening (Table 5[165]).

To a certain extent lead time effect persists after the first round of screening,[172] due to the intake of women aged 50 and 51, who are able to participate for the first time. As a result, the incidence remains higher than at the outset. The effect is not insignificant; in the Netherlands, new participants make up 11 per cent of all participants.[11,23]

When women reach the age of 75 and are no longer screened, there is a compensatory drop in incidence (deficit incidence), relative to what would have been expected if there were no screening (Figure 1). In actual fact, this deficit is not exactly equal to the earlier prevalence peak, suggesting overdiagnosis. The deficit incidence cannot be fully calculated until several years after the last woman in the age group with the opportunity to undergo screening (who therefore contributed to the surplus) has ceased to be screened. The number of years in question must be greater than the average lead time, since there is a wide range in lead times. Hence, accurate estimation of the degree of overdiagnosis requires a long observation period, ideally extending to the death of the women in question. However, five to ten years after the women have ceased to be screened is accepted as adequate.[50,161,173]

6.2.3 Research results

Research into overdiagnosis started early in the Netherlands. One of the first studies was the Nijmegen pilot project. The (cumulative) breast cancer incidence after six screening rounds (1975 to 1986) was compared with the incidence in the same period in a control group in Arnhem, the same geographical region, where screening had not been introduced.[174] The average overdiagnosis was calculated to be 11 per cent of the incidence to be expected without screening, falling from 30 per cent in the first four years to 1 to 3 per cent in the following two four-year periods. Because the study involved a relatively brief observation period, no correction was made for lead time, meaning that the 11 per cent figure must overestimate the true degree of overdiagnosis.

Figure 1 (Page 62) shows the incidence of breast cancer (invasive and DCIS) in women aged between 40 and 90 in the calendar years 1989, 1999 and 2009.[175] The peak is due to the fact that, from 1999, the target group for screening was extended to include women aged 70 to 75. The underlying data indicate that the lifetime risk of breast cancer in women in the Netherlands rose in this period, from of one in twelve to one in seven.[175] In the graph line for 2009, a clear dip is discernible in the section representing women between 70 and 85. That is attributable to the large number of women in whom breast cancer was already detected at a lower age as a result of screening.

The incidence of breast cancer in women of all ages observed in the Netherlands was modelled and compared with the incidence to be expected without screening.[176] For the period 1990 to 2006, an annual estimate of the degree of overdiagnosis was made to illustrate the significance of the screening phase in which overdiagnosis is measured. During the introductory phase, overdiagnosis peaks at more than 11 per cent (in 1993-1994), before falling to nearly 3 per cent in the stable phase (2006).[176] Averaged over a period of thirty years and taking account of the introduction of digital mammography, overdiagnosis worked out at roughly 3 per cent.[177] If the estimate is based not on women of all ages, but expressed as a fraction of screening carcinomas, the figure works out at a little more than 8 per cent (Table 2 in this study). Estimates for other countries are often considerably higher. One meta-analysis arrived at an overdiagnosis figure of 52 per cent in women in the target age group for screening.[178] An estimate for England and Wales even yielded a figure of 57 per cent overdiagnosis in the period 1993 to 1999 in women in the target age group for screening.[178]

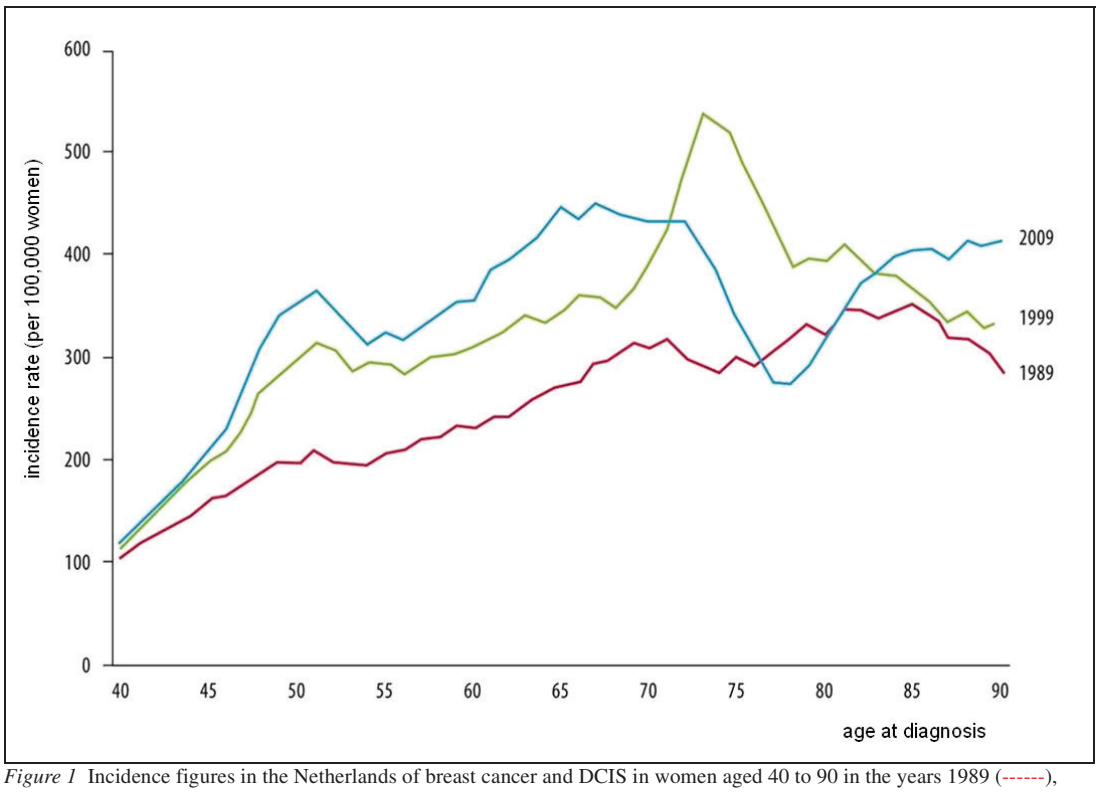

Figure 1 Incidence figures in the Netherlands of breast cancer and DCIS in women aged 40 to 90 in the years 1989 (------), 1999 (------) and 2009 (------). (source: IKN).[175]

These high estimates have been criticized, justified according to the Committee.[89,171,179] For example, the researchers assumed that, in England and Wales, breast cancer incidence was rising linear in women under the age of 65, until screening began in 1988. However, between 1984 and 1988, incidence was in fact increasing exponentially.[180] The Committee sees no good reason to disregard the exponential rise, since it cannot be explained by opportunistic screening or by the screening trial that began in 1979 in the UK. The rise in incidence not only involved women in the target age group for screening, but also women under the age of fifty. By assuming a continuation of the trend prior to 1984, the researchers underestimated the rise in incidence to be expected without screening (and therefore overestimated the degree of overdiagnosis). Furthermore, overdiagnosis was estimated only for the year 1999, while disregarding the preceding years of screening, when the incidence was lower. Finally, the estimate of overdiagnosis in the early years of screening was not

corrected for the deficit incidence in women older than the target age for screening.

A 57 per cent increase in incidence implies that more than one case in three of breast cancer diagnosed in the target age group was an instance of overdiagnosis. Given that, during the period under consideration, only 37 per cent of breast cancer cases came to light as a result of screening, the result of the meta-analysis effectively means that almost all breast cancer diagnoses resulting from screening constitute overdiagnosis.[89] That strikes the Committee as rather unlikely.

Another criticism of the meta-analysis is that it was concerned mainly with screening that was being introduced gradually, causing a prolonged prevalence peak, which is increasingly difficult to distinguish from overdiagnosis. Overdiagnosis can be most adequately estimated when screening is introduced quickly. In Copenhagen, the introductory period was within two years.[181] During the first round of screening in Copenhagen (1991 to 1993, participation rate 71 per cent), the prevalence peak was nearly twice as high as the expected incidence without screening, while the subsequent incidence was barely any higher than prior to the introduction of screening, allowing for the continuous intake of women aged 50 and 51 joining the programme in subsequent rounds.[181] These findings illustrate the implausibility of 52 per cent overdiagnosis.[178]

The authors of an earlier study, for which the screening regions Copenhagen and Funen were compared with the rest of Denmark (where there was no screening), concluded that overdiagnosis was 33 per cent.[182] However, the study only concerned the first ten to thirteen years after the introduction of screening, and disregarded the fact that, before screening, breast cancer incidence in Copenhagen and Funen was 8 per cent higher than in the rest of Denmark. After correction for that difference, the incidence in the target age group for screening was 25 per cent higher (386/286 per 100,000:1.08=1.25). The latter surplus fell to 16 per cent in the stable phase of the screening programme (2001 to 2003), while in women aged between 70 and 80 14 per cent fewer cases of breast cancer were detected in the screening regions: (327/367 per 100,000):(273/264 per 100,000)=0.86.[171] In other words, the incidence surplus in the stable phase is largely offset by an incidence deficit in women older than the screening age. Even then, the figure of overdiagnosis is overestimated, because these relatively simple calculations do not account for the rise in background risk. This was accounted for in a cohort study with a longer observation period (1991 to 2009).[173] That study showed a figure of 4 per cent overdiagnosis in the target age group for screening and 2 per cent in the subgroup with at least eight years follow-up after the cessation of screening.[173]

In a Norwegian study overdiagnosis of invasive breast cancer (i.e. excluding DCIS) in women in the target age group for screening (50 to 69) was estimated at 15 to 25 per cent during the introductory phase of the population screening programme (1996 to 2005).[183] And Table 2 of this study aimed to show, that the overdiagnosis estimates were as high as 30 to 50 per cent in some cases. However, this research raised debate.[184,185] First, it considered only the introductory phase of screening and not the subsequent stable phase. Furthermore, the short observation period will inevitably have led to underestimation of the deficit incidence.

Studies in which proper correction for lead time and background risk yield considerably lower figures: between 1 and 10 per cent overdiagnosis relative to the expected incidence without screening.[89,168,169,186-188] The Dutch study referred to earlier, with 3 per cent overdiagnosis (corresponding to just over 8 per cent of screening carcinomas), is within that range.[177]

6.3 Overtreatment

It has been suggested that screening leads to 20 per cent more breast amputations.[3,5,189,190] That suggestion[5] is based on Canadian and Swedish screening trials carried out when breast amputation was the standard treatment for breast cancer. While screening does lead to overdiagnosis and therefore inevitably also to overtreatment, the amount of overtreatment is not as great as suggested.

During the early years of population screening in the Netherlands, the number of breast amputations per 100,000 women aged 50 to 69 did indeed rise, as a consequence of the prevalence peak discussed above.[75] After 1993, however, the number fell, and since 1998 it has been lower than prior to population screening, in both absolute and relative terms. The same pattern was observed in women above the age of 70, while in women under the age of 50 there was a small rise.

In Italy, where the population screening was extended between 1997 and 2001 to include seventeen regions, the frequency of breast amputations decreased from 110 to 89 per 100,000 women between the ages of 50 and 70.[99] In the Irish Republic, in the first two years after population screening began (2000 and 2001) the number of breast amputations increased from 149 to 190 per 100,000 in women aged between 50 and 65 who were invited for screening, to subsequently decrease to 98 per 100,000 in 2009.[191] In Norway, the number of breast amputations decreased from 156 per 100,000 women in the target age group for screening during the years before the introduction of screening (1993

to 1995), to 106 per 100,000 after the introductory phase (2005 to 2008).[192] It may be argued that breast amputation would have become less frequent even without population screening. Nevertheless, a contemporary comparison in Norway during the introductory phase (1996 to 2007) found that, of the women invited for screening in whom breast cancer was detected, 48 per cent underwent breast amputation (and 38 per cent of participants diagnosed with breast cancer), whereas 58 per cent of women with breast cancer who had not been invited for screening underwent amputation.[97] Similarly, in Northern Ireland, screening was not found to lead to more breast amputations.[193]

6.4 Excess mortality due to causes other than breast cancer

6.4.1 Short and medium-term effects

The efficacy of breast cancer screening could be overestimated if measured exclusively on the basis of breast cancer mortality (slippery-linkage bias).[194] This bias occurs because any mortality caused by diagnostic or therapeutic procedures associated with screening cannot be calculated as breast cancer mortality. Do the original screening trials indicate slippery-linkage bias? Does general mortality (death from all causes) fall in line with the observed breast cancer mortality reduction?

The screening trials concluded that the availability of screening reduced breast cancer mortality by 20 per cent (or 25 per cent in women above the age of fifty). In the trial control groups, breast cancer accounted for 3 per cent of all mortality.[52] That figure is lower than the percentage in the general population (4.6 per cent in the Netherlands, www.cbs.nl), which is due to the fact that women with breast cancer were excluded from participation in the screening trials. Naturally, because screening is intended for women that exhibit no apparent symptoms of breast cancer.

On the basis of the trials, the invitation for screening should reduce general mortality by 1 to 2 per cent. The data confirm that this indeed is the case. The Swedish overview found that, after a follow-up of nearly sixteen years, general mortality was 2 per cent lower in women to whom screening had been made available: RR=0.98 (0.96-1.00).[52] That would suggest that screening does not result in excess mortality due to causes other than breast cancer.

The harms of overdiagnosis and treatment can manifest themselves only in women in whom breast cancer is detected as a result of screening. Consequently, the mortality pattern in such women provides a more accurate picture of any harm that screening may cause than the mortality pattern in all women invited for

screening.[47] Analyses of data on participants in the Two-County Trial revealed no difference, in terms of mortality due to causes other than breast cancer, between women with breast cancer invited for screening and women with breast cancer in the control group (age at the start of the trial 40 to 74).[139,165] Hence, the analysis found no evidence of excess mortality due to causes other than breast cancer.

In the Two-County Trial, after a twenty-year follow-up, general mortality in women with breast cancer who had been invited for screening was 13 per cent lower. That relative mortality reduction was in line with the 31 per cent relative reduction in breast cancer mortality observed in this trial.[139]

6.4.2 *Long-term effect*

Roughly 70 per cent of patients with invasive screening carcinomas receive radiotherapy after surgery, and more than 50 per cent receive adjuvant hormonal therapy and/or chemotherapy. Nearly all patients with invasive breast cancer who undergo breast-conserving surgery subsequently receive radiotherapy, as do roughly a third of patients with DCIS who undergo breast-conserving surgery.[75] Those supplementary therapies have considerably improved survival but can also have harmful long-term effects.[78,195,196]

The provision of radiotherapy to breast cancer patients reduces the breast cancer recurrence risk, breast cancer mortality and general mortality (death from all causes collectively).[196] Nevertheless, radiotherapy can increase the long-term risk of mortality due to cardiovascular disease, lung cancer and oesophageal cancer.[195,197] Some commentators have suggested that the harmful effects of screening (as associated with overtreatment resulting from overdiagnosis) could even negate what they regard as the marginal beneficial effect of screening.[198,199] However, that suggestion is based on the assumption that overdiagnosis is very high. Mortality due to cardiovascular disease in particular has in the past diminished the health benefit of radiotherapy. The radiation dose that the heart receives when radiotherapy is given for a tumour in the left breast is roughly twice as high as when a tumour in the right breast is treated.[200-202]

The observation that radiotherapy leads to excess mortality due to cardiovascular disease is derived from studies that began between 1960 and 1990. Since then, however, radiotherapy techniques have improved substantially. Developments such as the introduction of intensity-modulated radiotherapy (IMRT), breath-hold techniques and treatment in prone position have significantly reduced the exposure of the heart, coronary arteries and lungs.[203-206] In a US study conducted after 2005, the average cardiac dose associated with

prone radiotherapy (1.4 Gy) was less than a third of the dose in a Scandinavian study covering the treatment period 1958 to 2001.[200,201] In cases of cancer in the left breast, the average cardiac dose associated with prone radiotherapy (1.0 Gy) was half that associated with supine therapy (2.2 Gy).[200] In addition, partial breast irradiation – irradiation of only the region of the tumour, as opposed to the whole breast – is used to reduce the radiation exposure in cases where there is a low risk of local recurrence. The initial results obtained using the technique are promising; the results of large trials are expected within a few years.

The reduction in radiation exposure means that the harms have also reduced. Since the end of the seventies, mortality due to cardiovascular disease attributable to radiotherapy has declined sharply year on year.[207] In patients that received radiotherapy in the eighties, there was no longer any over-mortality due to cardiovascular disease, at least within a period of twelve years.[207] However, the harms concern long-term effects of radiotherapy, research with a longer follow-up is required. A large study of Danish and Swedish patients treated between 1976 and 2006, which involved a follow-up of up to thirty years, found no over-mortality due to cardiovascular disease attributable to radiotherapy.[202] A study in the Netherlands with a follow-up of eighteen years demonstrated that the risk of cardiovascular disease for patients treated between 1980 and 1986 was lower than in the seventies (and no longer statistically significantly elevated).[208]

According to some studies, radiotherapy received relatively early in life entails greater risk than radiotherapy later in life.[202,208] Measured over the period 1970 to 1986, patients over the age of 55 were not found to be at elevated risk of cardiovascular disease due to radiotherapy.[208] However, a recent Swedish-Danish study indicated that the risk per gray of radiation does not vary with age.[201] Nevertheless, caution with radiation exposure remains advisable. It is not yet clear whether there is a safe level of radiation exposure below 2 Gy.

Following early-stage breast cancer, adjuvant therapy considerably reduces the risk of local recurrence and breast cancer mortality.[78] However, medications such as Adriamycin can damage the heart. Excess mortality due to causes other than breast cancer is, as far as can be established after a fifteen-year follow up, relatively small – at most one or two cases per thousand treated patients.[196]

6.5 Adverse effects on quality of life

Individually mammographic examination hardly substantially influences the quality of life, but population based screening does affect a large number of people. Participants may find the examination stressful or worrying. Half of participants find the mammography procedure uncomfortable or even

painful.[209,210] For a clear image, the breast has to be compressed quite firmly during mammography. On the other hand, screening provides the vast majority of participants (98 per cent) with the reassurance that they exhibit no abnormalities. Initially, some authors argued that, in the longer term, population screening could lead to a disproportionate fear of breast cancer. However, in the medium term, screening does not appear to increase psychological morbidity in participants (as measured by the General Health Questionnaire).[209]

6.5.1 *The harms of (false) positive results*

If an abnormality is detected on a screening mammogram, the woman is referred to a breast clinic for further diagnostic testing and possibly treatment. Most women find the diagnostic process and, in particular, the waiting extremely stressful, even if imaging research and, where appropriate, needle biopsy (for tissue testing) ultimately reveal no evidence of breast cancer.[211] Therefore, it is important to arrange minimized waiting times with the breast clinics.

'False positive' results constitute a serious drawback of screening, without compensatory benefit for the affected subgroup of participants. A study found that, in nearly two thirds of false positive cases, non-invasive diagnostic imaging was sufficient to establish that cancer was not present.[11] During the diagnostic phase, depression and anxiety levels increase amongst the affected women, especially if biopsy is needed. Feelings of anxiety have typically subsided one month after a woman has been told that there is no evidence of cancer.[12,211] Longer term effects vary. Research using disease-specific outcomes has found that screening can have adverse psychological effects for up to three years, particularly if invasive follow-up diagnostic procedures are required.[212,213] However, when the status of women who had received a negative (i.e. favourable) screening result and the status of women who had received a false positive result were assessed using generic tools for the measurement of anxiety and depression, no difference between the two groups could be detected after six weeks or three months.[212]

Of note, more than half of the women with false positive results, after being diagnosed with a 'benign breast abnormality', kept returning for outpatient check-ups, sometimes as many as eight times in the first year.[211,214] According to a study in the south-eastern part of the Netherlands, 21 per cent of women with false positive results, were still receiving outpatient care four years later.[215] That observation begs the question: do these women attend because they are stressed or anxious and, if so, does the availability of outpatient checks constitute adequate support? Or should this be a task for their GPs? In view of the high

frequency, it would appear that most of these outpatient visits cannot constitute a medical necessity. The Committee believes that further research is needed to improve the support for women receiving false positive results.

European studies indicate that the receipt of a false positive result has little or no adverse effect on subsequent participation.[12,212] In the above-mentioned study in the south-eastern part of the Netherlands, participation was affected.[215] However, that may due to the tendency to keep women under outpatient supervision, following false positive screening results. This constitutes a deviation from the screening programme.

6.5.2 *Likelihood of a false positive result*

In 2012 in the Netherlands, 24 in every thousand screening participants were referred for further examination (Annex E, Table 2).[23] In 27 per cent of referred cases (6.3 in every thousand screening participants), the suspicion of breast cancer is confirmed. In the other 73 per cent (17.2 in every thousand participants), the screening result was therefore 'false positive'. In 62 per cent of false positive cases (10.6 in every thousand participants), diagnostic imaging was sufficient. Nearly a third of women with false positive results underwent invasive diagnostics (biopsy).

Over the last decade, the FPR in the Netherlands has risen from 0.5 per cent in the nineties to 17.2 in every thousand participants (1.7 per cent) in 2012. That rise is the downside of a (successful) policy geared to increase the detection rate and to reduce the number of interval carcinomas. Despite being a lot higher than ten years ago, the FPR in the Netherlands is exceptionally low, certainly compared with the United States, where (in women above the age of fifty) it is more than 20 per cent for a first screening and 9 per cent for a subsequent screening.[216]

Compared with other European countries with programmatic population based screening, the Netherlands also performs well (i.e. has a low FPR), while attaining a comparable detection rate.[12-17] In the United Kingdom, the FPR was found to be 3.4 per cent in 2010/2011: 7.2 per cent for a first screening and 2.3 per cent for a subsequent screening (http://www.cancerscreening.nhs.uk/breastscreen/publications/2012review.html). In Norway, the FPR was 3.5 per cent during the first round of screening and 2 per cent in the third round, discounting the additional 0.8 per cent of cases referred due to the mammogram being of insufficient quality or symptoms spontaneously reported by the participant.[14] In Spain, after extrapolation to yield a figure for women above the

age of fifty, the FPR was found to be 7 to 11 per cent in the first screening and 4 to 5 per cent in subsequent screening.[17,217,218,219]

Consequently, the risk of a woman regular participating in breast cancer screening receiving a false positive screening result at some time is quite substantial. In Florence, for example, after attending seven rounds of screening, a woman would have a 15 per cent chance of receiving a false positive result at some point.[16] Research in Norway and Spain found that, after attending ten rounds of screening, a woman had a 20 per cent chance of a false positive.[14,17] In the Netherlands, for a woman of fifty attending all thirteen rounds of screening (with digital mammography) the chance of a false positive is estimated to be 15 to 16 per cent.[75,220]

6.5.3 *The harms of false negative results*

A further drawback of screening is the risk of false negative results. Contrary to what the term might suggest, a 'false negative' does not necessarily involve an assessment error. A false negative is merely a negative result that is followed by a breast cancer diagnosis within two years (the screening interval), referred to as 'interval cancer'. Between 2004 and 2009, interval cancer was detected in 11,855 women (2.3 per thousand screened women, or per 2,000 woman-years), while 29,530 screening carcinomas were recorded. That yields a numerical ratio of 1:2.5 and a programme sensitivity of 71.4 per cent (Annex E, Table 2).

For quality assurance purposes, the LRCB reassesses a certain proportion of the mammograms of women in whom interval cancer was later detected.[221] The reassessed mammograms are placed into one of three categories: I. 'no significant abnormalities' (radiologically occult or non-manifest tumours), II. 'minimal signs' (minimal abnormalities, insufficient to have justified referral), and III. 'significant abnormalities' which would have justified referral. This review process shows that in half of interval cancer cases no significant abnormality was visible on the screening mammogram.[222] In a quarter, there were minimal signs and in the other quarter there were abnormalities that, in hindsight, should have led to the woman's referral.

When the pre-diagnosis screening mammograms are reassessed, the assessor is not told where the tumour appeared on the diagnostic mammogram. However, the procedure is not perfect, because the assessor does know that cancer was detected in the woman. The method has been in use for twenty-five years, including for cases involving advanced screening tumours. Nevertheless, it is clear that, the more prior knowledge the assessor has, the more likely he or she is to find abnormalities when reassessing a mammogram.[223,224] If an assessor

knows there is something to find, he or she looks at a mammogram in a different way than when assessing a mammogram in the context of primary screening (hindsight bias). Hindsight bias can be reduced by giving the reassessor a mixture of false negative mammograms and true negatives, preferably in proportions that correspond to the screening programme. The Committee believes that it is desirable to develop a review method that excludes prior knowledge.

The drawback of a false negative result is that the woman in question has gained nothing from undergoing screening. Furthermore, having been told that there is no evidence of cancer, a woman and her doctors may be less alert and slower to act when problems or symptoms develop that might point to breast cancer. Patient delay can lead to therapy starting later and to a poorer prognosis for the patient than would have been the case without screening. Although it has often been suggested that screening may lead to false reassurance,[209,225] the phenomenon has been studied only in Rotterdam. The researchers found that breast cancer patients who had undergone screening did not delay visiting their doctors any longer after first discerning breast abnormalities than patients who had not been screened.[162]

Other harms of false negative results can be: disappointment and incomprehension, sometimes leading to legal proceedings and, following the detection of interval cancer, loss of confidence in screening. The Committee is not aware of any research into those harms.

6.5.4 *The harms of true positive results*

Screening results in earlier diagnosis. While that can be beneficial in terms of life-years gained, it can also be a drawback, if a woman has to deal with the knowledge that she has breast cancer for several more years, without the prognosis being improved.

It has been suggested that only 3 to 13 per cent of the women in whom breast cancer is detected as a result of screening actually benefit in terms of life-years gained.[226] That is based estimates from the United States, where opportunistic screening is practised. Given the optimal implementation of population base screening and follow-up in the Netherlands, the calculations indicate that roughly 26 per cent of women in whom breast cancer is detected as a result of screening benefit from the diagnosis,[176] Because of the screening programme they will not experience the terminal phase of the disease and gain an average of 16.5 life-years.

In the other 74 per cent, although breast cancer is detected several years earlier than it would have been without screening, they experience no health benefit in terms of gained life years. Several would have survived the disease even without screening-related early detection (56 per cent), others die despite screening-related early detection (13 per cent), and others still would never have been diagnosed without screening as they would have died of some other cause before their cancer became symptomatic (6 per cent*).[176]

It is reasonable to assume that screening-related early diagnosis leads to more restricted treatment, but the Committee is unaware of any research on that topic.

6.6 Conclusion

A significant drawback of breast cancer screening is overdiagnosis: the detection of tumours that without screening would never have come to light. Estimates of the extent of overdiagnosis vary considerably and have been the subject of much debate. The Committee concludes that overdiagnosis will be substantially overestimated if insufficient account is taken of the lead time, the background incidence of breast cancer or the compensatory incidence drop in breast cancer in women after exiting the screening programme. For the Netherlands, overdiagnosis is estimated to be roughly 3 per cent of the expected incidence of breast cancer without screening, or over 8 per cent of screening carcinomas.

Overdiagnosis inevitably leads to overtreatment. Additionally it has been suggested that population screening leads to 20 per cent more breast amputations. However, the latter claim has been shown to be unfounded.

The Committee considers it unlikely that treatment for breast cancer leads to excess mortality due to causes other than breast cancer in the long term. However, because the harms of radiotherapy and adjuvant therapy can occur after a long time, research into the long-term consequences of treatment remains desirable.[197,201,208]

Screening results in earlier diagnosis. On the downside a women with breast cancer will have to live with that knowledge for several years longer than they would have without screening. In three of every four cases, there is no compensatory benefit in terms of life years gained. Quality of life can also be adversely affected by the anxiety and uncertainty associated with an abnormal screening result and the subsequent diagnostic testing, even if no breast cancer is ultimately detected. The risk of a false positive result is relatively small in the

* With digital mammography, 8 per cent of screening carcinoma cases involve overdiagnosis.[176]

Netherlands (1.7 per cent in 2012) but can rise to an estimated 15 to 16 per cent for women participating every two years over a period of 26 years.

Chapter 7

The efficiency of population screening

The preceding sections of this report have described, in turn, the benefits and harms of screening. This section addresses the minister's question regarding the relationship between the benefits and harms of the existing Dutch population screening programme. The Committee also considers the cost-effectiveness of population based screening in the Netherlands.

7.1 Number needed to screen

The efficiency of population screening depends to a considerable extent on the screening interval. In the Dutch programme, the interval between examinations is set at two years. In the trials carried out to determine the effectiveness of screening, the screening interval varied from 12 to 33 months. There is little evidence that annual screening of women above the age of fifty is significantly more effective than biennial screening, but annual screening inevitably means roughly twice the harms, such as false positives and unnecessary biopsies. If the interval is extended to more than two years, the efficiency quickly declines as the risk of interval cancer increases.[227,228]

Modelling indicates that, in the Netherlands, population based screening prevents an average of 775 breast cancer mortality cases per year (683 estimated for 2008 and 858 for 2018).[10,33] Roughly 1,200 women must each undergo screening once to prevent one death from breast cancer. That observation is broadly in line with, for example, the finding of the Swedish Two-County Trial,

that one fatality is prevented for every 1,300 to 1,700 screening examinations (for women aged *forty* and above).

According to a Cochrane review, at least 2,000 women must be invited for screening for ten years to prevent one case of breast cancer mortality.[5] Given a participation rate of 80 per cent and biennial screening, that equates to at least 9,600 (2,000x6x0.8) screening examinations (as opposed to 1,200).

That much more negative estimate is based on a meta-analysis of the few screening trials that were subjectively judged to be of sufficient quality (Malmö, UK Age and the two Canadian trials). Furthermore, the differences arise because of the duration of the screening and follow-up and the age of the screened women. The results of the Two-County Trial illustrate the effect of those factors. With a follow-up of up to 29 years, the trial found that between 400 and 500 women needed to be screened throughout the seven-year experimental screening programme to prevent one case of breast cancer mortality.[81] That ratio is referred to in the literature as the NNS: number of women needed to screen to prevent one breast cancer death. The NNS would be twice as high (indicating a lower level of efficiency) if calculated using a follow-up of only ten years.[81] The figures presented in the Cochrane review are based on a follow-up of seven years.[5]

The screening period is also important. If the screening programme had run for ten years instead of seven, the NNS would have been 300.[81] Age at the time of screening makes a difference as well. Based on data for women between the ages of 50 and 70, the NNS for a seven-year programme is just over 300, whereas with women aged 40 to 74 it is 400 to 500.[89]

7.2 Cost-effectiveness and cost-utility

A key outcome indicator for efficiency is cost-effectiveness. The cost-effectiveness of a population screening programme is a statement of its added value relative to an alternative (e.g. not screening), expressed as extra cost per life-year gained. Cost-effectiveness analysis involves the development and validation of a computer model which combines data on the natural course of a given disease, the associated disease burden and mortality in a given population, the benefits and harms of screening over a long period of time, and the expenses and cost savings attributable to the screening programme.

A cost-utility analysis goes a step further. Life-years gained are corrected for quality of life and expressed as QALYs: quality-adjusted life years. QALYs are calculated by multiplying life-years gained by a quality-of-life factor (utility) of between 0 and 1. Adverse effects on quality of life attributable to breast cancer screening and treatment are partially offset by the avoidance of the terminal

phase of the disease and the associated palliative treatment. Early detection means that less extensive primary treatment is required.

To secure health benefits and savings, it is necessary to incur costs (medical and non-medical, tangible and intangible). Where cost and effects occur not immediately but at a later date, they need to be discounted. That involves adjusting the value assigned to the costs and effects on the basis of their timing. Discounting serves to level out costs and effects in line with the fact that people generally prefer to realize benefits as soon as possible and to defer costs as long as possible (positive time preference). Cost and effects are discounted at a fixed annual rate, which level has been the subject of debate. The Pharmacoeconomic Guidelines published by the Health Care Insurance Board suggest discount rates of 4 per cent for costs and 1.5 per cent for health effects.[229,230] The UK Treasury applies a uniform rate of 3.5 per cent for both costs and effects.[230,231]

Modelling indicates that programmatic population based screening can reduce breast cancer mortality at a reasonable cost. The cost-effectiveness of biennial screening of women aged 50 to 70 (or 75) varies from 1,600 to 24,400 euros per life-year gained.[33,232-236] The range reflects not only international differences in disease burden, mortality and programme design, but also differences in the models used and the assumptions made. In the Netherlands, the National Advisory Council for Public Health has defined 80,000 euros per QALY as the upper limit for efficient collectively funded care, depending on the seriousness of the disease.[237]

The most recent estimate of the cost of breast cancer screening in the Netherlands is 1,600 euros per life-year gained.[235] The figure includes the direct costs incurred within the health care system (screening, diagnosis and treatment), as paid by the community. It does not include the direct costs incurred outside the health care system (the patient's travel, time and other expenses), indirect costs incurred outside the health care system (paid and unpaid work) or indirect costs incurred within the health care system (cost of associated diseases, expressed in life-years gained).

The cost-effectiveness of the programme in the Netherlands compares well with other countries. That is probably due to the background situation with a high historical level of breast cancer mortality, the high quality of the screening programme with a low referral rate combined with a high detection rate and therefore a high positive predictive value,[238] and the high participation rate.

Table 1 in this study contains data on the population based screening programme for breast cancer in the Netherlands compared with the existing (cytological) screening programme for cervical cancer in the Netherlands, the new-style

Table 1 Cost-effectiveness of population screening in the Netherlands. Euros per life-year gained.

	Discount rate	
	4%/1.5%	3%/3%
Cervical cancer (current programme, 7 rounds of cytology)[239]	5,900	11,300
Cervical cancer (GR recommended, 5 rounds of HPV testing)[239]	4,100-4,600	5,100-8,700
Breast cancer[235]	1,600	3,700
Bowel cancer[240]	2,600	2,200

cervical cancer screening programme proposed by the Health Council involving HPV testing and the recently started colorectal cancer screening programme.

7.3 Relationship between benefits and harms

Modelling has been used to estimate the efficacy of population based screening in the Netherlands.[41] It was estimated that screening prevents between 683 (in 2008) and 858 (in 2018) breast cancer mortality cases per year. On average these women gain 16.5 life-years and therefore, in total, roughly 11,000 to 14,000 life-years are gained. Screening has beneficial and adverse effects on the quality of life. Beneficial is the prevention of breast cancer mortality and the associated need for palliative treatment. On the other hand, there are various adverse effects. For every breast cancer death prevented, 1,200 women will undergo a screening examination, all but one of them without health benefit, other than the reassurance that no suspect abnormalities were found. In addition, for each breast cancer death prevented, 23 women per year are referred for further testing, of whom 16 prove to have received a false positive result. Of the seven true positive women, five gain no health benefit, excepting that early diagnosis may lead to less intensive treatment. Of the five women who do not benefit, 3.7 would have survived their breast cancer even without screening; 0.9 are destined to die from their breast cancer despite screening and 0.5 would never have been diagnosed with breast cancer if it were not for screening (overdiagnosis).

What does the foregoing ultimately mean in terms of the loss of quality of life attributable to screening? No recent data can answer that question. On the basis of a Dutch study, it has been estimated that correction for quality of life reduces the health benefit in terms of life-years gained by 4.7 per cent.[241]

In the debate regarding the efficacy of screening the relationship between the amount of breast cancer mortality prevented and the amount of overdiagnosis, is considered important. According to the estimates above, the number of women

whose death from breast cancer is prevented is roughly twice the number of women affected by overdiagnosis. Modelling of future screening effects indicated that the ratio was 3 to 1.[177] A study based on UK and Swedish data put the ratio at between 2 and 2.5 to 1 in women aged 50 to 69.[89]

The figures cited in the last paragraph paint a more favourable picture than the estimate of 1 to 3 made on the basis of other calculations for screening in the UK (women aged 50 to 69, screening interval of three years, participation rate 70 per cent).[50] The latter estimate assumes less effective screening (a relative breast cancer mortality reduction of 20 per cent, rather than the 25 per cent indicated by the meta-analysis) and a greater degree of overdiagnosis: 19 per cent of breast cancer cases in women aged fifty who are offered screening for twenty years.

7.4 Conclusion

In order to accurately assess the effects of screening, one needs to know how long screening continued, what the age of the target group was and what the length of the follow-up period was. Taking those factors into account, in the population based screening programme in the Netherlands roughly 1,200 women need to undergo mammography for every breast cancer death prevented. Cost-effectiveness analysis indicates that the cost per life-year gained amounts to 1,600 euros. The existing population based screening programme prevents 775 breast cancer deaths per year (the average of 683 in 2008 and 858 in 2012). That is between two and three times as many as the number of overdiagnosis cases.

Even now that many breast cancer patients receive adjuvant therapy, the beneficial effect of screening on breast cancer mortality remains substantial and significantly greater than the harms, such as overdiagnosis.

The following sections describe the developments that may further improve the efficacy of breast cancer screening in the short or longer term.

Chapter 8

Ways to improve screening in the short term

8.1 Alternative referral strategy

Currently in a randomized trial an alternative referral strategy for women whose screening result is BI-RADS category 0 (www.lrcb.nl/MASS-trial) is being researched. BI-RADS category 0 implies that the screening mammogram provides insufficient information for assessment. In such cases, further investigation is required, but the suspicion of breast cancer is low (PPV = 14 per cent).[20] More than half of all referred women, 24 per cent of screening carcinomas and 38 per cent of screening tumours smaller than 1 centimetre in diameter fall in this category.[20]

In more than half of women referred with BI-RADS-0, breast cancer can be excluded by imaging alone; invasive diagnostic procedures are not required.[20] Nevertheless, the current procedure involves the woman attending a breast clinic for examination.

The trial is investigating a new approach where, in the event of woman with a BI-RADS-0 result, further imaging is undertaken as soon as possible for final diagnosis, thus reducing anxiety, worry and quality-of-life impairment.

An approach that limits the number of referrals to a breast clinic can also limit the increasing pressure on the health care system. According to the Mamma Carcinoma Guidelines of the National Breast Cancer Forum (NABON), more than 90 per cent of referred women should be examined at a breast clinic within five working days.[242] That target is by no means attained everywhere; in some

cases, women have to wait five to six weeks.[21] In recent years, the lowering of the referral threshold and the introduction of digital screening mammography have driven up the referral rate from less than 1 per cent to 2.4 per cent (Annex E, Table 2).

8.2 The lower age limit of fifty

The screening for breast cancer of women under the age of fifty has always been controversial. Meta-analyses of the original screening trials showed that only after a longer follow-up period there was a statistically significant reduction in breast cancer mortality in women aged between forty and fifty as well.[5,243] Nevertheless, in both relative and absolute terms, the mortality reduction is less than in women above the age of fifty. The findings of the more recent UK Age Trial did not shed any further light on the matter: a statistically non-significant result was obtained: RR=0.83 (95% CI 0.66-1.04).[244]

In most screening programmes, the lower age limit is 50. In the United Kingdom, the limit has recently been decreased to 47, even though this change was not supported by the UK Age Trial .[244] Notably, the United States is moving in the opposite direction. Until recently, almost all the organizations involved in screening recommended starting at the age of 40, but in 2009 the US Preventive Services Task Force concluded that screening women under the age of 50 had little value and in general should not be done.[227]

A number of new developments have put lowering the age limit back on the agenda. First, the incidence of breast cancer in women above the age of 40 (particularly in those aged between 45 and 50) has risen markedly in recent years.[75] Second, the performance benefits of digital screening mammography are expected to be more significant for younger women.[245]

Against that background, a literature study has been conducted for the RIVM to assess the benefits and harms of lowering the age limit.[73,246] The study found that recent observational research in the Netherlands and Sweden indicate that analogue screening of women under the age of 50, particularly those between 45 and 50, is more effective than suggested by the randomized trials.[247,248]

In June 2010, the last of the Dutch screening units switched to digital mammography. It is not yet clear whether digital mammography is more sensitive for detecting breast cancer in women under the age of fifty than analogue mammography or whether it will increase the reduction in breast cancer mortality attributable to screening.[338]

However, lowering the age limit has potential drawbacks: more false positive results and overdiagnosis than associated with the screening of women over

fifty.[17,217] The potential drawbacks have not yet been adequately quantified and cannot readily be extrapolated from the results of digital mammography in the over-fifties or from the available data (mostly from the United States) on women under the age of fifty.

If the age limit was lowered, a screening interval of two years would appear appropriate.[247-250] The radiation exposure associated with digital screening in the Netherlands is very low and should not be a decisive factor in the debate.[149] All things considered, the Committee is unable to say what the relationship between the benefits and harms is likely to be. Overall, in screening below fifty the relationship is less balanced than in screening above fifty,[243,250] except for women below fifty (e.g. 46) with a clearly elevated risk of breast cancer.[250] On the other hand, there will be women above fifty with a lower risk than some women below fifty. The Committee will return to that point in subsection 9.1.

The conclusion of the literature study was that, while there are arguments in favour of lowering the age limit, such as the increased incidence of breast cancer in women below fifty,[246] the scientific evidence is insufficient for a proper assessment of the efficacy in the Netherlands. It is therefore important to link a knowledge-deepening study to the screening programme.[73] Two possible scenarios have been described. First, a one-off expansion of the population based screening programme to include one round of screening for half of all 48 year-old women in the Netherlands. Second, a regional pilot project in which women are invited from the age of 44.

The Committee concludes that there is insufficient evidence to support lowering the age limit. Furthermore, the Committee doubts whether the proposed knowledge-deepening studies are any more likely to yield a positive result than the UK Age Trial.[244] According to a recent cost-effectiveness analysis, lowering the age limit has a favourable cost-effectiveness ratio.[235] However, because primarily there is uncertainty about the effectiveness and the utility-risk ratio, a cost-effectiveness analysis of lowering the age limit can provide little useful evidence.

Given that the efficacy of screening in women below fifty whose risk of breast cancer on average is less favourable than the ratio of screening in women above fifty, a knowledge-deepening study focused on women below fifty with a clear elevated risk seems more valuable. The benefits of screening are more likely to outweigh the harms in high-risk groups than in the average-risk general population.[250] Such a study could be part of research into risk stratification if it included women below fifty (see also 9.1).

8.3 The upper age limit of seventy-five

In most screening programmes, women are invited for screening until they reach the age of 70. Only in England, France, the Netherlands and Sweden the upper age limit is higher (73, 74 or 75). England has a system of self-referral, in which older women may participate in screening if they take the initiative to make an appointment. Until 2012, screening in England stopped at the age of 70 and 4 per cent of women above the age of 70 took advantage of the opportunity to participate on their own initiative.[251] In the Dutch population based screening programme, the participation rate in women aged 70 to 75 is 78 per cent.[75]

There are no compelling scientific arguments for continuing screening above 75; there are barely any data on the effectiveness or efficacy of screening of women above the age of 75. Overall they clearly have a shorter life expectancy than younger women and a greater chance of having other diseases and dying from these. Screening would then more likely lead to overdiagnosis than in younger women. Most studies, but not all,[252] indicate that the speed of breast cancer growth declines with increasing age. In line with that observation, the frequency of interval cancer declines with age.[75,253,254]

One argument for raising the upper age limit is that many women of 75 still have a good life expectancy; half of them still have ten years or more to live. Research in Nijmegen in the Netherlands found that, amongst women who were able to participate in screening and had reached the age of 75, more than 3 per cent developed breast cancer before they reached the age of 85 and 30 per cent of those women died from the disease within ten years of diagnosis.[255] The study may however not reflect the current situation. It covers a period of more than thirty years (1975 to 2008), during which the treatment of breast cancer has improved and screening has been introduced and then extended to include women up to the age of 75. Those developments have brought about a significant decrease in breast cancer mortality in women aged between 75 and 85 (www.ikn.nl).

The main evidence for the effectiveness of screening in older women comes from retrospective studies of the correlation between recently having undergone screening and survival duration in breast cancer patients.[256,257] However, survival is not an adequate outcome indicator for research into the effectiveness of screening, because the results are liable to several forms of bias (lead-time bias, duration bias, selection bias), which are not corrected for in the studies. That is illustrated by the fact that survival in patients with diseases other than breast cancer was also found to be greater in older women diagnosed with breast

cancer as a result of screening.[257] Case-control research with an appropriate outcome parameter (breast cancer mortality) provides evidence for the effectiveness of screening below the age of 75, but not above that age.[258,259]

The Committee observes that no data are available from randomized trials into the effectiveness of screening above the age of 75. Observational research using an appropriate outcome parameter does not indicate that screening above 75 is effective. It is not known whether the benefits of screening after the age of 75 outweighs the harms. The Committee concludes that there are no sound scientific arguments to raise the upper age limit of 75.

8.4 Screening interval

The interval between two successive screening examinations is set at two years (see subsection 7.1). The aim is to invite over 80 per cent of the target group 'on time', i.e. within 24 ± 2 months of their previous examination.[75] In practice, it does not always prove possible to bring the mobile screening unit to the correct location on time. Furthermore, about 30 per cent of invited women exercise the option of changing the suggested screening date. In addition, women who move home can find themselves on a different screening schedule.

In 2011 the proportion of women invited on time decreased to 72 per cent, with a broad range between the screening regions (41 to 91 per cent).[11,23] In 2007, the actual screening interval was at least 27 months for 12 per cent of repeat examinations; the interval was at least thirty months for nearly 5 per cent.[75] In 2012, 75 per cent of consecutive invitations were on time.

Also diagnostic delay can occur. A study in the south of the Netherlands found that, in 6.5 per cent of the 1,503 screening carcinoma cases, final diagnosis took more than three months after the screening result became available.[260] The percentage was found to differ considerably from one hospital to another, the range being 4 to 11 per cent.

It is reasonable to assume that treatment delays can adversely affect the prognosis. Each month's delay equates to three to six months loss of life expectancy, according to calculations based on clinical data on the natural course of breast cancer.[261] It has been demonstrated that an excessive screening interval is associated with an elevated risk of advanced breast cancer.[95] That underlines the importance of preventing unnecessary delays in screening (and treatment).

In response to the major regional differences in actual screening intervals, steps have been taken to promote conformance to the standard. The Committee regards the practice of reporting the *average* interval uninformative. The

percentage of repeat examinations with an actual screening interval of at least twenty-seven months is more informative.

8.5 Compression paddles, film

Various studies have recently been started in the Netherlands, which should shortly provide information about possible ways of making mammography less painful, without adversely affecting mammogram quality or increasing the radiation dose. In 2011, a comparative study of the two types of compression paddle used in mammography was started.[262] The results indicated that use of a newer, flexible type of paddle should be discouraged, because with the new paddles less breast tissue was imaged than with the classic, non-flexible type, the glandular tissue was pushed backwards and the whole of the glandular tissue in the breast was less well imaged. Furthermore, the claim that women found flexible compression less uncomfortable could not be substantiated. A subanalysis in women with larger breasts obtained similar results.

In a second study, a new, pressure-controlled compression method – which involves adjusting compression to the size and firmness of the breast – was studied. The initial results are positive.[263] Compared with the standard power-driven compression method, the new pressure-controlled method was associated with considerably less pain, especially extreme pain (NRS score $\geq$7). The image quality of the mammogram remained high. However, extra training is required to use pressure-controlled compression properly.

A third study, involving the use of film on the compression paddle and the bucky to enable more breast tissue to be imaged, is still in progress.

8.6 Radiation exposure

Over time, the radiation dose from screening mammography has been reduced considerably. Still research is in progress, aimed to identify ways of achieving further reductions without adversely affecting test performance. A new technique known as spectral imaging or photon counting is under development, which is intended to drastically diminish scatter radiation. The signs are that the new technique may nearly halve the glandular dose.[264]

8.7 Informed decision-making

People are often inclined to overestimate relative chances (e.g. risk ratios and odds ratios) in comparison with absolute chances.[339]

Providing information about absolute chances in the information material is very important, in order to avoid presenting screening too favourable. Women are currently still told that participation in screening will halve their chances of dying of breast cancer, which in absolute terms means that a fifty-year-old woman can reduce her risk of dying from breast cancer before the age of eighty from 2.6 per cent to 1.3 per cent.

In a responsible screening programme, women must be given honest, understandable, and balanced information, to enable an informed decision whether or not to participate.[265] Informed decision-making is facilitated by explaining the benefits as well as the harms of screening, without causing a reduction in participation.[266] However, what constitutes 'informed'.

A group of experts, including patients' representatives, was surveyed to establish what every woman should know about breast cancer screening.[267] The screening programme information material has since been revised accordingly. Women who wish to have more than the basic essential information can find more detailed information on a special website (www.rivm.nl/Onderwerpen/B/Bevolkingsonderzoek_borstkanker, consulted 7 January 2014). Screening programme information material is regularly updated. In 2008 women of fifty generally seemed to have sufficient knowledge, but the study group was small and the response was only 50 per cent.[268] It needs to be established whether the provided information does indeed meet the requirements of support for informed decision-making.[269] There is also a need for a decision aid (translated into Dutch and focused on the situation in the Netherlands).

Chapter 9

Ways of improving screening in the longer term

The existing screening programme in the Netherlands offers all eligible women the same screening regime. The target group for screening is defined by a single breast cancer risk factor, albeit the most influential factor: age. However, women of a given age are not all at equal risk of developing breast cancer. Therefore, is it possible to accurately estimate individual risk and tailor screening accordingly, thus increasing its effectiveness and efficiency?

9.1 Risk stratification

Early attempts to restrict population based screening to women at increased risk of breast cancer (selective screening) were rejected as non-viable. Classic risk factors, such as childlessness and obesity proved insufficient for the definition of a narrower target group for selective screening without considerable loss of programme effectiveness.[7,270] The fact that breast cancer often occurs in women who, based on classical risk factors, should not be at increased risk, emphasizes how much remains unknown about breast cancer causation.

Nevertheless, it is possible that more refined risk stratification and tailored screening might increase the effectiveness and efficiency of the programme and improve the relationship between the benefits and harms.[271] Three studies are currently in progress with a view to exploring the selective screening option: the KARMA study in Stockholm, PROCAS in Manchester and PRISMA in the Netherlands (http://www.zonmw.nl/nl/projecten/project-detail/breast-cancer-

screening-from-one-size-fits-all-to-a-personalized-risk-based-approach/samenvatting/).[272,273]

By collating information about as many of each participant's risk factors as possible, the woman's individual risk is estimated. Computer models are then used to calculate the effect of intensively screening the high-risk group while providing less intensive screening for low-risk group.

Risk models

Researchers have been working for at least thirty-five years to develop risk factor-based models, with which to predict the development of a woman's individual (absolute) chance of developing breast cancer.[274,275] Well-known examples are the Gail model and the more comprehensive Eurocentric Tyrer-Cuzick model.[276,277] However, the modellers have not yet succeeded in accurately defining individual levels of risk. That is partly because various relatively important risk factors, such as obesity, weight gain and bone density, were usually ignored in the risk models.[278] Physical exercise and alcohol consumption, which are known moderate risk factors, were not included in any of the models. A further complication is that major differences in risk are required to justify differentiation in the screening provision.

Better estimates of individual risk could probably be made by using more comprehensive risk models. 'New' risk factors, such as the radiological density of the breast tissue and blood profiles based on sex hormones and genetic or other markers could add value to the modelling process.[279] Candidate markers and risk models would require validation in large cohort studies before they could responsibly be used outside the context of scientific research. The clinical utility of a model must be established before it is introduced.[280,281]

Density of breast tissue

A high tissue density means that the breast contains a lot of glandular and connective tissue relative to fatty tissue. Numerous studies have identified high tissue density as an important risk factor for the development of breast cancer, regardless of other risk factors. Moreover, as tissue density increases, the sensitivity of mammography as a means of detecting breast cancer declines.[282] Glandular and connective tissue attenuates X rays more than fatty tissue, obscuring any abnormalities present and consequently substantially increasing the chance of interval cancer.[283] Information about tissue density could be important for individualizing the screening method or screening interval.

According to a US study, 6 per cent of women between the ages of 40 and 70 have a breast density of 75 per cent or more.[283] They are four to five times as likely to develop breast cancer as the 27 per cent of women whose breast density is 10 per cent or less.[284] Cohort studies with a long follow-up period, in which density was determined at the time of recruitment, rather than at the time of diagnosis, confirm that density is a risk factor for breast cancer and breast cancer mortality.[167,282]

Since the appearance of the original Wolfe classification system for density, various qualitative and quantitative methods for determining the degree of density have been developed.[284] With digital screening mammography, the volume of dense tissue can possibly be determined more accurately, taking account of the compression of the breast and the projection angle. This volumetric method has been validated by means of MRI,[285] and may provide a basis for improved risk estimation, but research is currently still in progress. So far, adding density as a risk parameter to a risk model did not sufficiently enable clear differentiation of breast cancer risk.[273,286,287]

Blood testing for genetic variants

Fundamentally breast cancer is determined by the disturbance of biological processes in the epithelial cells of the mammary gland, resulting mainly from (usually non-hereditary) changes in the function of oncogenes and tumour suppressor genes. The disturbance is accompanied by changes in composition or quantity of molecules such as DNA, RNA and protein. Those molecules – here referred to as biomarkers – can be measured in the laboratory in tumour tissue, blood or saliva. Research in this field is concentrating on the development and large-scale validation of biomarkers with improved test characteristics, and on the optimization of test methods. Additionally methods are being developed to make tumour markers visible using imaging techniques, such as PET and MRI.

Biomarkers in the DNA of the germ line (gametes) have a particular significance. These markers determine hereditary disposition to breast cancer: not only ‘hereditary breast cancer’ caused by mutations in the BRCA1 or BRCA2 gene. In the general population, there is also variation in hereditary disposition. Variations in the coding sequence within certain genes or in the number of copies of a given gene per cell, are indicative of a woman’s chance of developing breast cancer. Germ line biomarkers could in principle be used to refine selection of the target group for screening. However, individual germ line biomarkers have little influence on the probability of breast cancer; only in

combination they define substantial risk. In part, this type of research is limited by the enormous inter-individual variation in the genome.

Increased opportunity for large-scale DNA analysis has opened the way for genome-wide association studies. In such studies, large groups of patients and controls are tested for determining differences in the presence of genetic factors at the SNP (single nucleotide polymorphism) level by simultaneously taking thousands of samples from the genome as a whole. An SNP is a variation in a nucleotide: a single-letter variation in the four bases (known as A, C, G, T) making up the DNA. The term 'polymorphism' indicates that, in a (large) human population, there sometimes (e.g. at least 5 per cent of the time) is one nucleotide at a given location and sometimes another. Millions of SNPs have been identified.

SNPs associated with breast cancer are being discovered all the time.[288] Each SNP confers only a slight increase or decrease in breast cancer risk. It is anticipated that, with combinations of SNPs the risk estimation may be improved. In a meta-analyses including 96 SNPs,[289] 41 proved to be statistically significantly associated with breast cancer risk. In a modelling study the combination of these 41 SNPs predicted breast cancer as reliably as the existing risk models. Another study demonstrated that, with a hypothetical screening strategy based on 67 genetic variants, 24 per cent fewer women would require screening than in the UK's current programme, but 14 per cent fewer cases of breast cancer would be detected.[290]

Various gene-environment interactions involving SNPs have already been defined.[291] A number of studies have shown that there is little interaction between SNPs and standard risk factors and that the addition of SNPs could modestly improve the predictive capability of risk models.[292] Contrary to expectations, using SNPs, on their own or in combination with density, has proved to be of little value.[273]

There are some known mutations that constitute an intermediate risk factor, between SNPs and BRCA1/2, which are fairly good risk predictors, although not nearly as good as BRCA1/2. One example is the 1100delC mutation in the tumour suppressor gene CHEK2, which has an odds ratio of 2.7 for heterozygote women and possibly 3.4 for homozygote women.[293] First-degree relatives of carriers have a similarly elevated risk of breast cancer.[294]

Blood testing for hormone levels

High concentrations of oestrogens or androgens in the blood are risk factors for breast cancer in postmenopausal women.[295,296] However, sex hormone

concentrations are also linked to certain other risk factors,[297,298] albeit not to density as it seems.[299] The added predictive value of hormone levels has yet to be investigated further.

9.2 Scope for long-term improvement of the screening programme by imaging

Making additional screening available to women whose screening mammograms exhibit dense breast tissue would appear to be an attractive form of risk stratification. The presence of dense breast tissue makes it harder to detect breast cancer by means of either analogue or digital mammography.[300] More frequent mammographic screening of such women may achieve little,[283,301] but using a more sensitive screening method could be advantageous.

One drawback of mammography is that it produces a two-dimensional image. Two-view mammography is now the norm, to avoid the superprojection of structures. However, it is only a partial solution. The use of various high-performance imaging techniques, such as tomosynthesis, MRI and ultrasonography, is therefore being investigated.

9.2.1 Additional screening by means of tomosynthesis

Tomosynthesis is a technique that involves the use of a mammography device modified to enable the tube to move in an arc and thus to make a series of low-dose projections. The technique was previously known as planigraphy or tomography. Algorithms are then used to reconstruct slices from the data, enables more or less three-dimensional radiographs, supplementing conventional two-dimensional mammograms.

As an advantage relevant structures are less likely to be hidden by superprojection. Undoubtedly as important, referral due to pseudotumours can also be avoided. Various manufacturers' devices are currently available on the European market, which can differ markedly in their design and image acquisition.

Research has shown that the use of tomosynthesis in combination with standard mammography (2D+3D) results in breast cancer being detected more often than when standard mammography (2D) is used on its own and also results in lower referral rates in the study areas (Oslo and Italy).[302-304] Currently, the drawbacks are: greater radiation exposure, increased duration of examination and therefore pain, and increased reading time for radiologists.

The extent to which those drawbacks can be avoided depends on the development of so-called '2Dsynthetic+3D tomosynthesis'. With that technique, the 2D-images are not acquired separately but are reconstructed from the data acquired for tomosynthesis. The reconstructed images currently lack somewhat in quality, but improvements are anticipated in the short term.[305] A new version is expected to become available shortly, matching the performance of 2D+3D.

Yet there is insufficient evidence to support the general adoption of additional screening by means of tomosynthesis.[306] It has not been demonstrated that the higher detection rate actually reduces the chance of interval cancer, or that a lower referral rate could be expected in the Netherlands, where the referral rate is already low. Nor is it known which setup will yield the most efficient cost-effectiveness ratio. Furthermore, no solutions are yet available for issues such as the need to process huge quantities of data, the need for independent quality control, and crucially for the key aspects of mammographic screening: comparison with previous results.

9.2.2 *Additional screening by means of MRI*

Magnetic resonance imaging (MRI) is a more sensitive technique than mammography, but yields false positive results more often. MRI images tissue by a different method, without involving ionizing radiation. Additional MRI screening has been shown to be beneficial for women with $BRCA_1$ or $BRCA_2$ mutations, but not (yet) for women at less risk of breast cancer. For example the benefits – in terms of reduced breast cancer mortality – and the efficacy of additional MRI screening for women with high breast density are unclear.

In 2012, a randomized trial was started in the Netherlands to investigate the value of additional MRI for screening participants whose conventional mammograms are negative, but who exhibit high breast density (75 per cent or more).[307] This so-called DENSE Trial aims to establish the effectiveness in terms of reduction of interval cancers. This outcome parameter is considered the surrogate for breast cancer mortality. Secondary concerns are the participation rate, cost-effectiveness and effect on quality of life. Because three screening rounds are required, the duration of the DENSE Trial will be quite considerable: a total of eight years.

9.2.3 *Additional screening by means of ultrasonography*

In the United States, 2,662 women with dense glandular tissue and at least one other risk factor were studied to establish how often breast cancer was detected

in three annual rounds of combined mammographic and ultrasonographic screening.[308] A subgroup additionally underwent MRI screening (with contrast). Within a year of screening, breast cancer had been detected in 111 of the women: 59 by mammography, 91 by mammography combined with ultrasonography. Screening sensitivity was increased from 53 to 76 per cent by the combination of mammography and ultrasonography, and to 100 per cent by the combination of all three techniques. The number of false positive results also rose, however: the combination with ultrasonography raised the referral rate from 11 to 27 per cent in the first round of screening and from 9 to 17 per cent in subsequent rounds. Referral often led to invasive diagnostic testing: the percentage of participants who underwent biopsy increased from 2 to 10 per cent in the first round and from 2 to 7 per cent in subsequent rounds. Of the women who also received MRI screening, 36 per cent were referred and 13 per cent underwent biopsy. The researchers concluded that combination of ultrasonography or MRI with mammography leads to breast cancer being detected more frequently in women who are at elevated risk, but at the cost of additional false positives. Furthermore, ultrasonography is time-consuming. Ultrasonography of both breasts, followed by assessment, requires about nineteen minutes.[309]

Improved detection is of little benefit, if these cases mainly concern dormant tumours that, without additional screening, would not lead to symptoms or mortality.[307] More is not necessarily better.[310] With a non-randomized study of the kind described,[308] or indeed from a retrospective study of the kind more recently carried out,[311] it is impossible to ascertain whether the ultimate aim of reduced breast cancer mortality is realized by such a combination of techniques. As the beneficial effects are not established, it is also impossible to ascertain whether any benefits outweigh the harms (the additional false positives and biopsies). Consequently, the use of ultrasonography in addition to the established screening method is justified only in the context of well-designed scientific research.

9.2.4 *ABUS*

The use of 3D ultrasonography (ABUS, automated breast ultrasound) to supplement mammography in women with dense breast tissue is considered very promising.[312] It should match the high sensitivity of manual breast cancer ultrasonography, without the poor reproducibility and standardization issues associated with the conventional technique. ABUS is also less time-consuming, requiring only five to seven minutes per breast. However, little research has yet been done with ABUS in the general population.[313] Under British supervision, a

screening trial is in progress, in which digital mammography plus ABUS is being compared with digital mammography only (using a matched-pair design).[312] The trial will involve between 22,000 and 25,000 women with high breast density (>50 per cent).

9.2.5 *Computer-aided detection*

Various forms of computer-aided detection (CAD) are available. CAD is potentially useful, first as a way of preventing relevant abnormalities from being overlooked (perception errors). And secondly, forms of CAD are under development that support the interpretation of observed abnormalities (analysis), for example by presenting information about the chance of disease being present or by displaying comparable abnormalities with known diagnoses. Currently CAD is not used in the screening programme of the Netherlands.

Although numerous CAD studies have been conducted in the United States, the results are hardly relevant for the effectiveness of CAD in the Netherlands.[314] The reason is that CAD is used in the United States exclusively to prevent perception errors: the radiologist first makes a decision without CAD, but may still refer the case if the CAD indicates an abnormality. However, the radiologist may not reconsider a referral decision if the CAD detects no abnormality. That approach inevitably reduces the specificity of CAD and increases the sensitivity of the screening.

Various studies have shown that the use of CAD to support the assessment of mammograms by a radiologist (single reading) increases the sensitivity of the process. However, almost no research has been conducted into the value of CAD in comparison with the practice of double reading by two mutually independent radiologists, which is the norm in the Netherlands.

In a randomized trial involving more than 30,000 women in the UK, double reading was compared with assessment by a radiologist supported by CAD.[315] CAD was not found to increase the (relative) sensitivity, despite a slightly higher referral rate. An economic analysis linked to the trial indicated that CAD was unlikely to make screening in the UK more cost-effective unless the performance was improved, for example by reducing the referral rate.[316] A systematic review concluded that there was insufficient evidence that CAD at least matched the performance of double reading by two mutually independent radiologists.[317]

As a result of digital screening mammography, CAD programmes are getting better all the time. Interactive CAD looks particularly promising, but has yet to be tested in a clinical setting.[318] Whether CAD also improves the screening method, has still to be demonstrated by formal evaluation of its efficacy.[319]

9.2.6 *Other imaging techniques*

New techniques regularly emerge, which may have advantages over mammography, such as digital infrared thermal imaging (DITI), optical mammography, breast-specific gamma imaging (BSGI), positron emission mammography (PEM), elastography, computerized breast imaging (CBI), electrical impedance scanning (EIS) and photoacoustic mammoscopy (PAM).[320,321] Numerous commercial products can also be found on the internet. Survey research found that respondents were persuaded by the advertising claims made for DITI, EIS and CBI and regarded them as attractive alternatives to mammography.[322] However, most of the new techniques are more likely to compete with breast MRI than to serve as an alternative to screening mammography.

A synoptic study of sixty publications led to the conclusion that there was insufficient evidence for the efficacy of these generally immature techniques when used as independent or supplementary screening methods.[323] Moreover, with some of the techniques (BSGI, PEM) the radiation exposure is twenty to thirty times higher than with digital mammography.[324]

The Committee concludes that none of the emerging techniques referred to appear to have the potential to match the performance of screening mammography.

9.3 Preventing overdiagnosis and overtreatment where possible

Strictly speaking, diagnosis and treatment are outside the domain of population base screening, but they are nevertheless considered here because they influence the efficacy of screening. It is important to avoid merely increasing overdiagnosis by exclusively focussing on increasing the sensitivity of screening.[325]

9.3.1 *Can overdiagnosis be controlled?*

Breast cancer and, in particular, DCIS are detected more often using digital mammography than using analogue screening (Annex E, Table 2). Little research has so far been conducted into the nature of the additional cases. Nevertheless, a study in the south-east of the Netherlands has suggested that most of the additional cases involve low-grade DCIS.[326] That may point to overdiagnosis, particularly since it has been shown that digital screening does not reduce the

number of interval cancers.[327] A larger, national study has demonstrated that the additional cases detected by digital screening included not only DCIS, but also small invasive breast cancers.[238] Another study found no evidence that digital screening leads to the identification of more low-grade DCIS.[238] Further research is required to clarify the impact of digital mammography. In that context, it would be useful to consider whether there is scope for raising the detection threshold.

Other possible harms of digital mammography screening, such as more false positive results and more diagnostic procedures,[328-330] have also not been adequately quantified. Further research is recommended.[331]

With a view to limiting overtreatment, randomized trials are being prepared in the UK and the Netherlands, which will compare conventional therapy with a conservative 'wait-and-see' policy following the diagnosis of DCIS with grade 1 malignancy.[332]

Screening carcinomas have a more favourable prognosis than breast cancers with the same characteristics detected as a result of the investigation of symptoms.[140] Research into the causes of that discrepancy could reduce overtreatment and is a pressing need. Also research to validate molecular tests to distinguish between biologically aggressive and non-aggressive breast cancers more reliably than with classic prognostic characteristics alone, is recommended.[325]

9.3.2 *Is radiotherapy always needed and should the whole breast always be irradiated?*

Radiotherapy is an important element of breast cancer treatment. Because the risk of cancer recurrence after breast-conserving surgery depends on the patient's age, research is performed to establish whether postsurgical radiotherapy is necessary for older patients with early-stage breast cancers.[333] It is clear that, for many patients, radiotherapy (or adjuvant chemotherapy) is not necessary. Unfortunately, the need for radiotherapy cannot be confidently predicted in individual cases, due to a lack of validated selection criteria.

Hypofractionation (the use of a shorter, more intensive radiotherapy schedule) reduces the treatment duration and the physical strain for many patients ($pT_{1-3a}N_{0-1}M_0$), as well as simplifying hospital logistics. Conventional postsurgical radiotherapy (25 fractions of 2 Gy) and shorter schedules (13 to 16 fractions) have been investigated in various trials. After a ten-year follow-up, no difference was detected in terms of breast cancer recurrence rate, survival, cosmetic outcome or side-effects.[334] Despite the greater dose per fraction in the

hypofractionation schedule, the cardiac dose was no greater than with the twenty-five-fraction schedule. The traditional duration of conventional radiotherapy following breast-conserving surgery (six to seven weeks) has now been reduced to three to four weeks in nearly all centres in the Netherlands. The boost schedule has also been revised, from eight to five radiotherapy sessions in a sixteen-fraction schedule.

In the FAST-FORWARD trial, the fifteen-fraction, three-week schedule used in the United Kingdom is being compared with a five-fraction, five-working-day schedule delivered in a single week (www.cancerresearchuk.org). The trial is set to run until spring 2016.

Another development that could reduce the harms of radiotherapy without limiting the benefits, is intraoperative radiotherapy (IORT). IORT is an experimental radiotherapy technique for patients with early-stage breast cancer eligible for breast-conserving surgery. After removal of the tumour, the tumour bed is irradiated in the body during the surgical procedure. If this one-off procedure proves to be effective, conventional external postsurgical radiotherapy lasting three to four weeks is not required. With IORT, there is almost no exposure of adjacent breast tissue, or of the heart, lungs or oesophagus to radiation. The provisional results are hopeful in terms of the prevention of metastases in other organs, survival and side-effects such as pain and dermal induration.[335,336] However, after a follow-up of more than five years, IORT appeared to be less effective in preventing recurrence of breast cancer to the affected breast than the standard procedure. Further research is required to determine which subgroups can benefit from IORT.

Chapter 10

Responses to the request for advice

10.1 How effective is the existing screening programme in the Netherlands?

In 1989, the Netherlands started introducing population based screening for breast cancer. The expectation was that, in due course, by around 2015, the screening programme would result in about 700 fewer breast cancer deaths per year than there would be without screening. That expectation was based on the results of screening trials in other countries and population based screening trials in Utrecht and Nijmegen in the Netherlands.

Currently the population based screening programme in the Netherlands is characterized by a high participation rate, a low referral rate and a high positive predictive value. The Committee concludes that population based screening for breast cancer continues to have a beneficial effect on mortality consistent with the original expectations, even though circumstances have changed over time, particularly in terms of the availability of improved treatment for breast cancer.

In the Netherlands between 1989 and 2011, breast cancer mortality decreased from 39 to 27 per 100,000 women (of all ages). To what extent that 31 per cent decrease may be attributed to population based screening is difficult to determine. With computer modelling, it has been estimated that about half (16 per cent) of the mortality decrease is due to screening and the rest to improved therapy. This constitutes to on average 775 breast cancer deaths prevented per year; i.e. consistent with the original expectation.

How great are the harms in the Netherlands?

Screening results in earlier diagnosis. In screening a woman diagnosed with breast cancer will learn she has the disease several years sooner than would otherwise have been the case. In three cases out of four, there is no compensatory benefit in terms of life-years gained, although there may be benefit in terms of less extensive therapy. Quality of life may also be adversely affected by anxiety and uncertainty associated with an abnormal screening result and the subsequent diagnostic testing, even if no evidence of breast cancer is ultimately found. Per examination, the chance of a false positive test result is relatively small in the Netherlands (1.7 per cent of participants in 2012, compared with more than 10 per cent in the US, for example), but can rise to more than 15 per cent for a woman who undergoes screening every two years for the entire 26 years that she is within the target age group.

For participants, the most significant drawback associated with breast cancer screening is overdiagnosis. The term implies the detection of a tumour that would never have become symptomatic during the patient's life without screening. In the Netherlands, just over 8 per cent of breast cancer cases detected by screening constitute overdiagnosis.

The Committee considers it unlikely that, in the long term, treatment for breast cancer leads to excess mortality due to other causes than breast cancer. As for example cardiovascular damage, induced by radiotherapy or chemotherapy does not become manifest for a long time, long-term studies remain desirable.

10.2 How do the benefits relate to the harms?

In the population based screening programme in the Netherlands, 1,200 women must undergo mammography to prevent one breast cancer death. When screening prevents the death of a woman from breast cancer, she is spared the terminal phase of the disease and on average gains 16.5 life-years. A cost-effectiveness analysis indicates that the cost per life-year gained is 1,600 euros. The existing population screening programme prevents an average of 775 breast cancer deaths per year. That is two to three times the estimated number of women affected by overdiagnosis.

The Committee concludes that, because of screening, fewer women die from breast cancer. The decrease in mortality relevantly outweighs the harms, which consist mainly of overdiagnosis and false positive screening results.

10.3 How can screening be improved in the short term?

According to the Committee, there is no scientific reason to adjust the limits of the target age group, the screening interval or the screening method. The Committee does not recommend further research into the effectiveness of screening women under the age of fifty, regardless of their breast cancer risk. It is, however, recommended to investigate the feasibility and desirability of selective screening of women under fifty who are at elevated risk of breast cancer, in the context of a risk stratification study.

Following the lowering of the referral threshold and the introduction of digital mammography in screening, the referral rate increased and more cases of breast cancer are now detected. However, the number of false positive referrals also rose substantially (Annex E, Table 2). That phenomenon requires further research. To reduce the adverse implications of referral to a breast clinic, both for the women concerned and for the health care system, an alternative referral procedure is being investigated for women whose screening result indicates a low suspicion of breast cancer (BI-RADS 0). If the results are favourable, introduction of the optimal referral strategy can considerably reduce the problem of false positive referrals.

Research has found that women who receive false positive screening results often return for outpatient check-ups after receiving the diagnosis ‘benign breast abnormality’. The Committee believes that further research is advisable, to improve the support of such women.

Various studies are in progress in the Netherlands, which in the near future are likely to point out how mammography can be made less painful, without adversely influencing the quality of the mammogram or the radiation dose.

10.4 How can screening be improved in the medium term?

10.4.1 Risk stratification

The Dutch existing population based screening programme offers the same screening regime to all women in the target group, defined on the basis of one risk factor: age. Adjusting screening in line with the individual’s estimated breast cancer risk appears attractive as a means of increasing the effectiveness and efficiency of the screening. However, with existing risk models, it has not yet proved possible to reliably categorize women according to individual risk. Risk estimation can probably be improved by using more comprehensive models. The

inclusion of risk factors such as breast tissue density and blood profiles based on sex hormones and genetic or other markers could add particular value to the modelling process. Numerous studies are in progress with the aim of assessing candidate markers and the validity of new models. Questions also need to be answered regarding the logistics of risk stratification within the population based screening programme and regarding the acceptability and effects of offering intensive screening (lower intake age, additional forms of screening) to the high-risk group and less intensive screening to the low-risk group.

10.4.2 Tomosynthesis

Tomosynthesis is a new technique, which involves supplementing the conventional two-dimensional mammogram with something similar to three-dimensional images of the breast. Tomosynthesis has considerable promise as a means of further improving screening test performance. Currently, the harms are greater radiation exposure, increased duration of examination and therefore pain, and increased reading time for radiologists. Several significant problems remain to be resolved before it would be appropriate to organize a screening trial.

Chapter 11

Recommendations

11.1 Policy recommendations

1 The Committee concludes that population based screening programme in the Netherlands provides considerable health benefit. Accordingly the programme should be continued and improved further.
2 There is insufficient evidence to support extending the age limits below 50 or above 75 (in the existing screening programme based on age, regardless of individual breast cancer risk; Subsections 8.2 and 8.3).
3 Regional differences, such as in the actual screening interval (Subsection 8.4), should be further reduced. Up-to-date information about the results, as contained in the annual evaluation reports, can support that aim.
4 Digitization of the screening programme can further improve the quality assurance, monitoring and evaluation, for example by expediting the

availability of information. In that context, the development of a review method that excludes hindsight bias is desirable (Subsection 6.5.3).
5 BI-RADS scores should be included in the evaluation reports – to improve monitoring.
6 The provision of balanced information materials and support for informed decision-making regarding screening participation requires ongoing attention (Subsection 8.7).

11.2 Research recommendations

1 Investigate why women who have received false positive screening results so often and so long continue to receive outpatient supervision, without being referred back to the screening programme (Subsection 6.5.1).
2 Investigate the feasibility and desirability of risk stratification (Subsection 9.1). Include women under fifty with elevated risk of breast cancer compared to contemporaries (Subsection 8.2).
3 Investigate the extent of and reasons for opportunistic screening (Subsection 2.2).
4 Investigate the optimal tumour detection threshold (Subsection 9.3). Has the higher breast cancer detection rate due to digital screening reduced the risk of interval cancer?
5 Investigate why the prognosis for screening carcinomas is intrinsically better than the prognosis for breast cancer with the same characteristics detected outside the context of screening (Subsection 9.3). The findings may facilitate the reduction of overtreatment.
6 Investigate prognostic characteristics which could be used to establish whether postoperative radiotherapy is necessary.

References

1 Health Council of the Netherlands. Screening: between hope and hype. The Hague: Health Council of the Netherlands, 2008; publication no. 2008/05E. Internet: www.gr.nl.

2 Geuns HA van. Waarde bevolkingsonderzoek op tuberculose. Medisch Contact 1983; 29: 891-893.

3 Baum M. Harms from breast cancer screening outweigh benefits if death caused by treatment is included. BMJ 2013; 346: f385.

4 Baum M, Thornton H, Götzsche PC, Bewley S, Jörgensen KJ, Barratt A et al. Breast cancer awareness month. Still awaiting screening facts. BMJ 2010; 341: c6152.

5 Götzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2011;(1): CD001877.

6 Bonneux LG, Autier P. Bevolkingsonderzoek naar borstkanker loont niet: screening heeft nauwelijks invloed op sterfte. Ned Tijdschr Geneeskd 2011; 155(35): A3774.

7 Gezondheidsraad. De vroege opsporing van borstkanker. Den Haag: Gezondheidsraad; 1987: 1987/09. Internet: www.gr.nl.

8 Health Council of the Netherlands: The benefit of population screening for breast cancer with mammography. The Hague: Health Council of the Netherlands, 2002; publication no. 2002/03E. Internet: www.gr.nl.

9 Health Council of the Netherlands. Annual report on screening for disease 2006. The Hague: Health Council of the Netherlands, 2006; publication no. 2006/10. Internet: www.gr.nl.

10 Gelder R de, Heijnsdijk EAM, Fracheboud J, Draisma G, de Koning HJ. The effects of population-based mammography screening starting between age 40 and 50 compared to the effects of adjuvant systematic therapy. In: de Gelder R, editor. Predicting the benefits and harms of breast cancer screening. Rotterdam: Thesis, Erasmus University; 2012: 113-132.

11 Landelijk Evaluatie Team voor bevolkingsonderzoek naar Borstkanker. LETB rapportage 2012. Belangrijkste resultaten 2010 bevolkingsonderzoek borstkanker. Rotterdam: Erasmus MC, afd. Maatschappelijke Gezondheidszorg, Radboud Universiteit Nijmegen, afdeling Epidemiologie, Biostatistiek en HTA.; 2012.

12 Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. Ann Intern Med 2007; 146(7): 502-510.

13 Fletcher SW, Elmore JG. False-positive mammograms--can the USA learn from Europe? Lancet 2005; 365(9453): 7-8.

14 Hofvind S, Thoresen S, Tretti S. The cumulative risk of a false-positive recall in the Norwegian breast cancer screening program. Cancer 2004; 101(7): 1501-1507.

15 Njor SH, Olsen AH, Schwartz W, Vejborg I, Lynge E. Predicting the risk of a false-positive test for women following a mammography screening programme. J Med Screen 2007; 14(2): 94-97.

16 Puliti D, Miccinesi G, Zappa M. More on screening mammography. N Engl J Med 2011; 364(3): 284-285.

17 Salas D, Ibanez J, Roman R, Cuevas D, Sala M, Ascunce N et al. Effect of start age of breast cancer screening mammography on the risk of false-positive results. Prev Med 2011; 53(1-2): 76-81.

18 Landelijk bevolkingsonderzoek naar borstkanker in Nederland. Optimalisatiestudie 1999-2001. Rotterdam/ Nijmegen: Landelijk Evaluatie Team voor bevolkingsonderzoek naar Borstkanker (LETB) en Landelijk Referentiecentrum voor bevolkingsonderzoek op Borstkanker (LRCB); 2002.

19 Otten JD, Karssemeijer N, Hendriks JH, Groenewoud JH, Fracheboud J, Verbeek AL et al. 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.

20 Timmers JM, van Doorne-Nagtegaal HJ, Zonderland HM, van Tinteren H, Visser O, Verbeek AL et al. 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(8): 1717-1723.

21 Timmers J, den Heeten G, Zonderland H, Verbeek A, Broeders M. Veranderingen in bevolkingsonderzoek borstkanker. Huisarts Wet 2012; 55(7): 296-300.

22 Louwman WJ, Lemmens VE, van de Poll-Franse LV, van de Sangen MJ, Nieuwenhuizen GAP, Voogd AC. Mammacarcinoom in onze IKZ-regio; Factsheet op basis van gegevens uit de Kankerregistratie van het IKZ, 2010 en 2011. Eindhoven: IKZ; 2013.

23 LETB. Landelijke evaluatie van bevolkingsonderzoek naar borstkanker in Nederland 1990-2012. LETB XIII. Rotterdam/Nijmegen: Erasmus MC/UMC St Radboud (ter perse); 2014.

24 Council of the European Union (16 December 2003). Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC). OJ L 2004; 327: 34-38.

25 Karsa L von, Anttila A, Ronco G, Ponti A, Malila N, Arbyn M et al. Cancer screening in the European Union. Report on the implementation of the Council Recommendation on cancer screening - First Report. Lyon: International Agency for Research on Cancer; 2008.

26 Lynge E, Tornberg S, von Karss L, Segnan N, van Delden JJ. Determinants of successful implementation of population-based cancer screening programmes. Eur J Cancer 2012; 48(5): 743-748.

27 Giordano L, Von Karsa L, Tomatis M, Majek O, de Wolf C, Lancucki L et al. Mammographic screening programmes in Europe: organization, coverage and participation. J Med Screen 2012; 19 Suppl 1: 72-82.

28 Broeders MJ, Scharpantgen A, Ascunce N, Gairard B, Olsen AH, Mantellini P et al. Comparison of early performance indicators for screening projects within the European Breast Cancer Network: 1989-2000. Eur J Cancer Prev 2005; 14(2): 107-116.

29 Fabri V, Remacle A, Boutsen M. Programma Borstkankerscreening. Vergelijking van de eerste drie rondes 2002-2003, 2004-2005 en 2006-2007. Rapport nr. 7 van het Intermutualistisch Agentschap. Brussel: 2010. Internet: www.nic-ima.be.

30 Nunen K van, van der Burg M, Teughels S, van Hal G. Preventiebonus borstkankerscreening bij huisartsen en gynaecologen uit de regio Zuiderkempen. Universiteit Antwerpen; 2011.

31 Majek O, Danes J, Skovajsova M, Bartonkova H, Buresova L, Klimes D et al. Breast cancer screening in the Czech Republic: time trends in performance indicators during the first seven years of the organised programme. BMC Public Health 2011; 11: 288.

32 Boncz I, Sebestyen A, Pinter I, Battyany I, Ember I. The effect of an organized, nationwide breast cancer screening programme on non-organized mammography activities. J Med Screen 2008; 15(1): 14-17.

33 Gelder R de , Bulliard JL, de Wolf C, Fracheboud J, Draisma G, Schopper D et al. Cost-effectiveness of opportunistic versus organised mammography screening in Switzerland. Eur J Cancer 2009; 45(1): 127-138.

34 Autier P, Ait OD. Determinants of the number of mammography units in 31 countries with significant mammography screening. Br J Cancer 2008; 99(7): 1185-1190.

35 Smith-Bindman R, Ballard-Barbash R, Miglioretti DL, Patnick J, Kerlikowske K. Comparing the performance of mammography screening in the USA and the UK. J Med Screen 2005; 12(1): 50-54.

36 Palencia L, Espelt A, Rodriguez-Sanz M, Puigpinos R, Pons-Vigues M, Pasarin MI et al. Socio-economic inequalities in breast and cervical cancer screening practices in Europe: influence of the type of screening program. Int J Epidemiol 2010; 39(3): 757-765.

37 Bihrmann K, Jensen A, Olsen AH, Njor S, Schwartz W, Vejborg I et al. Performance of systematic and non-systematic ('opportunistic') screening mammography: a comparative study from Denmark. J Med Screen 2008; 15(1): 23-26.

38 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.

39 Hofvind S, Geller BM, Skelly J, Vacek PM. Sensitivity and specificity of mammographic screening as practised in Vermont and Norway. Br J Radiol 2012; 85(1020): e1226-e1232.

40 Koning HJ de, van Ineveld BM, van Oortmarssen GJ, Boer R, Collette HJA, Verbeek AL. De kosten en effecten van bevolkingsonderzoek naar borstkanker. Rotterdam: Instituut Maatschappelijke Gezondheidszorg Erasmus Universiteit Rotterdam. Instituut Sociale Geneeskunde Katholieke Universiteit Nijmegen. Prevention, Vakgroep Gezondheidszorg en Epidemiologie Rijksuniversiteit Utrecht; 1990.

41 Gelder R de. Predicting the benefits and harms of breast cancer screening: current debates and future directions. Table 8.2 p170 [Proefschrift]. Rotterdam: Erasmus Universiteit.

42 Götzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? Lancet 2000; 355(9198): 129-134.

43 Duffy S, Tabár L, Smith RA. The mammographic screening trials: commentary on the recent work by Olsen and Götzsche. J Surg Oncol 2002; 81(4): 159-162.

44 Duffy SW. Interpretation of the breast screening trials: a commentary on the recent paper by Götzsche and Olsen. Breast 2001; 10(3): 209-212.

45 Miller AB. Screening for breast cancer with mammography. Lancet 2001; 358(9299): 2164-2168.

46 Holmberg L, Duffy SW, Yen AM, Tabár L, Vitak B, Nyström L et al. Differences in endpoints between the Swedish W-E (two county) trial of mammographic screening and the Swedish overview: methodological consequences. J Med Screen 2009; 16(2): 73-80.

47 Freedman DA, Petitti DB, Robins JM. On the efficacy of screening for breast cancer. Int J Epidemiol 2004; 33(1): 43-55.

48 Veen WA van, Knottnerus JA. Het nut van bevolkingsonderzoek naar borstkanker; een advies van de Gezondheidsraad. Ned Tijdschr Geneeskd 2002; 146(22): 1023-1026.

49 Glasziou P, Houssami N. The evidence base for breast cancer screening. Prev Med 2011; 53(3): 100-102.

50 The Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. London: A report jointly commissioned by Cancer Research UK and the Department of Health (England); 2012.

51 Nyström L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Ryden S et al. Breast cancer screening with mammography: overview of Swedish randomised trials. Lancet 1993; 341(8851): 973-978.

52 Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet 2002; 359(9310): 909-919.

53 Héry C, Ferlay J, Boniol M, Autier P. Changes in breast cancer incidence and mortality in middle-aged and elderly women in 28 countries with Caucasian majority populations. Ann Oncol 2008; 19(5): 1009-1018.

54 Tryggvadottir L, Sigvaldason H, Olafsdottir GH, Jonasson JG, Jonsson T, Tulinius H et al. Population-based study of changing breast cancer risk in Icelandic BRCA2 mutation carriers, 1920-2000. J Natl Cancer Inst 2006; 98(2): 116-122.

55 Beral V, Bull D, Doll R, Peto R, Reeves G. Breast cancer and abortion: collaborative reanalysis of data from 53 epidemiological studies, including 83,000 women with breast cancer from 16 countries. Lancet 2004; 363(9414): 1007-1016.

56 Huis M van. Cohortvruchtbaarheid van niet-westers allochtone vrouwen. Bevolkingstrends. Voorburg: Bevolkingstrends. CBS; 2013.

57 Nab H, Mulder P, Crommelin M, van der Heijden L, Coebergh J. Veranderingen in de incidentie van borstkanker per geboortencohort in Zuidoost-Nederland. Ned Tijdschr Geneeeskd 1995; 139(7): 332-335.

58 Soerjomataram I, Pukkala E, Brenner H, Coebergh JW. On the avoidability of breast cancer in industrialized societies: older mean age at first birth as an indicator of excess breast cancer risk. Breast Cancer Res Treat 2008; 111(2): 297-302.

59 Talma H, Schönbeck Y, van Dommelen P, Bakker B, van Buuren S, Hirasing RA. Trends in menarcheal age between 1955 and 2009 in the Netherlands. PLoS One 2013; 8(4): e60056.

60 Eaton SB, Pike MC, Short RV, Lee NC, Trussell J, Hatcher RA et al. Women's reproductive cancers in evolutionary context. Q Rev Biol 1994; 69(3): 353-367.

61 Huang Z, Hankinson SE, Colditz GA, Stampfer MJ, Hunter DJ, Manson JE et al. Dual effects of weight and weight gain on breast cancer risk. JAMA 1997; 278(17): 1407-1411.

62 Gaudet MM, Gapstur SM, Sun J, Diver WR, Hannan LM, Thun MJ. Active smoking and breast cancer risk: original cohort data and meta-Analysis. J Natl Cancer Inst 2013; Epub ahead of print.

63 Grundy A, Richardson H, Burstyn I, Lohrisch C, Sengupta SK, Lai AS et al. Increased risk of breast cancer associated with long-term shift work in Canada. Occup Environ Med 2013; 70(12): 831-838.

64 Gezondheidsraad. Nachtwerk en borstkanker: een oorzakelijk verband? Den Haag: Gezondheidsraad; 2006: publicatienr 2006/15. Internet: www.gr.nl.

65 Gezondheidsraad. Ethanol (ethyl alcohol). Den Haag: Gezondheidsraad; 2006: 2006/06OSH. Internet: www.gr.nl.

66 Parkin DM. Is the recent fall in incidence of post-menopausal breast cancer in UK related to changes in use of hormone replacement therapy? Eur J Cancer 2009; 45(9): 1649-1653.

67 Beral V. Breast cancer and hormone-replacement therapy in the Million Women Study. Lancet 2003; 362(9382): 419-427.

68 Chlebowski RT, Kuller LH, Prentice RL, Stefanick ML, Manson JE, Gass M et al. Breast cancer after use of estrogen plus progestin in postmenopausal women. N Engl J Med 2009; 360(6): 573-587.

69 Soerjomataram I, Coebergh JW, Louwman MW, Visser O, van Leeuwen FE. Does the decrease in hormone replacement therapy also affect breast cancer risk in the Netherlands? J Clin Oncol 2007; 25(31): 5038-5039.

70 Louwman WJ, Voogd AC, van Dijck JA, Nieuwenhuijzen GA, Ribot J, Pruijt JF et al. On the rising trends of incidence and prognosis for breast cancer patients diagnosed 1975-2004: a long-term population-based study in southeastern Netherlands. Cancer Causes Control 2008; 19(1): 97-106.

71 Paap E, Broeders MJ, van SG, Otten JD, Verbeek AL. Large increase in a Dutch woman's lifetime risk of developing breast cancer. Eur J Cancer 2008; 44(11): 1485-1487.

72 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(8): 1929-1934.

73 Broeders MJM, van Kalken PJF, Paap E, Ten Voorde M, Kroeze-Voogd EH, Dunning-van Faassen L. Verdiepende studie naar leeftijdsverlaging binnen het bevolkingsonderzoek borstkanker. Landelijk Referentiecentrum voor Bevolkingsonderzoek. Bevolkingsonderzoek oost, Business unit Borstkanker Enschede. Bevolkingsonderzoek Noord; 2012.

74 Voogd AC, Repelaer van Driel OJ, Roumen RMH, Crommelin MA, van Beek MW, Coebergh JWW. Changing attitudes towards breast-conserving treatment of early breast cancer in the Southeastern Netherlands: results of a survey among surgeons and a registry-based analysis of patterns of care. Eur J Surg Oncol 1997; 23: 134-138.

75 Landelijk Evaluatie Team voor bevolkingsonderzoek naar Borstkanker. Landelijke evaluatie van bevolkingsonderzoek naar borstkanker in Nederland 1990-2007. Rotterdam: Erasmus MC, afd. Maatschappelijke Gezondheidszorg; LETB XII, 2009.

76 Ernst MF, Voogd AC, Coebergh JW, Repelaer van Driel OJ, Roukema JA. Wel degelijk gunstig effect van borstkankerscreening op de trend in mammasparende chirurgie. Ned Tijdschr Geneeskd 2002; 146(36): 1712.

77 Struikmans H, Aarts MJ, Jobsen JJ, Koning CCE, Poortmans PMP, Louwman MWJ et al. Trends in toepassing van primaire radiotherapie voor kanker in Nederland. Ned Tijdschr Geneeskd 2012; 156(12): 508-513.

78 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(9472): 1687-1717.

79 Hurvitz SA, Hu Y, O'Brien N, Finn RS. Current approaches and future directions in the treatment of HER2-positive breast cancer. Cancer Treat Rev 2013; 39(3): 219-229.

80 Tabár L, Vitak B, Chen HH, Duffy SW, Yen MF, Chiang CF et al. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. Radiol Clin North Am 2000; 38(4): 625-651.

81 Tabár L, Vitak B, Chen TH, Yen AM, Cohen A, Tot T et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. Radiology 2011; 260(3): 658-663.

82 Tabár L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Grontoft O et al. Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. Lancet 1985; 1(8433): 829-832.

83 Autier P, Boniol M, La Vecchia C, Vatten L, Gavin A, Hery C et al. Disparities in breast cancer mortality trends between 30 European countries: retrospective trend analysis of WHO mortality database. BMJ 2010; 341: c3620.

84 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: 1-10.

85 Hery C, Ferlay J, Boniol M, Autier P. Quantification of changes in breast cancer incidence and mortality since 1990 in 35 countries with Caucasian-majority populations. Ann Oncol 2008; 19(6): 1187-1194.

86 Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. N Engl J Med 2012; 367(21): 1998-2005.

87 Jörgensen KJ, Zahl PH, Götzsche PC. Breast cancer mortality in organised mammography screening in Denmark: comparative study. BMJ 2010; 340: c1241.

88 Blanks RG, Moss SM, McGahan CE, Quinn MJ, Babb PJ. Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1980-8: comparison of observed with predicted mortality. BMJ 2000; 321: 665-669.

89 Duffy SW, Tabár L, Vitak B, Allgood PC, Chen TH, Yen AM et al. Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. J Med Screen 2010; 17(1): 25-30.

90 Pons-Vigues M, Puigpinos R, Cano-Serral G, Mari-Dell'Olmo M, Borrell C. Breast cancer mortality in Barcelona following implementation of a city breast cancer-screening program. Cancer Detect Prev 2008; 32(2): 162-167.

91 Ascunce EN, Moreno-Iribas C, Barcos UA, Ardanaz E, Ederra SM, Castilla J et al. Changes in breast cancer mortality in Navarre (Spain) after introduction of a screening programme. J Med Screen 2007; 14(1): 14-20.

92 Otto SJ, Fracheboud J, Looman CW, Broeders MJ, Boer R, Hendriks JH et al. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. Lancet 2003; 361(9367): 1411-1417.

93 Yen AM, Duffy SW, Chen TH, Chen LS, Chiu SY, Fann JC et al. Long-term incidence of breast cancer by trial arm in one county of the Swedish Two-County Trial of mammographic screening. Cancer 2012; 118(23): 5728-5732.

94 Fracheboud J, Otto SJ, van Dijck JA, Broeders MJ, Verbeek AL, de Koning HJ. Decreased rates of advanced breast cancer due to mammography screening in The Netherlands. Br J Cancer 2004; 91(5): 861-867.

95 Nederend J, Duijm LE, Voogd AC, Groeneveld JH, Jansen FH, Louwman MW. Trends in incidence and detection of advanced breast cancer at biennial screening mammography in the Netherlands: a population based study. Breast Cancer Research 2012; 14(1): R10.

96 Hofvind S, Sörum R, Thoresen S. Incidence and tumor characteristics of breast cancer diagnosed before and after implementation of a population-based screening-program. Acta Oncol 2008; 47(2): 225-231.

97 Hofvind S, Lee CI, Elmore JG. Stage-specific breast cancer incidence rates among participants and non-participants of a population-based mammographic screening program. Breast Cancer Res Treat 2012; 135(1): 291-299.
98 Swedish Organised Service Screening Evaluation Group. Effect of mammographic service screening on stage at presentation of breast cancers in Sweden. Cancer 2007; 109(11): 2205-2212.
99 Zorzi M, Puliti D, Vettorazzi M, De Lisi V, Falcini F, Federico M et al. Mastectomy rates are decreasing in the era of service screening: a population-based study in Italy (1997-2001). Br J Cancer 2006; 95(9): 1265-1268.
100 Foca F, Mancini S, Bucchi L, Puliti D, Zappa M, Naldoni C et al. Decreasing incidence of late-stage breast cancer after the introduction of organized mammography screening in Italy. Cancer 2013; 119(11): 2022-2028.
101 Sarkeala T, Heinavaara S, Anttila A. Breast cancer mortality with varying invitational policies in organised mammography. Br J Cancer 2008; 98(3): 641-645.
102 Kalager M, Zelen M, Langmark F, Adami HO. Effect of screening mammography on breast-cancer mortality in Norway. N Engl J Med 2010; 363(13): 1203-1210.
103 Lynge E, Braaten T, Njor SH, Olsen AH, Kumle M, Waaseth M et al. Mammography activity in Norway 1983 to 2008. Acta Oncol 2011; 50(7): 1062-1067.
104 Paci E, Giorgi D, Bianchi S, Vezzosi V, Zappa M, Crocetti E et al. Assessment of the early impact of the population-based breast cancer screening programme in Florence (Italy) using mortality and surrogate measures. Eur J Cancer 2002; 38(4): 568-573.
105 Hakama M, Pukkala E, Heikkila M, Kallio M. Effectiveness of the public health policy for breast cancer screening in Finland: population based cohort study. BMJ 1997; 314(7084): 864-867.
106 Olsen AH, Njor SH, Vejborg I, Schwartz W, Dalgaard P, Jensen M-B et al. Breast cancer mortality in Copenhagen after introduction of mammography screening: cohort study. BMJ 2005; 330: 220-222 (340:c1241), letters BMJ 2005; 330:1024-1025.
107 Götzsche PC, Thornton H, Jörgensen KJ. Reduction in mortality from breast cancer: presentation of benefits and harms needs to be balanced. BMJ 2005; 330(7498): 1024-1025.
108 Lynge E, Olsen AH. Presentation of benefits and harms needs to be balanced. BMJ 2005; 330: 1025.
109 Duffy SW, Tabár L, Chen H-H, Holmqvist M, Yen M-F. The impact of organized mammography service screening on breast carcinoma mortality in seven Swedish counties. A collaborative evaluation. Cancer 2002; 95(3): 458-469.
110 The Swedish organised service screening evaluation group. Reduction in breast cancer mortality from organized service screening with mammography: 1. Further confirmation with extended data. Cancer Epidemiol Biomarkers Prev 2006; 15(1): 45-51.
111 Tabár L, Vitak B, Chen HH, Yen MF, Duffy SW, Smith RA. Beyond randomized controlled trials: organized mammographic screening substantially reduces breast carcinoma mortality. Cancer 2001; 91(9): 1724-1731.
112 Anttila A, Koskela J, Hakama M. Programme sensitivity and effectiveness of mammography service screening in Helsinki, Finland. J Med Screen 2002; 9(4): 153-158.

113 Njor S, Nyström L, Moss S, Paci E, Broeders M, Segnan N et al. Breast cancer mortality in mammographic screening in Europe: a review of incidence-based mortality studies. J Med Screen 2012; 19(Suppl 1): 33-41.

114 Knol MJ. Weg met oddsratio's: risicoratio's in cohortonderzoek en gerandomiseerd gecontroleerd onderzoek. Ned Tijdschr Geneeeskd 2012; 156(28): 1136-1141.

115 Autier P, Boniol M. Mammography screening and breast cancer mortality. Cancer Epidemiol Biomarkers Prev 2012; 21(5): 869.

116 Clegg LX, Reichman ME, Miller BA, Hankey BF, Singh GK, Lin YD et al. Impact of socioeconomic status on cancer incidence and stage at diagnosis: selected findings from the surveillance, epidemiology, and end results: National Longitudinal Mortality Study. Cancer Causes Control 2009; 20(4): 417-435.

117 Dano H, Hansen KD, Jensen P, Petersen JH, Jacobsen R, Ewertz M et al. Fertility pattern does not explain social gradient in breast cancer in denmark. Int J Cancer 2004; 111(3): 451-456.

118 Aarts MJ, Voogd AC, Duijm LE, Coebergh JW, Louwman WJ. Socioeconomic inequalities in attending the mass screening for breast cancer in the south of the Netherlands--associations with stage at diagnosis and survival. Breast Cancer Res Treat 2011; 128(2): 517-525.

119 Bastiaannet E, Portielje JE, van de Velde C, de Craen AJ, van der Velde S, Kuppen PJ et al. Lack of survival gain for elderly women with breast cancer. Oncologist 2011; 16(4): 415-423.

120 Cronin KA, Weed DL, Connor RJ, Prorok PC. Case-control studies of cancer screening: theory and practice. J Natl Cancer Inst 1998; 90(7): 498-504.

121 Roder D, Houssami N, Farshid G, Gill G, Luke C, Downey P et al. Population screening and intensity of screening are associated with reduced breast cancer mortality: evidence of efficacy of mammography screening in Australia. Breast Cancer Res Treat 2008; 108(3): 409-416.

122 Stirbu I, Kunst AE, Vlems FA, Visser O, Bos V, Deville W et al. Cancer mortality rates among first and second generation migrants in the Netherlands: Convergence toward the rates of the native Dutch population. Int J Cancer 2006; 119(11): 2665-2672.

123 Visser O, van Peppen AM, Ory FG, van Leeuwen FE. Results of breast cancer screening in first generation migrants in Northwest Netherlands. Eur J Cancer Prev 2005; 14(3): 251-255.

124 Vermeer B, Van den Muijsenbergh ME. The attendance of migrant women at the national breast cancer screening in the Netherlands 1997-2008. Eur J Cancer Prev 2010; 19(3): 195-198.

125 Yilmaz HH, Yazihan N, Tunca D, Sevinc A, Olcayto EÖ, Özgül N et al. Cancer trends and incidence and mortality patterns in Turkey. Jpn J Clin Oncol 2011; 41(1): 10-16.

126 Middelkoop BJ, Struben HW, Burger I, Vroom-Jongerden JM. Urban cause-specific socioeconomic mortality differences. Which causes of death contribute most? Int J Epidemiol 2001; 30(2): 240-247.

127 Paap E, Holland R, den Heeten GJ, van Schoor G, Botterweck AA, Verbeek AL et al. A remarkable reduction of breast cancer deaths in screened versus unscreened women: a case-referent study. Cancer Causes Control 2010; 21(10): 1569-1573.

128 Schoor G van, Paap E, Broeders MJ, Verbeek AL. Residual confounding after adjustment for age: a minor issue in breast cancer screening effectiveness. Eur J Epidemiol 2011; 26(8): 585-588.

129 Paap E, Verbeek A, Puliti D, Broeders M, Paci E. Minor influence of self-selection bias on the effectiveness of breast cancer screening in case-control studies in the Netherlands. J Med Screen 2011; 18(3): 142-146.
130 Puliti D, Miccinesi G, Collina N, De Lisi V, Federico M, Ferretti S et al. Effectiveness of service screening: a case-control study to assess breast cancer mortality reduction. Br J Cancer 2008; 99(3): 423-427.
131 Elmore JG, Reisch LM, Barton MB, Barlow WE, Rolnick S, Harris EL et al. Efficacy of breast cancer screening in the community according to risk level. J Natl Cancer Inst 2005; 97(14): 1035-1043.
132 Gabe R, Tryggvadottir L, Sigfusson BF, Olafsdottir GH, Sigurdsson K, Duffy SW. A case-control study to estimate the impact of the Icelandic population-based mammography screening program on breast cancer death. Acta Radiol 2007; 48(9): 948-955.
133 Otto SJ, Fracheboud J, Verbeek AL, Boer R, Reijerink-Verheij JC, Otten JD et al. Mammography screening and risk of breast cancer death: a population-based case-control study. Cancer Epidemiol Biomarkers Prev 2012; 21(1): 66-73.
134 Schoor G van, Moss SM, Otten JD, Donders R, Paap E, den Heeten GJ et al. Increasingly strong reduction in breast cancer mortality due to screening. Br J Cancer 2011; 104(6): 910-914.
135 Allgood PC, Warwick J, Warren RM, Day NE, Duffy SW. A case-control study of the impact of the East Anglian breast screening programme on breast cancer mortality. Br J Cancer 2008; 98(1): 206-209.
136 Fielder HM, Warwick J, Brook D, Gower-Thomas K, Cuzick J, Monypenny I et al. A case-control study to estimate the impact on breast cancer death of the breast screening programme in Wales. J Med Screen 2004; 11(4): 194-198.
137 Paap E, Verbeek AL, Puliti D, Paci E, Broeders MJ. Breast cancer screening case-control study design: impact on breast cancer mortality. Ann Oncol 2011; 22(4): 863-869.
138 Kiemeney LALM, Lemmers FAMO, Verhoeven RHA, Aben KKH, Honing C, de Nooijer J et al. De kans op kanker voor Nederlanders. Ned Tijdschr Geneeskd 2008; 152(41): 2233-2241.
139 Tabár L, Duffy SW, Yen MF, Warwick J, Vitak B, Chen HH et al. All-cause mortality among breast cancer patients in a screening trial: support for breast cancer mortality as an end point. J Med Screen 2002; 9(4): 159-162.
140 Mook S, van 't Veer LJ, Rutgers EJ, Ravdin PM, van de Velde AO, van Leeuwen FE et al. Independent prognostic value of screen detection in invasive breast cancer. J Natl Cancer Inst 2011; 103(7): 585-597.
141 Vilaprinyo E, Puig T, Rue M. Contribution of early detection and adjuvant treatments to breast cancer mortality reduction in Catalonia, Spain. PLoS One 2012; 7(1): e30157.
142 Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M et al. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med 2005; 353(17): 1784-1792.
143 Paci E, Ponti A, Zappa M, Patriarca S, Falini P, Delmastro G et al. Early diagnosis, not differential treatment, explains better survival in service screening. Eur J Cancer 2005; 41(17): 2728-2734.

144 Yaffe MJ, Mainprize JG, Jong RA. Technical developments in mammography. Health Phys 2008; 95(5): 599-611.

145 Landelijk Evaluatie Team voor bevolkingsonderzoek naar Borstkanker in Nederland. LETB tussenrapportage 2011. Belangrijkste resultaten 2008-2009 bevolkingsonderzoek borstkanker. Rotterdam: LETB; 2011.

146 Zoetelief J, Veldkamp WJ, Thijssen MA, Jansen JT. Glandularity and mean glandular dose determined for individual women at four regional breast cancer screening units in the Netherlands. Phys Med Biol 2006; 51(7): 1807-1817.

147 Hendrick RE, Pisano ED, Averbukh A, Moran C, Berns EA, Yaffe MJ et al. Comparison of acquisition parameters and breast dose in digital mammography and screen-film mammography in the American College of Radiology Imaging Network digital mammographic imaging screening trial. AJR Am J Roentgenol 2010; 194(2): 362-369.

148 National Research Council of the National Academies. Health risks from exposure to low levels of ionizing radiation: BEIR VII Phase 2ed. Washington, D.C.: 2006.

149 Gelder R de, Draisma G, Heijnsdijk EA, de Koning HJ. Population-based mammography screening below age 50: balancing radiation-induced vs prevented breast cancer deaths. Br J Cancer 2011; 104(7): 1214-1220.

150 Bijwaard H, Brenner A, Dekkers F, van Dillen T, Land CE, Boice JD, Jr. Breast cancer risk from different mammography screening practices. Radiat Res 2010; 174(3): 367-376.

151 Brito JP, Morris JC, Montori VM. Thyroid cancer: zealous imaging has increased detection and treatment of low risk tumours. BMJ 2013; 347:f4706.

152 Draisma G, Boer R, Otto SJ, van der Cruijsen I, Damhuis RA, Schröder FH et al. Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. J Natl Cancer Inst 2003; 95(12): 868-878.

153 Schilling FH, Spix C, Berthold F, Erttmann R, Fehse N, Hero B et al. Neuroblastoma screening at one year of age. N Engl J Med 2002; 346(14): 1047-1053.

154 Zahl PH, Götzsche PC, Maehlen J. Natural history of breast cancers detected in the Swedish mammography screening programme: a cohort study. Lancet Oncol 2011; 12(12): 1118-1124.

155 Horii R, Akiyama F, Kasumi F, Koike M, Sakamoto G. Spontaneous “healing” of breast cancer. Breast Cancer 2005; 12(2): 140-144.

156 Larsen SU, Rose C. [Spontaneous remission of breast cancer. A literature review]. Ugeskr Laeger 1999; 161(26): 4001-4004.

157 Ross MB, Buzdar AU, Hortobagyi GN, Lukeman JM. Spontaneous regression of breast carcinoma: follow-up report and literature review. J Surg Oncol 1982; 19(1): 22-24.

158 Nickerson HJ, Matthay KK, Seeger RC, Brodeur GM, Shimada H, Perez C et al. Favorable biology and outcome of stage IV-S neuroblastoma with supportive care or minimal therapy: a Children’s Cancer Group study. J Clin Oncol 2000; 18(3): 477-486.

159 Collins LC, Tamimi RM, Baer HJ, Connolly JL, Colditz GA, Schnitt SJ. Outcome of patients with ductal carcinoma in situ untreated after diagnostic biopsy: results from the Nurses' Health Study. Cancer 2005; 103(9): 1778-1784.

160 Sanders ME, Schuyler PA, Dupont WD, Page DL. The natural history of low-grade ductal carcinoma in situ of the breast in women treated by biopsy only revealed over 30 years of long-term follow-up. Cancer 2005; 103(12): 2481-2484.

161 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(12): 1129-1138.

162 Gelder R de, van As E, Tilanus-Linthorst MM, Bartels CC, Boer R, Draisma G et al. Breast cancer screening: evidence for false reassurance? Int J Cancer 2008; 123(3): 680-686.

163 Götzsche PC, Jörgensen KJ, Zahl P-H, Maehlen J. Why mammography screening has not lived up to expectations from the randomised trials. Cancer Causes Control 2012; 23: 15-21.

164 Weedon-Fekjaer H, Vatten LJ, Aalen OO, Lindqvist B, Tretli S. Estimating mean sojourn time and screening test sensitivity in breast cancer mammography screening: new results. J Med Screen 2005; 12(4): 172-178.

165 Tabár L, Fagerberg G, Duffy SW, Day NE, Gad A, Grontoft O. Update of the Swedish two-county program of mammographic screening for breast cancer. Radiol Clin North Am 1992; 30(1): 187-210.

166 Wu JC, Hakama M, Anttila A, Yen AM, Malila N, Sarkeala T et al. Estimation of natural history parameters of breast cancer based on non-randomized organized screening data: subsidiary analysis of effects of inter-screening interval, sensitivity, and attendance rate on reduction of advanced cancer. Breast Cancer Res Treat 2010; 122(2): 553-566.

167 Chiu SY-H, Duffy S, Yen AM-F, Tabár L, Smith RA, Chen H-H. Effect of baseline breast density on breast cancer incidence, stage, mortality, and screening parameters: 25-year follow-up of Swedish mammographic screening. Cancer Epidemiol Biomakers Prev 2010; 19(5): 1219-1228.

168 Paci E, Miccinesi G, Puliti D, Baldazzi P, De Lisi V, Falcini F et al. Estimate of overdiagnosis of breast cancer due to mammography after adjustment for lead time. A service screening study in Italy. Breast Cancer Res 2006; 8(6): R68.

169 Olsen AH, Agbaje OF, Myles JP, Lynge E, Duffy SW. Overdiagnosis, sojourn time, and sensitivity in the Copenhagen mammography screening program. Breast J 2006; 12(4): 338-342.

170 Etzioni R, Gulati R, Mallinger L, Mandelblatt J. Influence of study features and methods on overdiagnosis estimates in breast and prostate cancer screening. Ann Intern Med 2013; 158(11): 831-838.

171 Puliti D, Duffy SW, Miccinesi G, de Koning H., Lynge E, Zappa M et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. J Med Screen 2012; 19 Suppl 1: 42-56.

172 Duffy SW, Lynge E, Jonsson H, Ayyaz S, Olsen AH. Complexities in the estimation of overdiagnosis in breast cancer screening. Br J Cancer 2008; 99(7): 1176-1178.

173 Njor SH, Olsen AH, Blichert-Toft M, Schwartz W, Vejborg I, Lynge E. Overdiagnosis in screening mammography in Denmark: population based cohort study. BMJ 2013; 346: f1064.

174 Peeters PH, Verbeek AL, Straatman H, Holland R, Hendriks JH, Mravunac M et al. Evaluation of overdiagnosis of breast cancer in screening with mammography: results of the Nijmegen programme. Int J Epidemiol 1989; 18(2): 295-299.

175 Verbeek AL, Broeders MJ, Otto SJ, Frachebou J, Otten JD, Holland R et al. Effecten van het bevolkingsonderzoek naar borstkanker. Ned Tijdschr Geneeskd 2013; 157(10): A5218.

176 Gelder R de, Heijnsdijk EA, van Ravesteyn NT, Fracheboud J, Draisma G, de Koning HJ. Interpreting overdiagnosis estimates in population-based mammography screening. Epidemiol Rev 2011; 33(1): 111-121.

177 Gelder R de, Fracheboud J, Heijnsdijk EA, den Heeten G, Verbeek AL, Broeders MJ et al. Digital mammography screening: weighing reduced mortality against increased overdiagnosis. Prev Med 2011; 53(3): 134-140.

178 Jörgensen KJ, Götzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. BMJ 2009; 339: b2587.

179 Kopans DB, Smith RA, Duffy SW. Mammographic screening and “overdiagnosis”. Radiology 2011; 260(3): 616-620.

180 Johnson A, Shekhdar J. Breast cancer incidence: what do the figures mean? J Eval Clin Pract 2005; 11(1): 27-31.

181 Svendsen AL, Olsen AH, von Euler-Chelpin M, Lynge E. Breast cancer incidence after the introduction of mammography screening: what should be expected? Cancer 2006; 106(9): 1883-1890.

182 Jörgensen KJ, Zahl P-H, Götzsche PC. Overdiagnosis in organised mammography screening in Denmark. A comparative study. BMC Women’s Health 2009; 9(36): .

183 Kalager M, Adami HO, Bretthauer M, Tamimi RM. Overdiagnosis of invasive breast cancer due to mammography screening: results from the Norwegian screening program. Ann Intern Med 2012; 156(7): 491-499.

184 Falk RS, Skaane P, Hofvind S. Overdiagnosis of Invasive Breast Cancer due to Mammography Screening [letter]. Annals of Internal Medicine 2012; 157(3):219.

185 Haldorsen T, Tretli S, Ursin G. Overdiagnosis of Invasive Breast Cancer due to Mammography Screening [letter]. Annals of Internal Medicine 2012; 157(3):220.

186 Waller M, Moss S, Watson J, Moller H. The effect of mammographic screening and hormone replacement therapy use on breast cancer incidence in England and Wales. Cancer Epidemiol Biomarkers Prev 2007; 16(11): 2257-2261.

187 Puliti D, Zappa M, Miccinesi G, Falini P, Crocetti E, Paci E. An estimate of overdiagnosis 15 years after the start of mammographic screening in Florence. Eur J Cancer 2009; 45(18): 3166-3171.

188 Seigneurin A, François O, Labarère J, Oudeville P, Monlong J, Colonna M. Overdiagnosis from non-progressive cancer detected by screening mammography: stochastic simulation study with calibration to population based registry data. BMJ 2011; 343: d7017.

189 Olsen O, Götzsche PC. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2001;(4): CD001877.
190 Jörgensen KJ, Keen JD, Götzsche PC. Is mammographic screening justifiable considering its substantial overdiagnosis rate and minor effect on mortality? Radiology 2011; 260(3): 621-627.
191 Hayes BD, Fitzpatrick P, Quinn C. Mastectomy rates in the Irish national breast screening programme, 2000-2010. Br Med J (www bmj com/content/343/bmj d4692?tab=responses) 2011;
192 Suhrke P, Maehlen J, Schlichting E, Jörgensen KJ, Götzsche PC, Zahl PH. Effect of mammography screening on surgical treatment for breast cancer in Norway: comparative analysis of cancer registry data. BMJ 2011; 343: d4692.
193 Walsh PM, McCarron P, Middleton RJ, Comber H, Gavin AT, Murray L. Influence of mammographic screening on trends in breast-conserving surgery in Ireland. Eur J Cancer Prev 2006; 15(2): 138-148.
194 Juffs HG, Tannock IF. Screening trials are even more difficult than we thought they were. J Natl Cancer Inst 2002; 94(3): 156-157.
195 Darby SC, McGale P, Taylor CW, Peto R. Long-term mortality from heart disease and lung cancer after radiotherapy for early breast cancer: prospective cohort study of about 300,000 women in US SEER cancer registries. Lancet Oncol 2005; 6(8): 557-565.
196 Darby S, McGale P, Correa C, Taylor C, Arriagada R, Clarke M et al. Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. Lancet 2011; 378(9804): 1707-1716.
197 Morton LM, Gilbert ES, Hall P, Andersson M, Joensuu H, Vaalavirta L et al. Risk of treatment-related esophageal cancer among breast cancer survivors. Ann Oncol 2012; 23(12): 3081-3091.
198 Giard R. Twijfel aan nut van borstkankerscreening. Huisarts Wet 2002; 45(03): 110-112.
199 Götzsche PC. Time to stop mammography screening? CMAJ 2011; 183(17): 1957-1958.
200 Brenner DJ, Shuryak I, Jozsef G, Dewyngaert KJ, Formenti SC. Risk and risk reduction of major coronary events associated with contemporary breast radiotherapy. JAMA Intern Med 2014; 174(1): 158-160.
201 Darby SC, Ewertz M, McGale P, Bennet AM, Blom-Goldman U, Bronnum D et al. Risk of ischemic heart disease in women after radiotherapy for breast cancer. N Engl J Med 2013; 368(11): 987-998.
202 McGale P, Darby SC, Hall P, Adolfsson J, Bengtsson NO, Bennet AM et al. Incidence of heart disease in 35,000 women treated with radiotherapy for breast cancer in Denmark and Sweden. Radiother Oncol 2011; 100(2): 167-175.
203 Formenti SC, DeWyngaert JK. Positioning during radiotherapy for breast cancer. JAMA 2013; 309(2): 137.
204 Formenti SC, DeWyngaert JK, Jozsef G, Goldberg JD. Prone vs supine positioning for breast cancer radiotherapy. JAMA 2012; 308(9): 861-863.
205 Hayden AJ, Rains M, Tiver K. Deep inspiration breath hold technique reduces heart dose from radiotherapy for left-sided breast cancer. J Med Imaging Radiat Oncol 2012; 56(4): 464-472.

206 Alderliesten T, Betgen A, Elkhuizen PH, van Vliet-Vroegindeweij C, Remeijer P. Estimation of heart-position variability in 3D-surface-image-guided deep-inspiration breath-hold radiation therapy for left-sided breast cancer. Radiother Oncol 2013; 109(3): 442-447.
207 Giordano SH, Kuo YF, Freeman JL, Buchholz TA, Hortobagyi GN, Goodwin JS. Risk of cardiac death after adjuvant radiotherapy for breast cancer. J Natl Cancer Inst 2005; 97(6): 419-424.
208 Hooning MJ, Botma A, Aleman BM, Baaijens MH, Bartelink H, Klijn JG et al. Long-term risk of cardiovascular disease in 10-year survivors of breast cancer. J Natl Cancer Inst 2007; 99(5): 365-375.
209 Drossaert CHC, Boekema AG, Boer H, Seydel ER. Psychosociale gevolgen van borstkankerscreening. Tijdschrift voor gezondheidswetenschappen 1998; 76(2): 59-66.
210 Poulos A, McLean D, Rickard M, Heard R. Breast compression in mammography: how much is enough? Australas Radiol 2003; 47(2): 121-126.
211 Keyzer-Dekker CMG, De Vries J, Roukema JA, Van der Steeg AFW. Nadelige effecten van een afwijkend screeningsmammogram op kwaliteit van leven en angstbeleving. Ned Tijdschr Geneeskd 2011; 155(28): 1243-1248.
212 Bond M, Pavey T, Welch K, Cooper C, Garside R, Dean S et al. Psychological consequences of false-positive screening mammograms in the UK. Evid Based Med 2013; 18(2): 54-61.
213 Brodersen J, Siersma VD. Long-term psychosocial consequences of false-positive screening mammography. Ann Fam Med 2013; 11(2): 106-115.
214 Roukema JA. Niets is zeker. Ned Tijdschr Geneeeskd 2013; 157: A5384
215 Setz-Pels W, Duijm LE, Coebergh JW, Rutten M, Nederend J, Voogd AC. Re-attendance after false-positive screening mammography: a population-based study in the Netherlands. Br J Cancer 2013; 109(8): 2044-2050.
216 Hubbard RA, Kerlikowske K, Flowers CI, Yankaskas BC, Zhu W, Miglioretti DL. Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. Ann Intern Med 2011; 155(8): 481-492.
217 Sala M, Salas D, Belvis F, Sanchez M, Ferrer J, Ibanez J et al. Reduction in false-positive results after introduction of digital mammography: analysis from four population-based breast cancer screening programs in Spain. Radiology 2011; 258(2): 388-395.
218 Castells X, Molins E, Macia F. Cumulative false positive recall rate and association with participant related factors in a population based breast cancer screening programme. J Epidemiol Community Health 2006; 60(4): 316-321.
219 Roman R, Sala M, Salas D, Ascunce N, Zubizarreta R, Castells X. Effect of protocol-related variables and women's characteristics on the cumulative false-positive risk in breast cancer screening. Ann Oncol 2012; 23(1): 104-111.
220 Otten JD, Fracheboud J, den Heeten GJ, Otto SJ, Holland R, de Koning HJ et 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(10): 2501-2506.
221 Horst F van der, Hendriks JHCL, Rijken HJTM, Holland R. Breast cancer screening in the Netherlands: audit and training of radiologists. Seminars in Breast Disease 2003; 6(3): 114-122.

222 Broeders MJ, Onland-Moret NC, Rijken HJ, Hendriks JH, Verbeek AL, Holland R. Use of previous screening mammograms to identify features indicating cases that would have a possible gain in prognosis following earlier detection. Eur J Cancer 2003; 39(12): 1770-1775.

223 Ciatto S, Catarzi S, Lamberini MP, Risso G, Saguatti G, Abbattista T et al. Interval breast cancers in screening: the effect of mammography review method on classification. Breast 2007; 16(6): 646-652.

224 Rijke JM de, Schouten LJ, Schreutelkamp JL, Jochem I, Verbeek AL. A blind review and an informed review of interval breast cancer cases in the Limburg screening programme, the Netherlands. J Med Screen 2000; 7(1): 19-23.

225 Webster P, Austoker J. Women's knowledge about breast cancer risk and their views of the purpose and implications of breast screening--a questionnaire survey. J Public Health (Oxf) 2006; 28(3): 197-202.

226 Welch HG, Frankel BA. Likelihood that a woman with screen-detected breast cancer has had her “life saved” by that screening. Arch Intern Med 2011; 171(22): 2043-2046.

227 U.S.Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2009; 151(10): 716-236.

228 Tabár L, Faberberg G, Day NE, Holmberg L. What is the optimum interval between mammographic screening examinations? An analysis based on the latest results of the Swedish two-county breast cancer screening trial. Br J Cancer 1987; 55(5): 547-551.

229 College voor zorgverzekeringen. Richtlijnen voor farmaco-economisch onderzoek; evaluatie en actualisatie. Diemen: CvZ; 2005.

230 Oostenbrink JB, Bouwmans CAM, Koopmanschap MA, Rutten FFH. Handleiding voor kostenonderzoek: methoden en richtlijnprijzen voor economische evaluatie in de gezondheidszorg. Amstelveen: 2004.

231 Earnshaw J, Lewis G. NICE guide to the methods of technology appraisal: pharmaceutical industry perspective. Pharmacoeconomics 2008; 26(9): 725-727.

232 Carles M, Vilaprinyo E, Cots F, Gregori A, Pla R, Roman R et al. Cost-effectiveness of early detection of breast cancer in Catalonia (Spain). BMC Cancer 2011; 11: 192.

233 Norum J. Breast cancer screening by mammography in Norway. Is it cost-effective? Ann Oncol 1999; 10(2): 197-203.

234 Leivo T, Sintonen H, Tuominen R, Hakama M, Pukkala E, Heinonen OP. The cost-effectiveness of nationwide breast carcinoma screening in Finland, 1987-1992. Cancer 1999; 86(4): 638-646.

235 Sankatsing VD, Heijnsdijk E, van Luijt PA, van Ravesteyn NT, Fracheboud J, de Koning HJ. Kosteneffectiviteitsanalyse met betrekking tot het screenen van vrouwen onder de 50 jaar in Nederland. Erasmus MC, Maatschappelijke Gezondheidszorg. November 2013.

236 Pharoah PD, Sewell B, Fitzsimmons D, Bennett HS, Pashayan N. Cost effectiveness of the NHS breast screening programme: life table model. BMJ 2013; 346: f2618.

237 Raad voor de Volksgezondheid en Zorg. Zinnige en duurzame zorg. Den Haag: Raad voor de Volksgezondheid en Zorg; 2006: Advies nr. 06/06.

238 Luijt PA van, Fracheboud J, Heijnsdijk EA, den Heeten GJ, de Koning HJ. 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.

239 Gezondheidsraad. Screening op baarmoederhalskanker. Den Haag: Gezondheidsraad; 2011: publicatie nr 2011/07.

240 Gezondheidsraad. Bevolkingsonderzoek naar darmkanker. Den Haag: Gezondheidsraad; 2009: 2009/13.

241 Haes JC de, de Koning HJ, van Oortmarssen GJ, van Agt HM, de Bruyn AE, van der Maas PJ. The impact of a breast cancer screening programme on quality-adjusted life-years. Int J Cancer 1991; 49(4): 538-544.

242 Nationaal Borstkanker Overleg Nederland. Richtlijn Mammacarcinoom. Amsterdam: NABON; 2008.

243 Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. Ann Intern Med 2009; 151(10): 727-742.

244 Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet 2006; 368(9552): 2053-2060.

245 Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. N Engl J Med 2005; 353(17): 1773-1783.

246 Broeders M, van Schoor G, Bluekens A, Visser R, Otten JDM, Paap E et al. Rapportage over een eventuele verlaging van de onderste leeftijdsgrens van het bevolkingsonderzoek op borstkanker. Een analyse van de literatuur en vertaling van de bevindingen naar de Nederlanse situatie. Nijmegen: Landelijk Referentiecentrum voor het Bevolkingsonderzoek op Borstkanker en Afd. Epidemiologie, Biostatistiek en HTA, UMC St Radboud, 2009.

247 Hellquist BN, Duffy SW, Abdsaleh S, Björneld L, Bordás P, Tabár L et al. 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(4): 714-722.

248 Schoor G van, Moss SM, Otten JD, Donders R, Paap E, den Heeten GJ et al. Effective biennial mammographic screening in women aged 40-49. Eur J Cancer 2010; 46(18): 3137-3140.

249 Mandelblatt JS, Cronin KA, Berry DA, Chang Y, de Koning HJ, Lee SJ et al. Modeling the impact of population screening on breast cancer mortality in the United States. Breast 2011; 20 Suppl 3: S75-S81.

250 Ravesteyn NT van, Miglioretti DL, Stout NK, Lee SJ, Schechter CB, Buist DS 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(9): 609-617.

251 Bennett RL, Moss SM. Screening outcomes in women over age 70 who self-refer in the NHSBSP in England. J Med Screen 2011; 18(2): 91-95.

252 Schoor G van. Breast cancer screening: Increasing benefits and extending scope [Proefschrift]. Radboud Universiteit Nijmegen; 2012.
253 Hofvind S, Bjurstam N, Sorum R, Bjorndal H, Thoresen S, Skaane P. Number and characteristics of breast cancer cases diagnosed in four periods in the screening interval of a biennial population-based screening programme. J Med Screen 2006; 13(4): 192-196.
254 Törnberg S, Kemetli L, Ascunce N, Hofvind S, Anttila A, Sèradour B et al. A pooled analysis of interval cancer rates in six European countries. Eur J Cancer Prev 2010; 19(2): 87-93.
255 Schoor G van, Otten JD, den Heeten GJ, Holland R, Broeders MJ, Verbeek AL. Breast cancer among women over 75 years: an important public health problem? Eur J Public Health 2012; 22(3): 422-424.
256 McPherson CP, Swenson KK, Lee MW. The effects of mammographic detection and comorbidity on the survival of older women with breast cancer. J Am Geriatr Soc 2002; 50(6): 1061-1068.
257 Badgwell BD, Giordano SH, Duan ZZ, Fang S, Bedrosian I, Kuerer HM et al. Mammography before diagnosis among women age 80 years and older with breast cancer. J Clin Oncol 2008; 26(15): 2482-2488.
258 Dijck JA van, Holland R, Verbeek AL, Hendriks JH, Mravunac M. Efficacy of mammographic screening of the elderly: a case-referent study in the Nijmegen program in The Netherlands. J Natl Cancer Inst 1994; 86(12): 934-938.
259 Dijck JA van, Verbeek AL, Beex LV, Hendriks JH, Holland R, Mravunac M et al. Mammographic screening after the age of 65 years: evidence for a reduction in breast cancer mortality. Int J Cancer 1996; 66(6): 727-731.
260 Duijm LE, Groenewoud JH, de Koning HJ, Coebergh JW, van Beek M., Hooijen MJ et al. Delayed diagnosis of breast cancer in women recalled for suspicious screening mammography. Eur J Cancer 2009; 45(5): 774-781.
261 Kievit J. De invloed van uitstel van behandeling op de prognose bij patiënten met borstkanker. Ned Tijdschr Geneeskd 2002; 146(22): 1031-1034.
262 Voorde M ten, Kroeze-Voogd EH. Vergelijking F.A.S.T. compressie paddle en standaard paddle. Enschede: Bevolkingsonderzoek Oost, Business unit Borstkanker Enschede/Landelijk Referentiecentrum voor Bevolkingsonderzoek; 2011.
263 Broeders MJM, Den Heeten GJ, Grimbergen CA. Is pain an unavoidable side effect of mammography? A 'proof-of-principle' study on pressure-guided compression. Nijmegen: Landelijk Referentiecentrum voor Bevolkingsonderzoek; 2013.
264 Cole EB, Toledano AY, Lundqvist M, Pisano ED. Comparison of radiologist performance with photon-counting full-field digital mammography to conventional full-field digital mammography. Acad Radiol 2012; 19(8): 916-922.
265 Marteau TM, Dormandy E, Michie S. A measure of informed choice. Health Expect 2001; 4(2): 99-108.
266 Steckelberg A, Hulfenhaus C, Haastert B, Muhlhauser I. Effect of evidence based risk information on "informed choice" in colorectal cancer screening: randomised controlled trial. BMJ 2011; 342: d3193.

267 Agt HME van, Fracheboud J, Rebolj M, Korfage IJ, de Koning HJ. Volledige, evenwichtige en eerlijke voorlichting over nut en risico's van bevolkingsonderzoek naar kanker. Rotterdam: Erasmus MC, afdeling Maatschappelijke Gezondheidszorg; 2008.

268 Agt H van, Fracheboud J, van der Steen A, de Koning HJ. Do women make an informed choice about participating in breast cancer screening? A survey among women invited for a first mammography screening examination. Patient Educ Couns 2012; 89(2): 353-359.

269 Essink-Bot ML, Korfage IJ, de Koning HJ. 'Geïnformeerde keuze'. Uitdaging voor bevolkingsonderzoek borstkanker. Ned Tijdschr Geneeskd 2013; 157(9): A5756.

270 Waard F de, Collette HJ, Rombach JJ, Collette C. Breast cancer screening, with particular reference to the concept of 'high risk' groups. Breast Cancer Res Treat 1988; 11(2): 125-132.

271 Dent T, Jbilou J, Rafi I, Segnan N, Törnberg S, Chowdhury S et al. Stratified cancer screening: the practicalities of implementation. Public Health Genomics 2013; 16(3): 94-99.

272 Evans DG, Warwick J, Astley SM, Stavrinos P, Sahin S, Ingham S et al. Assessing individual breast cancer risk within the U.K. National Health Service Breast Screening Program: a new paradigm for cancer prevention. Cancer Prev Res (Phila) 2012; 5(7): 943-951.

273 Howell A, Astley S, Warwick J, Stavrinos P, Sahin S, Ingham S et al. Prevention of breast cancer in the context of a national breast screening programme. J Intern Med 2012; 271(4): 321-330.

274 Soini I, Hakama M. Failure of selective screening for breast cancer by combining risk factors. Int J Cancer 1978; 22(3): 275-281.

275 Pfeiffer RM, Park Y, Kreimer AR, Lacey JV, Jr., Pee D, Greenlee RT et al. Risk prediction for breast, edometrial, and ovarian cancer in white women aged 50 y or older: derivation and validation from population-based cohort studies. PLoS Med 2013; 10(7): e1001492.

276 Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 1989; 81(24): 1879-1886.

277 Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. Stat Med 2004; 23(7): 1111-1130.

278 Santen RJ, Boyd NF, Chlebowski RT, Cummings S, Cuzick J, Dowsett M et al. Critical assessment of new risk factors for breast cancer: considerations for development of an improved risk prediction model. Endocr Relat Cancer 2007; 14(2): 169-187.

279 Brennan K, Garcia-Closas M, Orr N, Fletcher O, Jones M, Ashworth A et al. Intragenic ATM methylation in peripheral blood DNA as a biomarker of breast cancer risk. Cancer Res 2012; 72(9): 2304-2313.

280 Holmberg L, Vickers A. Evaluation of prediction models for decision-making: beyond calibration and discrimination. PLoS Med 2013; 10(7): e1001491.

281 Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010; 21(1): 128-138.

282 Olsen AH, Bihrmann K, Jensen MB, Vejborg I, Lynge E. Breast density and outcome of mammography screening: a cohort study. Br J Cancer 2009; 100(7): 1205-1208.
283 Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med 2007; 356(3): 227-236.
284 Boyd NF, Martin LJ, Bronskill M, Yaffe MJ, Duric N, Minkin S. Breast tissue composition and susceptibility to breast cancer. J Natl Cancer Inst 2010; 102(16): 1224-1237.
285 Engeland S van, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. IEEE Trans Med Imaging 2006; 25(3): 273-282.
286 Cecchini RS, Costantino JP, Cauley JA, Cronin WM, Wickerham DL, Bandos H et al. Baseline mammographic breast density and the risk of invasive breast cancer in postmenopausal women participating in the NSABP study of tamoxifen and raloxifene (STAR). Cancer Prev Res (Phila) 2012; 5(11): 1321-1329.
287 Vachon CM, van Gils CH, Sellers TA, Ghosh K, Pruthi S, Brandt KR et al. Mammographic density, breast cancer risk and risk prediction. Breast Cancer Res 2007; 9(6): 217.
288 Milne RL, Herranz J, Michailidou K, Dennis J, Tyrer JP, Zamora MP et al. A large-scale assessment of two-way SNP interactions in breast cancer susceptibility using 46,450 cases and 42,461 controls from the Breast Cancer Association Consortium. Hum Mol Genet 2013; doi: 10.1093/hmg/ddt581 [Epub ahead of print].
289 Zitteren M van, van der Net JB, Kundu S, Freedman AN, van Duijn CM, Janssens AC. Genome-based prediction of breast cancer risk in the general population: a modeling study based on meta-analyses of genetic associations. Cancer Epidemiol Biomarkers Prev 2011; 20(1): 9-22.
290 Pashayan N, Duffy SW, Chowdhury S, Dent T, Burton H, Neal DE et al. Polygenic susceptibility to prostate and breast cancer: implications for personalised screening. Br J Cancer 2011; 104(10): 1656-1663.
291 Nickels S, Truong T, Hein R, Stevens K, Buck K, Behrens S et al. Evidence of gene-environment interactions between common breast cancer susceptibility loci and established environmental risk factors. PLoS Genet 2013; 9(3): e1003284.
292 Wacholder S, Hartge P, Prentice R, Garcia-Closas M, Feigelson HS, Diver WR et al. Performance of common genetic variants in breast-cancer risk models. N Engl J Med 2010; 362(11): 986-993.
293 Huijts PE, Hollestelle A, Balliu B, Houwing-Duistermaat JJ, Meijers CM, Blom JC et al. CHEK2*1100delC homozygosity in the Netherlands-prevalence and risk of breast and lung cancer. Eur J Hum Genet 2013; doi: 10.1038/ejhg.2013.85. [Epub ahead of print]
294 Adank MA, Verhoef S, Oldenburg RA, Schmidt MK, Hooning MJ, Martens JW et al. Excess breast cancer risk in first degree relatives of CHEK2 *1100delC positive familial breast cancer cases. Eur J Cancer 2013; 49(8): 1993-1999.
295 Cummings SR, Lee JS, Lui LY, Stone K, Ljung BM, Cauleys JA. Sex hormones, risk factors, and risk of estrogen receptor-positive breast cancer in older women: a long-term prospective study. Cancer Epidemiol Biomarkers Prev 2005; 14(5): 1047-1051.

296 Zhang X, Tworoger SS, Eliassen AH, Hankinson SE. Postmenopausal plasma sex hormone levels and breast cancer risk over 20 years of follow-up. Breast Cancer Res Treat 2013; 137(3): 883-892.

297 Key TJ, Appleby PN, Reeves GK, Roddam AW, Helzlsouer KJ, Alberg AJ et al. Circulating sex hormones and breast cancer risk factors in postmenopausal women: reanalysis of 13 studies. Br J Cancer 2011; 105(5): 709-722.

298 Woolcott CG, Courneya KS, Boyd NF, Yaffe MJ, McTiernan A, Brant R et al. Association between sex hormones, glucose homeostasis, adipokines, and inflammatory markers and mammographic density among postmenopausal women. Breast Cancer Res Treat 2013; 139(1): 255-265.

299 Varghese JS, Smith PL, Folkerd E, Brown J, Leyland J, Audley T et al. The heritability of mammographic breast density and circulating sex-hormone levels: two independent breast cancer risk factors. Cancer Epidemiol Biomarkers Prev 2012; 21(12): 2167-2175.

300 Kerlikowske K, Hubbard RA, Miglioretti DL, Geller BM, Yankaskas BC, Lehman CD et al. Comparative effectiveness of digital versus film-screen mammography in community practice in the United States: a cohort study. Ann Intern Med 2011; 155(8): 493-502.

301 White E, Miglioretti DL, Yankaskas BC, Geller BM, Rosenberg RD, Kerlikowske K et al. Biennial versus annual mammography and the risk of late-stage breast cancer. J Natl Cancer Inst 2004; 96(24): 1832-1839.

302 Skaane P, Bandos AI, Gullien R, Eben EB, Ekseth U, Haakenaasen U et al. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. Radiology 2013; 267(1): 47-56.

303 Rose SL, Tidwell AL, Bujnoch LJ, Kushwaha AC, Nordmann AS, Sexton R, Jr. Implementation of breast tomosynthesis in a routine screening practice: an observational study. AJR Am J Roentgenol 2013; 200(6): 1401-1408.

304 Ciatto S, Houssami N, Bernardi D, Caumo F, Pellegrini M, Brunelli S et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. Lancet Oncol 2013; 14(7): 583-589.

305 Gur D, Zuley ML, Anello MI, Rathfon GY, Chough DM, Ganott MA et al. Dose reduction in digital breast tomosynthesis (DBT) screening using synthetically reconstructed projection images: an observer performance study. Acad Radiol 2012; 19(2): 166-171.

306 Houssami N, Skaane P. Overview of the evidence on digital breast tomosynthesis in breast cancer detection. Breast 2013; 22(2): 101-108.

307 Gils CH van, Veldhuis WB, Peeters PH. Borstkankerscreenng op maat met echo en MRI? Ned Tijdschr Geneeskd 2012; 156(37): A5313.

308 Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. JAMA 2012; 307(13): 1394-1404.

309 Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Bohm-Velez M et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. JAMA 2008; 299(18): 2151-2163.

310 Verkooijen HM, Pijnappel R. Borstkankerscreening: meer is niet per se beter. Ned Tijdschr Geneeskd 2012; 156: A5098.
311 Girardi V, Tonegutti M, Ciatto S, Bonetti F. Breast ultrasound in 22,131 asymptomatic women with negative mammography. Breast 2013; 22(5): 806-809.
312 Lander MR, Tabár L. Automated 3-D breast ultrasound as a promising adjunctive screening tool for examining dense breast tissue. Semin Roentgenol 2011; 46(4): 302-308.
313 Giuliano V, Giuliano C. Improved breast cancer detection in asymptomatic women using 3D-automated breast ultrasound in mammographically dense breasts. Clin Imaging 2013; 37(3): 480-486.
314 Heeten GJ den, Karssemeijer N. Computerbeoordeling van screeningsmammogrammen. Ned Tijdschr Geneeskd 2011; 155(18): A3035.
315 Gilbert FJ, Astley SM, Gillan MG, Agbaje OF, Wallis MG, James J et al. Single reading with computer-aided detection for screening mammography. N Engl J Med 2008; 359(16): 1675-1684.
316 Guerriero C, Gillan MG, Cairns J, Wallis MG, Gilbert FJ. Is computer aided detection (CAD) cost effective in screening mammography? A model based on the CADET II study. BMC Health Serv Res 2011; 11: 11.
317 Azavedo E, Zackrisson S, Mejare I, Heibert AM. Is single reading with computer-aided detection (CAD) as good as double reading in mammography screening? A systematic review. BMC Med Imaging 2012; 12: 22.
318 Hupse R, Samulski M, Lobbes MB, Mann RM, Mus R, den Heeten GJ et al. Computer-aided detection of masses at mammography: interactive decision support versus prompts. Radiology 2013; 266(1): 123-129.
319 Jong PA de, van der Graaf Y, Bossuyt PM. Te weinig evaluatie van beeldvormende onderzoeken. Ned Tijdschr Geneeskd 2011; 155: A2996.
320 Arora N, Martins D, Ruggerio D, Tousimis E, Swistel AJ, Osborne MP et al. Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer. Am J Surg 2008; 196(4): 523-526.
321 Heijblom M, Piras D, Xia W, van Hespen JC, Klaase JM, van den Engh FM et al. Visualizing breast cancer using the Twente photoacoustic mammoscope: what do we learn from twelve new patient measurements? Opt Express 2012; 20(11): 11582-11597.
322 Johnson R, Jalleh G, Pratt IS, Donovan RJ, Lin C, Saunders C et al. Online advertising by three commercial breast imaging services: Message takeout and effectiveness. Breast 2013; 22(5): 780-786.
323 Vreugdenburg TD, Willis CD, Mundy L, Hiller JE. A systematic review of elastography, electrical impedance scanning, and digital infrared thermography for breast cancer screening and diagnosis. Breast Cancer Res Treat 2013; 137(3): 665-676.
324 Hendrick RE. Radiation doses and cancer risks from breast imaging studies. Radiology 2010; 257(1): 246-253.

325 Esserman L, Shieh Y, Thompson I. Rethinking screening for breast cancer and prostate cancer. JAMA 2009; 302(15): 1685-1692.

326 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.

327 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.

328 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.

329 Glynn CG, Farria DM, Monsees BS, Salcman JT, Wiele KN, Hildebolt CF. Effect of transition to digital mammography on clinical outcomes. Radiology 2011; 260(3): 664-670.

330 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.

331 Visser R, Veldkamp WJ, Beijerinck D, Bun PA, Deurenberg JJ, Imhof-Tas MW et al. Increase in perceived case suspiciousness due to local contrast optimisation in digital screening mammography. Eur Radiol 2012; 22(4): 908-914.

332 Soumian S, Verghese ET, Booth M, Sharma N, Chaudhri S, Bradley S et al. Concordance between vacuum assisted biopsy and postoperative histology: Implications for the proposed Low Risk DCIS Trial (LORIS). Eur J Surg Oncol 2013; 39(12): 1337-1340.

333 Hughes KS, Schnaper LA, Bellon JR, Cirrincione CT, Berry DA, McCormick B et al. Lumpectomy plus tamoxifen with or without irradiation in women age 70 years or older with early breast cancer: Long-term follow-up of CALGB 9343. J Clin Oncol 2013; 31(19): 2382-2387.

334 Mouw KW, Harris JR. Hypofractionation in the era of modulated radiotherapy (RT). Breast 2013; 22 Suppl 2: S129-S136.

335 Vaidya JS, Wenz F, Bulsara M, Tobias JS, Joseph DJ, Keshtgar M et al. Risk-adapted targeted intraoperative radiotherapy versus whole-breast radiotherapy for breast cancer: 5-year results. Lancet 2013; doi: 10.1016/S0140-6736(13)61950-9 [Epub ahead of print].

336 Veronesi U, Orecchia R, Maisonneuve P, Viale G, Rotmensz N, Sangalli C et al. Intraoperative radiotherapy versus external radiotherapy for early breast cancer (ELIOT): a randomised controlled equivalence trial. Lancet Oncol 2013; 14(13): 1269-1277.

337 Groot JE de, Broeders MJ, Branderhorst W, den Heeten GJ, Grimbergen CA. A novel approach to mammographic breast compression: Improved standardization and reduced discomfort by controlling pressure instead of force. Med Phys 2013; 40(8): 081901.

338 Sala M. Reduction in false-positive results after introduction of digital mammography. Radiol 2011; 258(2): 388-395.

339 Kreuter MW. Dealing with competing and conflicting risks in cancer communication. J Natl Cancer Inst Monogr. 1999; (25): 27-35.

Annexes

Annex **A**

Request for advice

June 15th 2012 the president of the Health Council of the Netherlands received a request for advice by the Minister of Health, Welfare and Sport about population based screening for breast cancer. Translated from the original Dutch request the Minister asked (letter PG/OGZ 3117093):

[…]I ask you to review and interpret the current scientific knowledge and to report your findings on the efficacy, the balance between the benefits and harms of population based screening [for breast cancer] in Netherlands.

[…]The Government and it's acting stakeholders continually strive to improve the screening programme. Are we doing the right things and do we do them well? So I ask you not only to assess the current population based screening programme, but also to assess the short and medium term developments. Which are the priorities for optimizing the screening programme? And what are the possible changes and improvements in the screening programme?

Annex **B**

The Committee

- Prof. dr. J.J.M. van Delden, MD, PhD, *chairman*
  Professor of Medical Ethics, University Medical Center Utrecht, Utrecht
- Prof. dr. J. Gussekloo, MD, PhD,
  Professor of General Medicine, Leiden University Medical Center, Leiden,
- Dr. E.M.M. Adang
  Health Economist, Radboud University Medical Center, Nijmegen
- Dr. M.M. Boere-Boonekamp, MD, PhD
  Physician, University Twente, Enschede
- Prof. dr. M.C. Cornel, MD, PhD
  Professor of Community Genetics and Public Health Genomics, VU Medisch Centrum, Amsterdam
- Dr. W.J. Dondorp, PhD
  Ethicist, Maastricht University Medical Centre, Maastricht
- Prof. mr. J.C.J. Dute, PhD
  Professor of Health Law, Radboud University, Nijmegen
- Prof. dr. L.P. ten Kate, MD, PhD
  Emeritus professor in Clinical Genetics, VU University Medical Centre, Amsterdam
- Prof. dr. I.M. van Langen, MD, PhD
  Professor of Clinical Genetics, University Medical Center Groningen
- Prof. dr. B.J.C. Middelkoop, MD, PhD
  Professor of Public Health, Leiden University Medical Center, Leiden

- Mr. dr. M.C. Ploem, PhD
  Lecturer of Health Law, Academic Medical Center, Amsterdam
- Prof. dr. A.M. Stiggelbout, MD, PhD,
  Professor of Medical Decision Making, Leiden University Medical Center
- W.A. van Veen, MD, Delft
- Prof. dr. G.J. van der Wilt, PhD
  Professor of Health Technology Assessment, Radboud university medical center, Nijmegen
- Dr. C.H. van Gils, PhD
  Epidemiologist, University Medical Center Utrecht
- Prof. dr. A.L.M. Verbeek, MD, PhD
  Professor of Clinical Epidemiology, Radboud university medical center, Nijmegen
- Mr. A. Rendering, LLM, *observer*
  Ministry for Health, Welfare and Sport, The Hague
- Dr. S.J.W. Kunst, MD, PhD, *scientific secretary*
  Health Council of Netherlands, The Hague
- Dr. L.G.M. van Rossum, PhD, *scientific secretary*
  Epidemiologist, Health Council of Netherlands, The Hague

The Health Council and interests

Members of Health Council Committees are appointed in a personal capacity because of their special expertise in the matters to be addressed. Nonetheless, it is precisely because of this expertise that they may also have interests. This in itself does not necessarily present an obstacle for membership of a Health Council Committee. Transparency regarding possible conflicts of interest is nonetheless important, both for the chairperson and members of a Committee and for the President of the Health Council. On being invited to join a Committee, members are asked to submit a form detailing the functions they hold and any other material and immaterial interests which could be relevant for the Committee's work. It is the responsibility of the President of the Health Council to assess whether the interests indicated constitute grounds for non-appointment. An advisorship will then sometimes make it possible to exploit the expertise of the specialist involved. During the inaugural meeting the declarations issued are discussed, so that all members of the Committee are aware of each other's possible interests.

Annex **C**

Experts consulted

- Prof. dr. H. Bartelink, MD, PhD
  Emeritus Professor of Clinical Experimental Radiotherapy, Academic Medical Center, Amsterdam; The Netherlands Cancer Institute, Amsterdam
- Prof. dr. G.J. den Heeten, MD, PhD
  Professor of Radiology, Academic Medical Center, Amsterdam; National Expert and Training Centre For Breast Cancer Screening, Nijmegen
- Prof. dr. W.P.T. Mali, MD, PhD
  Professor of Radiology, University Medical Center Utrecht

Annex **D**

Abreviations and terminology

Only in Dutch in the Dutch report.

Annex **E**

Results of population based screening for breast cancer in the Netherlands

See Table 2 on next page.

Table 2 Key figures population based screening for breast cancer. The Netherlands 1998-2012[a]

	1998-2007	2012
Number of women invited	10,318,763	1,266,559
Number of women screened	8,282,990	1,007,966
first screening (%)	47	11
subsequent screening (%) <2,5 jaar	51	85
subsequent screening (%) ≥2,5 jaar	1.6	4
Participation (%)	80.2	79.6
Reinvitations 'on time' = within 24 ± 2 months (%)	75.6	75.0
Referrals	113,424	23,681
Referral rate per 1,000 women participating	13.7	23.5
False positive results per 1,000 women participating	8.3	17.2
after non-invasive diagnostics per 1,000 women participating	4.8	10.6
after invasive diagnostics per 1,000 women participating	3.1	5.3
Breast cancer detected by screening	41,288	6,301
Breast cancer detected per 1,000 women participating	5.0	6.3
DCIS (%)	14.4	20.1
Invasive (%)	84.0	76.7
Unknown morphology of the tumour (%)	1.6	3.2
PPV of referrals (%)	36	27
Number of interval carcinoma 2004-2009	11,855	-
Interval carcinoma per 1,000 women screened	2.3	-
Programme sensitivity 2004-2009 (%)	71.4	-
Specificity 2004-2009 (%)	98.9	-
Total costs (million € per year)	40.0	64.6
Costs per screening (€)	45.90	64.05
Breast cancer mortality per 100,000 women 50-74 years (ESR)	75.8	61.8

PPV= positive predictive value; DCIS=ductal carcinoma in situ.

[a] LETB. National evaluation of breast cancer screening in the Netherlands 1990-2011/2012. LETB XIII. Rotterdam/Nijmegen: Erasmus MC/Radboudumc (in Dutch; tables and graphics in Dutch and English); 2014.