THE EFFECTS AND COSTS OF BREAST CANCER SCREENING

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Koning, Henricus Johannes de

The effects and costs of breast cancer screening /
Henricus Johannes de Koning. - [S.l. : s.n.]. - Fig., tab.
Proefschrift Rotterdam. - Met lit. opg.
ISBN 90-9006070-7
Trefw.: borstkanker / bevolkingsonderzoek ; Nederland.

Cover: torso by Philippe Merloz
Photography: Anton van der Riet Fotostudio, Den Haag
Printer: Pasmans offsetdrukkerij B.V., Den Haag
Layout: Arry de Bruyn and secretariat Dept. of Public Health

THE EFFECTS AND COSTS OF BREAST CANCER SCREENING

Effecten en kosten van bevolkingsonderzoek naar borstkanker

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam

op gezag van de rector magnificus
Prof.Dr. C.J. Rijnvos
en volgens besluit van het College van Dekanen.

De openbare verdediging zal plaatsvinden op
woensdag 9 juni 1993 om 16.00 uur

door

Henricus Johannes de Koning

geboren te 's-Gravenhage

Promotiecommissie

Promotores:	Prof.Dr. P.J. van der Maas
	Prof.Dr. J.A. van Dongen
Overige leden:	Prof.Dr. F. de Waard
	Dr. T.A.W. Splinter

Omdat ik weet dat je trots op me zou zijn geweest.

'Hemeldonderbliksem!' riep de hoogleraar uit. 'Hier heeft iets schrikkelijks plaats gegrepen! Wat hebt u nu gemaakt?'
'Het is niets', hernam de vreemdeling geruststellend. 'Ra-ra, de kwintenstraler heeft de samenstelling van uw tuigje veranderd. En nu gaat Ep-lep, de waarnemer, even uw kukel meten.'
Zo sprekende beklom hij met zijn metgezel de weke massa. De laatste tilde de geleerde de hoed van de schedel en de eerste zette hem een soort stetoscoop op de kruin.
'Geen plus', verklaarde hij, na een ogenblik aandachtig geluisterd te hebben. 'Een min kukel. Dank u!'
Met deze woorden borg hij teleurgesteld zijn instrumentje op en de beide ventjes verdwenen met een korte groet.

Uit: Het kukel. In: Als je begrijpt wat ik bedoel. Marten Toonder, Amsterdam, 1967

Contents

1 **Introduction**
Introduction 1
Mass screening 2
Mass screening for breast cancer 3
This thesis 5

Part I Reduction in mortality and prevention of morbidity

2 **Effectiveness**
Introduction 11
An assessment of the main screening trials 11
Expectation of the Dutch screening programme 17
Discussion 17
Conclusions 20

3 **Advanced breast cancer and its prevention by screening**
Introduction 21
Patients and methods 23
Results 25
Discussion 31

Part II Impact on quality of life and on health care

4 **Impact on clinical medicine**
Introduction 37
Material and methods 38
Results 41
Discussion 46

5 **Impact on quality-adjusted life-years**
- Introduction 51
- The consequences of breast cancer (screening) and value assignment; a literature review 53
- Material and methods 54
- Results 57
- Discussion 64

6 **Changes in use of breast conserving therapy in years 1983-2000**
- Introduction 67
- Material and methods 68
- Results 71
- Discussion 75

Part III Cost-effectiveness

7 **The cost-effectiveness**
- Introduction 81
- Material and methods 82
- Results 84
- Discussion 90

8 **Cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors**
- Introduction 95
- Material and methods 96
- Results 98
- Discussion 107

9 **Discussion**
- Introduction 111
- Update screening trials 111
- Priorities for future research 113
- Decisions for Europe 120
- Spin-offs to other research 121
- Politics 121

10 **Conclusions** 123

Literature 129
Summary 139
Samenvatting 143
Acknowledgements 147
About the author 149
Appendices

Publications

Chapters 2-8 are based on the following (published or submitted) articles:

2 Effectiveness of breast cancer screening; reduction of mortality in the Netherlands and other countries (in Dutch); NTvG 1990a;134:2240-2245[1)]
HJ de Koning, R Boer, PJ van der Maas, BM van Ineveld, HJA Collette and JHCL Hendriks.

3 Advanced breast cancer and its prevention by screening; Br J Cancer 1992a;65:950-955[2)]
HJ de Koning, BM van Ineveld, JCJM de Haes, GJ van Oortmarssen, JGM Klijn and PJ van der Maas.

4 Breast cancer screening: its impact on clinical medicine; Br J Cancer 1990c;61:292-297[2)]
HJ de Koning, GJ van Oortmarssen, BM van Ineveld and PJ van der Maas.

5 The impact of a breast cancer screening programme on quality-adjusted life-years; Int J Cancer 1991;49:538-544[3)]
JCJM de Haes, HJ de Koning, GJ van Oortmarssen, HME van Agt, AE de Bruyn and PJ van der Maas.

6 Changes in use of breast conserving therapy in years 1983-2000; submitted
HJ de Koning, JA van Dongen and PJ van der Maas.

7 The cost-effectiveness of breast cancer screening; Int J Cancer 1989; 43:1055-1060[3)]
PJ van der Maas, HJ de Koning, BM van Ineveld, GJ van Oortmarssen, JDF Habbema, JThN Lubbe, AT Geerts, HJA Collette, ALM Verbeek, JHCL Hendriks and JJ Rombach.

8 Breast cancer screening and cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors; Int J Cancer 1991;49:531-537[3)]
HJ de Koning, BM van Ineveld, GJ van Oortmarssen, JCJM de Haes, HJA Collette, JHCL Hendriks and PJ van der Maas.

1 Introduction

"There is a fallacy in your analysis;
every woman will eventually die somehow."

Introduction

Breast cancer is the most common form of cancer among women in many developed countries. Breast cancer is generally diagnosed on the basis of symptoms or signs. The woman may have felt a lump or change in the breast. Sometimes the finding is a coincidence when the woman is examined for other reasons. In recent years, breast cancer is also being diagnosed in women who have had mammography at their own request. In the Netherlands, 7,968 new cases of invasive breast cancer and 247 ductal carcinomas in situ had been registered in 1989 (Netherlands Cancer Registry, 1992). In the same year the Central Bureau of Statistics registered 3365 women to have died from the disease. Breast cancer accounts for more than 5% of all deaths among women and for more than 21% of all cancer deaths among women in this country (CBS, 1991). With these statistics, the Netherlands is one of the countries with the highest age-standardized breast cancer mortality rates in the world.

Tumours, generally diagnosed with a diameter of 2 cm or more, are treated either by surgery, radiotherapy, adjuvant systemic treatment or a combination of these modalities. Breast conserving therapy or mastectomy, possibly followed by radiotherapy, are treatments with curative intent. In addition, adjuvant systemic treatment has shown to improve survival, especially after longer follow-up, which may indicate this treatment to be effective against small and/or slow growing micro-metastases present at the time of primary diagnosis. The main prognostic factors in operable breast cancer patients are the presence or absence of axillary lymph node metastases, the number of nodes involved, oestrogen receptor status and size of the primary tumour. In case of metastatic disease, cure is not possible anymore.

This brief description of the disease breast cancer underscores the need for either primary prevention, more effective treatment or early detection of the disease. Mass screening for cancer is meant to detect malignant abnormalities at a relatively early stage, in which the earlier treatment is assumed to lead to a better prognosis, compared to treatment at a later stage.

Consequently, less women will develop advanced disease. If so, the probability of dying from the disease is lowered for persons participating in a screening programme. Several mammographic screening trials have shown this to be the case for women aged 50 years and over.

Also realizing the suffering when breast cancer metastasizes, major goals in present day health care and health care research should be to decrease breast cancer morbidity and mortality. Some breakthrough in the near future can be expected from adjuvant systemic treatment for node positive patients (Early Breast Cancer Trialists' Collaborative Group, 1992). As long as there is no highly effective treatment for advanced disease and since most of the known risk factors for breast cancer are either impossible or culturally unacceptable to modify (Harris et al., 1992), secondary prevention seems, although only the third best solution, worth investing in.

Mass screening

Screening involves the application of a screening test to individuals in order to decide whether a certain disease might be present, at a time that there are no clinical signs or symptoms. Abnormal cells shown in a PAP smear of the cervix uteri, for instance, or microcalcifications on a mammogram may thus indicate early lesions of cancer that are screen-detectable but not yet clinically apparent. Mass screening means applying such a screening test to a large proportion of the seemingly healthy population. A test with a high sensitivity will detect the vast majority of the screen-detectable lesions.

Whether screening for cancer will eventually lead to a mortality reduction for that type of cancer, depends basically on 3 steps: the screening should result in earlier diagnoses compared to no screening; the earlier diagnosis should reflect an improvement in prognosis; and other causes of death should not outweigh the disease-specific cause of death (as is the case in older age groups). Earlier detection does in reality mean more efficacious treatment if the risk that the tumour has already disseminated is reduced. If more patients are detected before the tumour has had the opportunity to metastasize, cancer mortality reduction is to be expected.

In some cancers like lung cancer, symptoms or signs evolve relatively late in the natural history of disease and are often already related to metastatic disease. A small advancing of the detection by screening which seems possible from trials (Walter et al., 1992) might then not be very efficacious since dissemination has already taken place. Only techniques that would substantially advance the diagnosis could lead to significant mortality reductions for this disease.

At an individual level one can never tell which person will benefit and which will not. An effective and acceptable screening programme is to be defined as one that, on average, will produce a significant reduction of the

health problem for a group of individuals to which it is intended. At the same time, it should not cause too much harm to those who will not gain from participating.

Mass screening for breast cancer

In breast cancer it has been shown that the majority of lesions can be detected much earlier by mammography than by palpation. The modern screening technique of mammography has enlarged the detectable preclinical period (of course without altering the natural history of disease). In theory, the improvement in prognosis for women with mammographically screen-detected breast cancer might thus be large, if the period that the diagnosis is brought forward is relatively long and/or when survival is strongly time-dependent (see figure 1.1; the maximum interval Q-S).

Studies have shown that for breast cancer, the size of the primary tumour at initial treatment and therefore the time of detection correlates strongly with the probability of metastatic dissemination (Atkinson et al., 1986). The greater the clinical volume, the higher the proportion of metastases appearing later during the course of the disease. More than 55% of the patients with primary tumours between 3.5 and 4.5 cm in diameter developed distant metastases in the period of 25 years following primary treatment, compared to only 25% of women with tumours between 1 and 2.5 cm (Koscielny et al., 1984; Tubiana and Koscielny, 1991).

Also more recent studies, both in experimental research and in screening trials, have proven that this theoretical concept of screening does apply to breast cancer. Counting of microvessels in 103 primary breast cancers showed significantly higher vascularisation in tumours than in normal tissue and it was significantly associated with node metastases. The results suggested that angiogenesis is closely linked to metastasis, that it is acquired at a critical density of vessels, and that this process occurs as tumours enlarge (or become poorly differentiated). Vascular counts correlated with early death (Horak et al., 1992).

The detection of cancers before a critical number of blood vessels has been induced could then lead to a cancer mortality reduction. The proportion of screen-detected cancers smaller than 10 mm in diameter, and the fact that these are associated with 12-year survival rates of about 95% (not corrected for lead time) in Swedish screening trials gives evidence on the efficacy of breast cancer screening (Tabár et al., 1992a).

However, it is not only point Q, the moment the tumour is first detectable, that may vary with the disease and the screening test applied. Point S, the moment of clinical diagnosis may also vary. The published trend in earlier diagnosis in different countries in recent years (Bennett et al., 1990) or in

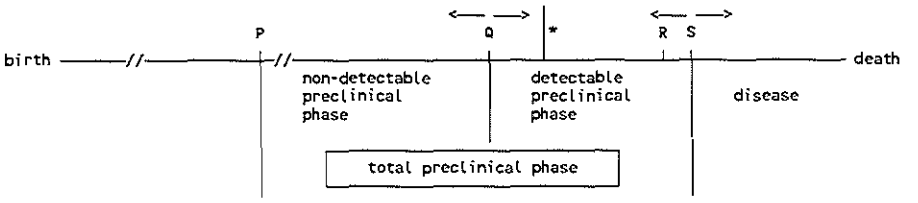

Figure 1.1 Concept of early detection of disease (adapted from Vandenbroucke en Hofman, 1990).
P = first cancer cell, Q = tumour first detectable, R = first symptoms, S = clinical diagnosis, * = screening
Note: the time axis is not linear (exponential growth rates)

improved survival for breast cancer patients (Adami et al., 1986) may, at least partly, be a result of higher awareness of women and physicians of breast problems, resulting in an earlier diagnosis, irrespective of screening. In one region of the Netherlands too, a striking increase of clinically diagnosed early stages of breast cancer has been reported (Coebergh et al., 1990).

The concept of screening for cancer has been incorporated in several mathematical or simulation models used for the evaluation of screening, in order to take account of a number of complexities involved such as the natural history of disease, current knowledge of the epidemiology and the characteristics of screening. For breast cancer, the computer simulation package MISCAN has been developed at the Department of Public Health of the Erasmus Universiteit Rotterdam. The breast cancer model was originally designed as a 2-stage disease process in which assumptions on the duration of the preclinical stages and the sensitivity of palpation and mammography were tested using data from the first American breast cancer screening trial, the Health Insurance Plan of Greater New York-study (Habbema et al., 1986; van Oortmarssen et al., 1990a). The MISCAN breast cancer model has been the underlying model for part of the research presented in this thesis.

Figure 1.2 shows the current structure in the model of the natural history of breast cancer and the screening. Basically, we distinguish invasive breast cancer to be clinically diagnosed in either of 3 categories (based on tumour size). Each stage has its own preclinical period with of course on average longer preclinical periods for the larger tumours. In the present model, screen-detectable ductal carcinoma in situ (DCIS) has been assumed to be the precursor of some of the invasive pre-clinical cancers. Although biologically DCIS might be the precursor of almost all invasive breast cancers, probably only a small part of all these precursor-lesions will be screen-detectable ductal carcinoma in situ.

If applying screening to a population, part of the detectable lesions will become screen-detected cancers whereas the number of clinically diagnosed cancers becomes smaller. The shift from diagnosing and treating relatively

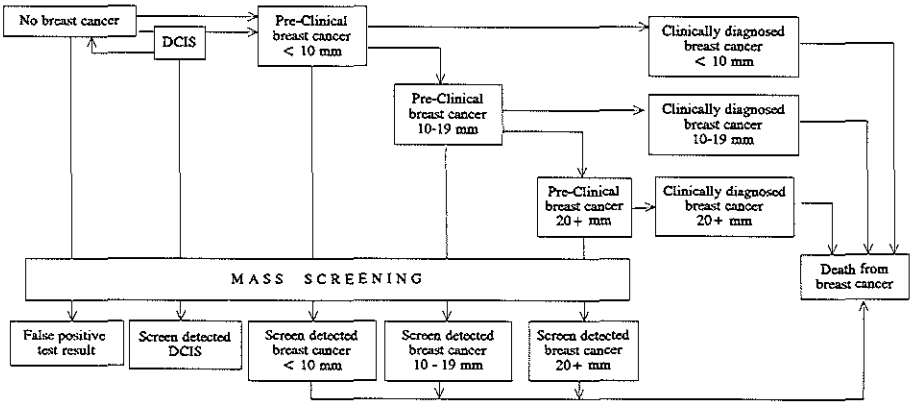

Figure 1.2 Structure of the disease model for breast cancer and the stages (tumour diameter categories) used in the model. The possible courses of the disease are indicated. The state 'death from other causes', which can be attained from all states is not shown (adapted from van Oortmarssen et al., 1990b).

large clinical cancers towards earlier screen-detected stages is reflected in a decrease in breast cancer mortality, as shown in breast cancer screening trials. More details about these model assumptions are published elsewhere (van Oortmarssen et al., 1990b) and will be described in a separate thesis.

This thesis

In 1986, the Dutch Ministry of Welfare, Health and Cultural Affairs asked a research group to investigate the expected effect of breast cancer screening on mortality and possibly morbidity, if implemented in the Netherlands. The research group consisted of members from 3 centres, the Dept. of Public Health (Erasmus Universiteit Rotterdam), the Dept. of Public Health and Epidemiology/Preventicon (Rijksuniversiteit Utrecht) and the Dept. of Epidemiology/Radiology (Katholieke Universiteit Nijmegen). Data from the two Dutch breast cancer screening projects, the DOM-projects in Utrecht (Rombach, 1980; Collette et al., 1984; de Waard et al., 1984) and the Nijmegen project (Hendriks, 1982; Verbeek et al., 1984; Peeters, 1989) which started in 1974/1975, was made available in order to first establish more precise estimates for the parameters in the above mentioned model for the disease process of breast cancer and screening in the Netherlands. Together with research on the cost, one of the first Dutch cost-effectiveness analyses in health care was started.

The research presented in this thesis in parts I, II and III addresses the following 3 main questions:

1 The first question in breast cancer screening concerns the reduction in breast cancer mortality and prevention of advanced disease, the main favourable and envisaged effects of nationwide breast cancer screening (chapters 2 and 3);

2 If screening is shown to be an effective means of reducing breast cancer mortality, could the decision to implement screening be otherwise when taking into account the impact of a screening programme on quality of life and on health care (chapters 4, 5 and 6);

3 The balance between unfavourable effects and favourable effects other than mortality reduction, and the relationship between effects and costs are finally additional guides for decisions on screening policy, evaluation and/or further research and follow up (chapters 7 and 8).

- Chapter 2 reviews the evidence on breast cancer mortality reduction as reported in several screening trials,
- chapter 3 analyzes the period of advanced breast cancer, and the effect of screening on the number of and care for women with advanced disease,
- chapter 4 explains the approach of integrating assessment and primary treatment changes in screening analyses,
- chapter 5 analyzes the expected impact of a screening programme on quality of life,
- chapter 6 shows the trend in health care, exemplified in detail for breast conserving therapy,
- chapter 7 analyzes the influence of different screening intervals on cost-effectiveness,
- chapter 8 comprises the cost-effectiveness of breast cancer screening, including the possible impact of uncertain factors,
- the discussion and conclusions on all above questions are presented in chapters 9 and 10.

With the research described in this thesis, both the concept of screening and the underlying model earlier described have been extended to include the impact on almost all aspects after or beyond the screening examination itself: impact on advanced disease late in life, on assessment and primary treatment for women and for health care, on quality of life, impact on financial cost and impact on the situation of the disease in a population outside the invited or screened group.

The organisation and structure of breast cancer screening in the Netherlands is presented in several reports of advisory committees (Dutch National Council for Public Health, 1987; Health Insurance Executive Board, 1988): the population registries are used to identify and invite all eligible women. Screening takes place in special screening units. In case of suspected abnormalities on the mammogram, women are referred to their general practitioners, who in turn will refer them to an out-patient department for further assessment. The estimation of the cost of the screening examination is published in more detail elsewhere (de Koning et al., 1990a; van Ineveld, 1992; NETB, 1992) and will form part of another thesis.

In 1988, the possible future Dutch screening policy was advised to be a 2-yearly mammographic screening programme for women aged 50 and over, partly based on the research described in chapter 7. The definitive decision on the introduction of breast cancer screening for all women aged 50-69 in the Netherlands was taken in 1993.

The quotation on page 1 (a comment at one of my presentations on breast cancer screening) suggests all preventive actions to be pointless. That every woman will eventually die is a fact. Research presented in this thesis hopefully shows women that they might have some influence in the way they live and time they die.

PART I

2 Effectiveness

Introduction

There are insufficient means of curative therapy for breast cancer that is detected at an advanced stage. The prognosis for treatment of a malignant tumour that is detected at an early stage is generally more favourable than for one that is detected at a later stage. This means that breast cancer screening is a method that should seriously be considered. Breast cancer screening is designed to detect tumours at such an early stage that they can still be treated effectively. The chance of dying of breast cancer is thus reduced for those women who participate in screening.

Various trial projects in which women were invited to participate in mammographic screening, both in the Netherlands and in other countries, have shown that a reduction in breast cancer mortality is indeed achieved. Following American research in the sixties, trials were set up in various countries in the seventies including the Netherlands, Sweden, Italy and the United Kingdom, and in Canada in the eighties. All the results that have been published to date indicate that women aged 50 years and over who participate in screening have a reduced chance of dying of breast cancer.

However the size of this reduction varies considerably from one trial to the next. In this chapter the main outcomes of breast cancer screening of women aged 50 years and over are discussed, as are the factors that influenced these outcomes. Based on this we shall predict the degree to which the mass breast screening programme in the Netherlands may reduce breast cancer mortality in the nineties and subsequent years. We shall also briefly discuss the less favourable results for women aged under 50.

An assessment of the main screening trials

Table 2.1 gives a summary of the characteristics of the most important trials that are described in this chapter.

Chapter based on de Koning et al., NTvG, 1990a.

Table 2.1 Summary of the characteristics of breast cancer screening trials in different countries

characteristics	*trials*							
Place	New York	Kopparberg/ Östergötland	Malmö	Edinburgh	Stockholm	Utrecht	Nijmegen	Florence
year of start	1963	1977/78	1976	1979	1981	1974	1975	1970
total number of women	63,000	133,092	42,283	45,130	60,261	20,555	30,700	24,813
age group (years)	40-64	40-74	45-69	45-64	40-64	50-64	> 35	40-69
randomized	yes	yes	yes	yes	yes	C-C	C-C	C-C

C-C = 'case-control'-studies

The New York trial

In 1963 the randomized Health Insurance Plan (HIP) trial was set up in New York (Shapiro et al., 1988). More than 60,000 women aged 40-65 who were insured under the Health Insurance Plan of Greater New York were randomly divided into two groups. Half the women were invited for mammographic screening and clinical examination (the study population). The other half was not, acting as the control population. The women were screened a total of four times, at yearly intervals.

After 10 years the mortality from breast cancer of breast cancer patients (diagnosed within 5 years after entry) aged 50 and over was 32% lower in the study population than in the control population. An estimate such as this is to some extent dependent on the actual attendance. Sixty-seven per cent of the study population attended the initial examination. Only they were invited for subsequent screens. A fresh evaluation was carried out after a follow-up period of 18 years, when a difference of 21% in mortality from breast cancer was found between the study group and the control group.

The Swedish trials

These were set up towards the end of the seventies. The Kopparberg/Östergötland study (W/E-trial) covered two rural areas in the south-east and mid-west of Sweden (Tabár et al., 1985). Randomisation was carried out by dividing each of nineteen homogenous socio-economic areas into one or more areas with a control group and a study group. In the study group the women aged 50 and over were invited for mammographic screening at intervals of approximately 33 months. The control group was not invited.

After an average follow-up period of 6 years for each woman offered screening, the first results were published showing a significant reduction in mortality from breast cancer of 39% in the study population for women aged 50-74. Data has been published with a further follow-up lasting two respectively 5 years longer, which showed roughly the same level of reduction (Tabár et al., 1989; Tabár et al., 1992b).

Screening trials were also started in a number of Swedish urban areas: Malmö, Stockholm and Göteborg. In Malmö randomisation was at an individual level and screening was carried out at intervals of 1.5 to 2 years. In Stockholm women were invited every 2 years (Frisell et al., 1986), and in Göteborg every 1.5 years (Bjurstam et al., 1987).

The first results from Malmö were published in 1988 (Andersson et al., 1988). They had decided beforehand at what stage they would start to analyze the initial data (Andersson et al., 1989b). The results after an average study period of more than 8 years were disappointing: in women aged 55 and over the mortality reduction achieved in the study population was 21%, which was lower than in the W/E-trial and not statistically significant. The provisional figures up to and including 1988 were consistent with the published data for this age group (Andersson, 1989a). The 95% confidence intervals of the

estimated impact on breast cancer mortality of the W/E- and Malmö trials are wide and largely overlap, so that the difference between the two could be coincidental. However an important question is whether there are differences between the two trials which might explain the differences between the point estimates that were published. These two trials are summarised in Table 2.2, together with the Edinburgh trial (Roberts et al., 1990). The Table places the emphasis on relevant features of the studies and the results achieved in the age group from 50 to 70 years; this is comparable with the target group in the Netherlands.

Table 2.2 Factors that influence the impact on mortality reduction in 3 randomized European breast cancer screening trials: Kopparberg/Östergötland (W/E), Malmö and Edinburgh

Characteristics/results	*Kopparberg/ Östergötland*	*Malmö*	*Edinburgh*
age at randomisation	50-69	55-69	50-64
size study group	46,897	13,107	17,313
size control group	33,074	13,113	16,094
randomisation	19 socio-economic homogenous blocs	individual	84 general practices
region	rural	city	city
attendance 1st round (%)	87	71	60
screening interval (mnths)	33	18-24	12-24*
mean evaluation period (yr)	7.9	8.8	7
% women screened outside programme total follow-up period (in control group)**	13	24	unknown
breast cancer mortality reduction (see age group) (%)	39	21	20
relative risk (95%-confidence interval)	0.61 (0.48-0.83)	0.79 (0.59-1.29)	0.80 (0.54-1.17)

* Alternate years clinical examination and mammography
** Age specific figures not known in all trials
References: Tabár et al., 1989; Andersson et al., 1988; Roberts et al., 1990

One notable difference is that the W/E-trial was set up on a much larger scale than the Malmö trial. In Malmö randomisation was at an individual level; in the W/E-trial it was at the level of small communities. In principle the former method produces more accurate estimates for a given sample size.

Attendance in Malmö was decidedly lower, which would tend to reduce the expected effect. On the other hand the women were screened more frequently, which would increase the expected effect on mortality reduction. Model calculations indicate that these two factors approximately cancel one another out (de Koning et al., 1991).

In Malmö, 24% of the women who were not invited for screening did in fact undergo mammographic screening (in the study period). In the rural W/E-areas only 13% of the control population had a mammogram made. Preventive mammography carried out independently of the screening programme would reduce the relative effect of screening. Finally, there might be a difference in the incidence or stage distribution of breast cancer between urban and rural areas.

Data currently available in the literature does not demonstrate that the combination of these factors results in a significant difference between the 2 Swedish trials. Important data about a possible difference in the quality of the mammograms in the two trials is not available. Therefore an average of the published estimates of the impact on breast cancer mortality was determined, weighted according to the confidence intervals. This average is 32% with a 95% confidence interval of 14-46%. In other words in these Swedish trials breast cancer mortality is on average 32% lower in the study groups than in the control groups. The combined analysis of all Swedish trials has recently been presented (Nyström, 1993) and will be discussed further in chapter 9.

The other trials also show a reduction in mortality from breast cancer in women aged 50 and over. However the way these trials were set up means they are often less suited to giving an unbiased estimate of the size of the reduction than are the Swedish trials. For example, where there was no randomisation, it cannot be ruled out that part of the effect on mortality might be due to selection.

The United Kingdom trial

In the United Kingdom, women were invited for mammographic screening in two areas, they were encouraged to carry out breast self-examination in a further two areas, and there were four control areas (UK Trial, 1988). The mortality reduction in women aged 45-64 years as a result of two-yearly mammographic screening (with clinical examination in alternate years) was estimated at 20% after an average follow-up of 6.5 years. It turned out that there were wide socio-economic differences between the areas, for which corrections were made.

The data from Edinburgh has also separately been published, one of the two areas in the United Kingdom where mammographic screening was carried out. Although the trial in the United Kingdom was not randomized, some element of randomisation was added at the beginning in Edinburgh in 1979: 84 general practices were randomly distributed into practices where women were invited for screening and those where they were not (see Table 2.2).

After an average follow-up of 7 years a statistically non-significant mortality reduction of 20% was found for women aged 50-64 (Roberts et al., 1990). There are clear factors which detracted from the effect. Attendance was low: 60% at the first examination, falling to 52% by the third. In addition, as the authors themselves write, the standard of mammography was fairly poor, particularly at the beginning of the trial. Only in 1982 did they start working with improved equipment. The stage distribution of the tumours detected by screening was worse than those of the Swedish W/E-trial. Consequently the authors concluded that a longer period of evaluation is necessary.

The Dutch and Italian trials

The Dutch and Italian screening trials were not randomized and did not use a control population that had been determined beforehand. They were evaluated by means of 'case-control' studies (Collette et al., 1984; Verbeek et al., 1984; Palli et al., 1986). In Utrecht and Nijmegen the screening histories of women who had died of breast cancer since the start of screening (case) were compared with those of women born in the same year, living in Utrecht and Nijmegen, who had not (control). The estimated impact on breast cancer mortality was represented as the relative risk of dying of breast cancer in women who had at some stage been screened compared with those who had never been screened. The relative risk for women aged 50-64 was found to be 0.35 in Utrecht and 0.30 in Nijmegen. In other words the chance of women who participated in the trial dying of breast cancer was almost one-third of that of women who did not.

These estimates will be higher than those of randomized studies because the effect is only estimated for participants. The estimates in randomized studies take account of the total study population (which is eligible for screening), and this includes women who do not attend screening. According to these case-control studies, and assuming an attendance of 70% in the Netherlands, breast cancer mortality should fall approximately 47% in the group that was offered screening, if there is no selection. The possibility of selection cannot be ruled out in case-control studies. Although this has not been demonstrated in the Netherlands, there may have been some measure of self-selection. This could mean that the women who participate in screening are those who are healthy and (or) have a favourable chance of survival. Despite the possible methodological limitations of case-control studies, they can well be used to provide supplementary evidence (Day, 1989a).

Expectation of the Dutch screening programme

It is hard to apply the projected reduction in mortality in the American HIP trial to the current situation in the Netherlands. The quality of the X-ray images (film without a screen) was much worse in the sixties than that of the present mammograms. This means that future screening might (in theory) produce even better results. On the other hand the prognosis for breast cancer was worse in the US in the sixties than in the Netherlands now, because breast cancer was generally diagnosed at a later stage then. This could lead to a possible overestimate of the reduction in mortality that could be achieved in the Netherlands now.

The Swedish trials are more comparable with the Dutch situation. The quality of the mammograms, the organisation of screening and the health care services in general are very similar. As far as breast cancer is concerned the clinical picture in Sweden is comparable with that in the Netherlands, in terms of stage distribution at diagnosis and survival rates. The expected mortality reduction for the Dutch population can be calculated, based on the assumptions in the two Swedish trials about the improved prognoses achieved for specific age groups (Figure 2.1). Assuming a programme of 2-yearly screening for women aged 50-69 started in 1990, with an average attendance of 70% and a high standard of mammography, a reduction in mortality from breast cancer of approximately 16% should be possible for the total female population (all ages) around the year 2015. Approximately 700 fewer women would die of breast cancer in that year than if screening was not carried out (see also chapter 8 for description of method).

Discussion

Misunderstandings

After the favourable W/E-trial results the lower mortality reduction in Malmö and the United Kingdom produced a stream of critical reactions on the question whether or not breast cancer screening is in fact worthwhile. Some reactions reminded us again how important the education of women is going to be, how important the quality of a screening programme is, and what can and what cannot be expected of a national screening programme (Warren, 1988; Roberts, 1989). Others often lost sight of the overall picture (Skrabanek, 1988; Schmidt, 1990).

In this discussion, we refer to three important mechanisms which are inherent in any programme of screening, but which often give rise to misunderstandings. At the same time we indicate to what extent these mechanisms are implicitly present in the results that have been published up to now, and do not change the conclusions derived from these studies.

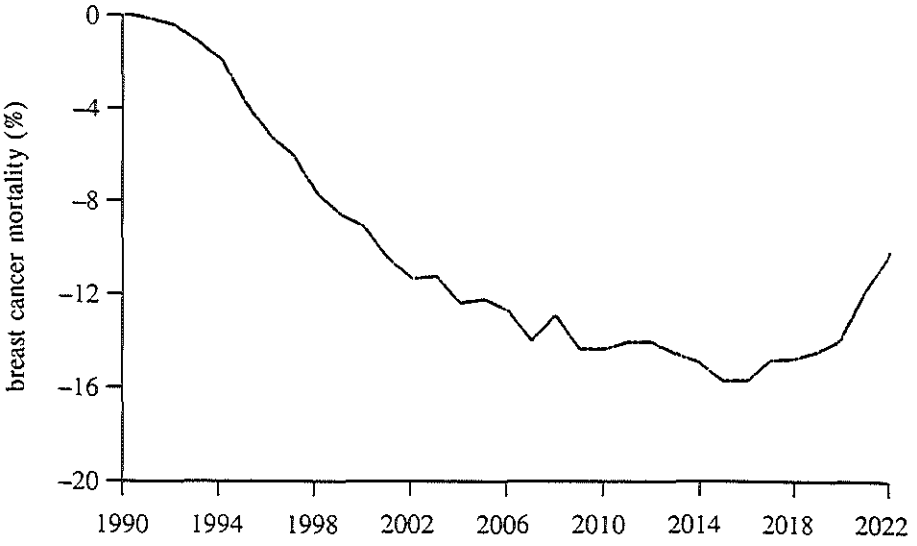

Figure 2.1 Expected reduction in mortality from breast cancer in the Netherlands due to mass screening (expected duration: 1990-2016), indicated as a percentage of breast cancer mortality if screening is not carried out. The improvement in prognosis of tumours detected by screening, based on the results of the Swedish trials in Kopparberg/Östergötland and Malmö.

In many cases screening results in breast cancer being diagnosed at an earlier stage than it otherwise would. The time gain in arriving at the diagnosis is known as the 'lead time'. The results of the trials do not generally include survival rate curves. Part of the difference shown in such a figure between a group that has been screened and one that has not is due solely to the earlier diagnosis. This then requires correction (Collette et al., 1992). In randomized and case-control studies the number of deaths as a result of breast cancer in the total study population is compared with that in the control population. The lead time is not a complicating factor in these comparisons.Where screening takes place at a certain interval, the chance that slow growing tumours will be detected is greater than the chance that fast growing ones will. This is known as 'length biased sampling'. It arises because there is only a short period during which fast growing tumours can be detected using mammography before they are diagnosed clinically. Thus they are less often detected by screens that are carried out at, for example, two-yearly intervals. The chance of detecting a number of tumours with a relatively longer preclinical phase and possibly a frequently low level of malignancy is particularly great at the first examination (Kallioniemi et al., 1988). To what extent this plays a part in breast cancer screening is not yet clear (Holmberg et al., 1986b; Tabár et al., 1987b). Be that as it may, length bias does not distort the effect of either randomized studies or case-control studies. The total number

of deaths as a result of breast cancer in the study population (thus including tumours that were not detected during screening) is counted.

Finally there has been debate about the effects of participating in screening on the chance a woman will die of other causes (Tijmstra, 1988). The question is raised whether mortality due to other causes increases where screening is carried out. A programme results in breast cancer being diagnosed in more women, particularly at the start of the programme, and because of lead time and improved prognosis, the total number of woman-years in breast cancer patients in the study population would be considerably greater than that of the control population. This means that, even if each woman has the same chance of dying, the expected number of deaths among breast cancer patients due to other causes would be greater in the study population than in the control population. It is therefore important to look at woman-years of the study population in comparison with the control population when comparing the mortality due to other causes. Neither the W/E-trial nor the Malmö trial showed a clear difference between the pattern of mortality from other causes in women who participated in screening and in those who did not (Andersson et al., 1988; Tabár et al., 1989).

Screening of women aged under 50

Whether screening of younger women also causes a reduction in mortality from breast cancer is not as yet clear. There are important differences between screening at a younger age and screening at an older age. These are to do both with the disease itself and with the screening test (active glandular tissue complicates assessment of the mammograms; probably higher growth rates). Breast cancer is less frequent in younger women. It would require a very large number of participants to demonstrate a significant mortality reduction for this group. There are some indications that the prognosis for women aged 40-49 is relatively favourable, in other words that they will survive longer (Adami et al., 1986; Høst and Lund, 1986), which means that the absolute benefit from screening in terms of mortality reduction will be limited.

The American HIP trial showed that screening had a favourable impact for women under 50 after 10 years of follow-up (Shapiro et al., 1982; Chu et al., 1988). The differences described above between America and the Netherlands in terms of the quality of mammography and the prognosis of breast cancer make it hard to make any firm pronouncements on screening of this group in the Netherlands. After a follow-up of 8 years, the publication from the W/E-trial described a small, non-significant reduction in breast cancer mortality (8%) in women aged 40-49 years. However the confidence intervals are very wide (Tabár et al., 1989). No other study demonstrates a favourable effect. Based on the publications to date, the effectiveness of screening in terms of a reduction in the mortality from breast cancer in women under 50 seems at best very limited at this moment in time. Swedish new data are being discussed in chapter 9.

Conclusions

All screening trials, both in the Netherlands and in other countries, confirm the expectation that introducing breast cancer screening for women aged 50 years and over will lead to a reduction in breast cancer mortality. The precise extent of the reduction will depend on various factors. A high attendance rate and a high quality of mammographic screening will increase the reduction. Attendance can be influenced by informing and educating people about the screening programme. Quality depends on effective quality control and proper training of the staff involved in screening. On the other hand, the benefits will be comparatively less if the prognosis of breast cancer has already improved, either because women are visiting their doctors earlier or because they decide to have mammograms made on their own initiative, without waiting for screening. The effectiveness of screening for women under 50 is still debatable. The results that have been published to date do not yet offer good reasons for inviting women under 50 to participate in a national screening programme.

3 Advanced breast cancer and its prevention by screening

Introduction

Population screening for breast cancer has now been introduced in several countries. Trials have shown that it is possible to reduce breast cancer mortality by mammographic screening of women of 50 years and older (Wald et al., 1991). The seemingly less favourable mortality reductions achieved in two trials (Andersson et al., 1988; Roberts et al., 1990) resulted in articles that emphasize possible disadvantages of breast cancer screening, and attention has been focused on the morbidity generated (Roberts, 1989; Dixon and John, 1992). However, the impact of screening on the prevention of advanced disease and its morbidity is generally neglected. In contrast to the research in costs and possible adverse effects of screening (Ellman et al., 1989), the impact on the prevention of morbidity had not yet been analyzed in detail nor quantified.

Little is known about the actual amount of care and treatment given to all women with advanced breast cancer, because of the wide variety, which makes it difficult to estimate the actual impact of screening on this group. Most published case series relate to disease-free or post-recurrence survival only. For this reason, we conducted a detailed study on systemic and palliative treatments using breast cancer patient files and using national sources from the Netherlands. It provides an overview of advanced disease and its course and identifies types and durations of diagnostic and treatment procedures, amount of care involved and the associated medical cost. At the same time, we calculated quality of life estimates for the different forms of palliative care. There has been extensive investigation into the psychological sequelae of early breast cancer and its treatment, but few data concerning patients with advanced disease (Hopwood et al., 1991). Population effects of screening with respect to the prevention of advanced disease are illustrated by simulating two different screening policies and calculating its impact on three aspects: morbidity (by a quality of life-index), consequences for health care and for medical cost.

Chapter based on de Koning et al., Br J Cancer, 1992a

Table 3.1 Main characteristics of patients (a), primary tumour (b), primary treatment (c), and first recurrence (d) from 68 women who died from breast cancer in 1985-1989. Three hospitals

(a)	Number of women from hospital		1	21
			2	27
			3	20
	Mean age at first diagnosis of breast cancer			61 (32-76)
(b)	Postoperative TN	pT_1, pT_2		28%, 56%
		pT_3, pT_4, pT_x		12%, 1%, 3%
		pN_o, pN_+, pN_x		25%, 72%, 3%
	Oestrogen receptor	positive ($\geq$ 10 fmol/mg^{-1})		60%
		negative		24%
		unknown		16%
(c)	Mastectomy			79%
	Breast conserving therapy			21%
	Postoperative radiotherapy			93%
	Adjuvant chemotherapy			10%
	Adjuvant hormonal treatment			3%
(d)	Locoregional recurrence			15%
	Distant metastases:	bone		44%
		lung		23%
		liver		9%
		other		9%
	Disease free interval	(mean)		24 months (1.5-81)
		(median)		21 months

Patients and methods

Patient files (both in- and out-patient) were analyzed from three public hospitals in the Netherlands: Antoni van Leeuwenhoek cancer centre (Amsterdam), Dr Daniel den Hoed cancer centre (Rotterdam) and Sint Anna Hospital (Oss). The first two have computerized data on all breast cancer patients diagnosed since 1981. The third is a general hospital, in which similar (non-computerized) data have been collected since 1973. From the two computerized registries, 60 female cases which met following criteria: (a) died from breast cancer in the period 1985-1989, (b) had received primary treatment with curative intent, (c) first diagnosis of breast cancer between the ages 50 and 80, were randomly taken. All 20 available female cases from the smaller hospital which met criteria (a) and (b) were used. All patients had developed distant metastases. A total of 68 patient files (see Table 3.1) were sufficiently documented to follow the full course of patient's illness. Data regarding the actual palliative and/or systemic treatment (in these or other hospitals) was recorded from the date of first diagnosis of advanced disease (distant metastases and/or locoregional recurrence after mastectomy) up to the date of death. This included hospital admissions, duration and dosage of medication, surgery, type and duration of radiotherapy, and diagnostic and pathological procedures, excluding items clearly not related to the disease. Data of one patient with a local recurrence after breast conserving therapy was included from the moment that the (second) recurrence after ablative surgery had been diagnosed.

Literature on quality of life for women with advanced breast cancer was collected as part of a larger study on breast cancer screening and quality of life. A total of 15 phases, including screening, primary treatment, etc. have been investigated (de Haes et al., 1991). Advanced disease was divided into five episodes: those in which respectively hormonal treatment, chemotherapy, palliative radiotherapy or surgery is the main treatment modality, and in a terminal phase. A questionnaire was constructed in which the complaints and symptoms, found in 39 articles were summarized for these five episodes of advanced breast cancer, considering the three dimensions physical, psychological and social. Eighteen public health professionals and 13 breast cancer experts were asked to value each of the described episodes on a visual analogue scale (VAS). The best health state is assumed to have a value of 1, whereas the worst state has a zero value. For quality-adjusted life-year analysis, the utility measure should reflect the subject's willingness to trade off quality against length of life, which is the case when using the so-called 'standard gamble' or 'time trade off' technique. Scores on direct scaling methods such as the VAS are systematically lower, but have been found to be related by a power function to the scores of the 'time trade off' method. Time Trade Off = $1-(1-VAS)^{1.82}$ (Bombardier et al., 1982). Using this function, the utilities for the five different states were computed from the scores of the

direct scaling method (utility = the relative desirability of different states; the relative evaluation of alternative outcomes). The average loss in quality of life for female breast cancer patients with advanced disease was calculated by combining the utilities with the mean durations of the different disease episodes found in the 68 patients.

Calculations of costs have been restricted to direct medical cost and were based on the detailed data from the files. Only the exact number of out-patient visits, if patients did not receive chemotherapy, was estimated (Holli and Hakama, 1989). Actual cost was calculated for radiotherapy. Cost for medical procedures and out-patient visits were based on the most recent Dutch tariffs. Cost of hormonal treatment and chemotherapy was based on (cheapest) retail-prices of the drugs and a fixed amount per prescription and calculated for the exact dosages and durations prescribed. Costs are expressed in US dollars using purchasing-power parities (1990, 1 US dollar = 2 Dfl). Cost of hospital nursing was based on the mean number of in-patient days, as analyzed in our file study, and the mean all-out tariff in the Netherlands of 235 US dollars per day (Dutch Sickness Funds Association, personal communication). Cost for spending a day in a nursing home is 100 US dollars, but 235 US dollars in the terminal phase. Three sources were used to analyze the representativeness of the hospital files. A questionnaire was sent to all 20 Dutch radiotherapy departments, concerning the number of female breast cancer patients that had been treated by radiation for recurrences or distant metastases in the years 1986-1988. Data concerning all hospital admissions in the Netherlands for women with breast cancer during 1985-1988 and data on 80% of the admissions to nursing homes in 1986-1988 (Centre for Health Care Information) was analyzed to obtain independent estimates of the number of in-patient days.

The MISCAN breast cancer model was used to predict the decrease in the number of women with advanced disease and in breast cancer mortality when introducing nationwide screening in the Netherlands (van Oortmarssen et al., 1990b). It is assumed in the model that all women, who had (or would have) died from breast cancer, had or would have been treated for advanced disease. Key parameters of this model were derived from an analysis of all results from the Dutch screening trials (de Koning et al., 1991). The improvement in prognosis after early detection was based on the 32% reduction (weighted average) in breast cancer mortality achieved in the randomized trials in Kopparberg/ Östergötland (Tabár et al., 1989) and in Malmö (Andersson et al., 1988) after 8-9 years for women in the study group aged 50-70. Effects are presented in this chapter for 2-yearly screening for women aged 50-70 (Dutch policy) and for 3-yearly for women aged 50-65 (Vessey, 1991) (attendance rate 70%). Cases of advanced disease that are prevented by screening are assumed to be a random subset of all cases of advanced disease.

Results

Systemic and palliative treatment

Hormonal treatment was the first-line treatment in 57% of all women with advanced disease and chemotherapy in only 16%. Most patients received more than one treatment modality during the period from recurrence till death. Outpatient drug treatment was the most frequently applied modality: 84% of women received hormonal treatment and 69% received chemotherapy during one or more time periods (Table 3.2). Tamoxifen, aminoglutethimide, high dose progestins and CMF (cyclophosphamide, methotrexate, 5-fluorouracil) were used most frequently. If Tamoxifen had been prescribed to a woman, she had taken it during 9.5 months on average. Aminoglutethimide, mostly used in the second or third line, and CMF were prescribed for respectively 6 and 5.5 months on average. In the lower part of the Table, figures are averaged out over all 68 women: drug treatment is the main modality during 18 months.

Radiotherapy, especially effective in reducing pain and invalidity caused by bone metastases, had been given to 59% of the women. Surgery played only a minor part in treatment and had often been undertaken with diagnostic intent. During the period of advanced disease, women remained in hospital for a mean period of one and a half months. The period from first recurrence until death was on average 21.4 months. Figure 3.1 shows the survival curve of all 68 women starting from the month of first recurrence. The three patient groups from the separate centres show a striking resemblance. The mean post-recurrence survival is 19.7 months in hospital 1, 21.7 months in hospital 2 and 22.7 months in hospital 3 (not significantly different; log-rank test).

Quality of life

The experts' assessment of quality of life did not in general differ significantly from that of the other persons (health professionals), both in ranking order and in mean or median value (de Haes et al., 1991). They differed only in their assessment of the impact of palliative hormonal treatment on quality of life (clinicians' rating implied less impact than of non-clinicians)(Table 3.3, b1 versus b2). An ANOVA (analysis of variance) showed no systematic difference between the two groups of respondents, which was the reason for combining the values in all further calculations. Using the questionnaires of 27 respondents, this resulted in (combined) median values on the visual analogue scale between 0.45 for advanced breast cancer in a period with mainly hormonal treatment and 0.17 for the terminal phase (b3).

Each row in Table 3.3 represents the loss in quality of life for a patient with advanced disease in relation to each of the treatment episodes and to the terminal period. The Table shows that a woman in the terminal period has a 71% loss in utility as compared to a healthy woman (column c) and the length of this period is 3 to 4 weeks on average (column d). In the preceding episodes, the utility is calculated to be 34-47% lower compared to a healthy woman

Table 3.2 Synopsis of treatment and amount of health care for women with advanced breast cancer from first recurrence until death (thus excluding primary therapy)

(a)	First measure	hormonal treatment	57%
		radiotherapy	19%
		chemotherapy	16%
		surgery	6%
		none	2%
	At least once	hormonal treatment	84%
		chemotherapy	69%
		radiotherapy	59%
(b)	Number of hospital admissions		2.5
	Total in-patients days	(mean)	45
		(median)	40
	Total days in nursing home		8.3
	Duration of hormonal treatment		14 months
	Duration of chemotherapy		4 months
	Diagnostic procedures	X-rays	22
		isotopic scans	2
		CT	0.8
		pathological procedures	1.9
	Post-recurrence survival period till death (mean)		21.4 months (0.5-70.5)
		(median)	20.3 months

(a) all women, (b) average per woman, based on 68 patient files from three hospitals.

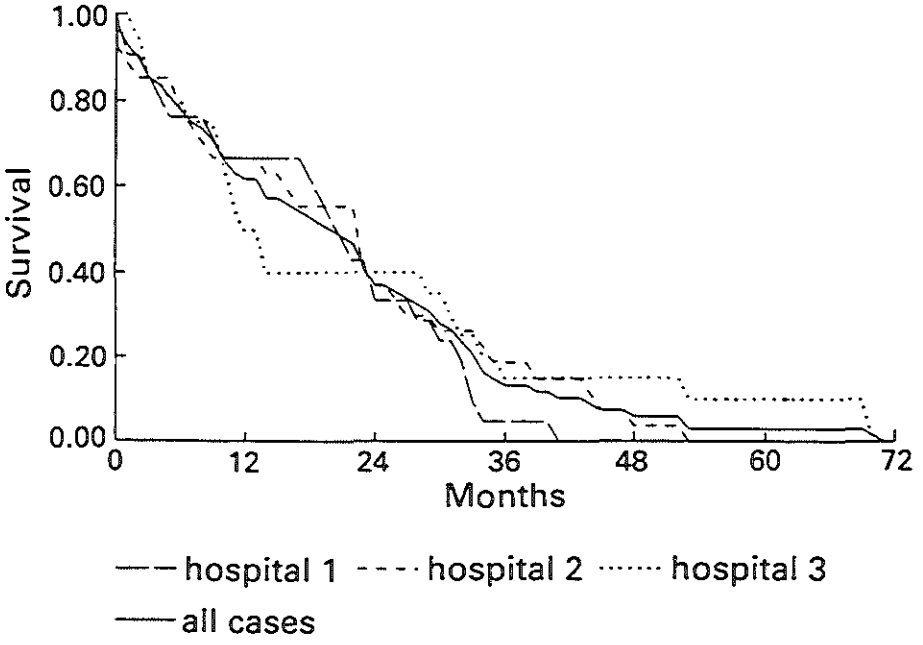

Figure 3.1 Post-recurrence survival of 68 female breast cancer patients. No significant differences (log-rank test) Note: death of breast cancer was a selection criterium in this study.

with a larger loss in quality of life in the following ascending order: hormonal treatment, palliative surgery, radiotherapy and chemotherapy. The relatively long duration of hormonal treatment strongly influences the overall quality of life. A woman with advanced breast cancer is estimated to have a quality of life which is on average 39% lower compared to that of healthy women during 21 months which corresponds to 1.75 - 0.68 = 1.07 quality-adjusted life-year with advanced disease. The lower and upper limits for this loss, based on the extreme mean or median values in either of the respondent groups, are 27% and 45% respectively.

Cost of treatment

Total intramural medical cost from diagnosis of advanced disease until death amounted to an average of 17,100 US dollars per woman, of which 62% is attributable to hospital nursing (Table 3.4). Radiotherapy (9%) was used frequently but mostly for a relatively short period (6.9 sessions per palliative radiation treatment). Diagnostic testing is also costly (10%); e.g. approximately one X-ray was made every month to follow the course of metastases or to find the source of complaints (Table 3.2). The other 19% are cost for out-hospital visits to specialists, cost of drugs and cost for nursing homes. Despite the fact that medication is the most frequently applied treatment modality in advanced disease, it is responsible for only 8% of the total costs.

Table 3.3 Calculation of the impact of advanced breast cancer and its treatment on quality of life

a	*b1*	*b2*	*b3*	*c*	*d*	*e*
Hormonal treatment	0.59	0.37	0.45	0.34	1.16	0.39
Palliative surgery	0.47	0.40	0.41	0.38	0.10*	0.04
Radiotherapy	0.38	0.38	0.38	0.41	0.08	0.03
Chemotherapy	0.34	0.34	0.34	0.47	0.33	0.16
Terminal stage	0.19	0.16	0.17	0.71	0.08	0.06
					___(+)	___(+)
Total period					1.75	0.68

Disease episode considered (a), median values on a visual analogue scale (b1-b3), (1-utility)-measure corresponding to b3 applicable to each period (c), duration of treatment as found in file study in years (d) and the loss in quality of life years during each phase (e = c x d). * Approximation based on in-patient days,
b1 = breast cancer experts, b2 = health professionals, b3 = combined.

Table 3.4 Mean medical cost for women treated for advanced breast cancer based on 68 patient files (extramural cost not included)*

Hospital nursing	10,575.-	62%
Diagnostic procedures	1,700.-	10%
Radiotherapy	1,626.-	9%
Nursing home	1,250.-	7%
Hormonal treatment	760.-	5%
Specialists	640.-	4%
Chemotherapy	550.-	3%
Total per woman	17,100.-	100%

Mean cost in US dollars per woman based on all patients. Mean duration of disease 21.4 months. * Extramural cost per woman: 3860 US dollars (Koopmanschap et al., 1992).

Impact of nationwide screening

With ten 2-yearly invitations for screening to women between ages 50 and 70, total breast cancer mortality is expected to be reduced by 16% in the Dutch population. However, it takes considerable time before this reduction will be reached. Without a screening programme, the total number of women that will die from breast cancer will rise to approximately 4245 in the year 2014. Screening will lower this figure by 655 in that year. As nearly all women with breast cancer in an advanced stage will ultimately die from the disease, the number of new patients for treatment of advanced disease will decrease proportionally to the expected decrease in breast cancer mortality. Table 3.5 summarises the impact of nationwide screening in the Netherlands, especially concerning the prevention of advanced disease in the year 2014. The lower breast cancer mortality as a result of the Dutch screening policy reflects 6630 life-years gained in 2014. In addition (655 x 1.75 =) 1150 woman-years with the diagnosis of and treatment for advanced breast cancer will be prevented, which is apart from the life-years gained the second most important favourable effect of screening.

By including the gain in (1150 x 39% =) 440 Quality-Adjusted Life-Years (QALYs) due to the prevention of advanced disease, screening as a whole is estimated to lead to a relatively minor overall (negative) impact on quality of life. Unfavourable aspects of screening are responsible for a decrease in quality of life of 10%, due to the loss in quality of life during the screening examination (+ 85 QALYs) and (more) treatment and follow up (+ 570 QALYs). However, this effect is almost entirely counterbalanced by the gain in quality of life due to the prevention of advanced breast cancer. The total number of QALYs (6425) is only 3% lower than the total number of life-years gained (6630).

The lower part of table 3.5 states the estimated screening and total medical cost for all women who will need assessment, primary treatment or treatment for advanced breast cancer in 2014. In the absence of screening, the cost for treating all women with advanced disease is (4245 x 21,000 dollars =) 90 mln US dollars when including extramural cost. This represents 42% of total expenditures related to breast cancer. Forty-seven per cent of the annual - cost of screening (30 mln US dollars) will be offset by savings due to a decrease in the number of women treated for advanced disease (= 14 mln US dollars). If costs generated by screening (false positives, more primary treatment and longer follow up) are taken into account, one third of the cost of screening will be offset by savings. The impact on the demand for health care services is another aspect. The decrease in systemic and palliative treatments results in a prevention of drug treatment and in a decrease in hospital admissions and in-patient days, primarily noticed by the physician in internal medicine. However, changes in palliative surgical procedures or radiothera-

Table 3.5 Impact of nationwide breast cancer screening on prevention of advanced disease

	No screening	*50-70 (2 yr)*	*Difference 50-70 (2 yr)*	*Difference 50-65 (3 yr)*
Number of new patients with advanced disease	4245	3590	- 655	- 405
Total number of life-years gained			6630	4460
QALY-adjustment, due to:				
Screening	0	85	+ 85	+ 45
Assessment/biopsy	135	125	- 10	- 15
Primary treatment/follow up	6700	7270	+ 570	+ 335
Advanced disease	2900	2460	- 440	- 275
Total			+ 205	+ 90
Total quality-adjusted life-years (QALYs) gained			6425	4370
Total cost of:				
Screening	0	30	+ 30	+ 18
Assessment/biopsy	47	45	- 2	- 2
Primary treatment/follow up	76	82	+ 6	+ 3
Advanced disease	90	76	- 14	- 9
Total	213	233	+ 20	+ 10

Change in number of new patients, in quality of life and in cost (in mln US dollars). Year 2014 for three situations: without screening, 2-yearly screening of women aged 50-70 (and difference), and difference with respectively 3-yearly screening of women aged 50-65. Costs are annual total costs for all applicable women in the Netherlands. No discounting.

peutic sessions will be small, which is quite in contrast to the large and immediate changes in primary treatment as a result of the implementation of screening (see part II of this thesis).

For the sake of comparison, the consequences of a 3-yearly screening interval for women aged 50-65 (United Kingdom policy) if implemented in the Netherlands are presented in the right column. The smaller age range invited is responsible for a 28% lower number of breast cancer deaths prevented. Due to the fact that women will have less screens with a 3-year interval (given the same time period), the number of prevented treatments will be reduced by another 13%, resulting in a total of 405 (versus 655 in the other policy). But again, the impact of preventing advanced disease on the total effects and expenditures is clear: 275 quality-adjusted life-years are gained due to less advanced disease and no less than 44% of the annual cost of screening are offset.

For both policies in Table 3.5, the estimated mortality reduction is based on the average reductions achieved in Kopparberg/Östergötland and Malmö. The published update of screening trials (Wald et al., 1991) indicates an approximately 25% reduction, if the HIP-trial (Health Insurance Plan of Greater New York) and the TEDBC (UK Trial of Early Detection of Breast Cancer) are included and the actual differences in screening interval or attendance rate are neglected. We have considered the HIP-results to be too outdated for predicting the effect of present European programmes. In our estimate, we corrected the results of both Swedish randomized trials for the specific differences in age groups, screening interval, attendance rate, follow up period, and trial size. Considering the other point estimates on mortality reduction from other trials, our "Swedish average" from Kopparberg/Östergötland and Malmö appears a realistic one (de Koning et al., 1991).

Discussion

Average treatment

A comparison with literature and some national data was made to check whether the selection of patient files had not resulted in a serious bias. One of the main prognostic discriminants is the site and number of metastases, which makes it difficult to compare (inter-)national data deriving from different hospitals. Nevertheless, median post-recurrence survival is mostly in the range of 18-24 months (Powles et al., 1980; Patel et al., 1986; Tomin and Donegan, 1987; Perez et al., 1990; Dixon et al., 1991), which strongly resembles our data (20.3 mo). Almost 60% of the women in the three hospitals had received radiotherapy for advanced disease. Our national survey amongst radiotherapy departments indicated that approximately 1925 female patients have had radiotherapy in 1988 for advanced breast cancer, which is also approximately 60% of the annual number of women that died from breast cancer. In general,

these data support our findings from the file study.

Literature on the cost of treating women with advanced disease is scarce. In the Netherlands, de Waard analyzed 52 files of women treated in 1973-74 for primary breast cancer, who had developed metastases in the ten years after (de Waard et al., 1986). With tariffs based on the public health insurance scheme, which are lower than those for privately insured patients, the cost of treating metastases was calculated to be 16,000 US dollars (total duration of disease unknown). The average number of in-patient days was 66, which highly influenced the cost. The lower number of 45 in our study is certainly a reflection of the policy in the eighties of not admitting all women to hospital for chemotherapy any more. Medicare data from the United States have shown that the cost of care for breast cancer patients delivered during the last 6 months of life are 15,137 US dollars (Baker et al., 1991). Cost for continuing care in the months before is 483 US dollars per month, but this is not just attributable to cancer. A baseline estimate for all Medicare patients was 235 US dollars per month. Total medical expenditure for women surviving 21 months can thus be calculated to be approximately 19,000 US dollars.

The most important part of the cost is related to the number of in-patient days. Our national estimate on all admissions to hospital shows fewer in-patient days (in the range of 35 to 40), which is more in line with the median number of days in our patient files, but one can not select as specifically to advanced breast cancer as one is able to do with actual patient files. The same applies to in-patient days in nursing homes, although the national data are very close to those in the file study (9.5 instead of 8.3 days). However, in general our cost analysis is a moderate estimate, as cost for analgesics, expenditure incurred by patients and their families or extramural cost have not been included in the file study. Furthermore, the cost of medication was based on the cheapest retail prices. A separate analysis has shown that extramural cost are at least 3,860 US dollars per woman (Koopmanschap et al., 1992). Total cost for women treated for advanced breast cancer would then become approximately 21,000 US dollars, which resembles the scarce international data quite well.

Uncertainties

The years that mass screening will have a significant impact on diminishing the number of treatments for advanced disease are at least 5-10 years ahead. New treatment modalities like hyperthermia, new systemic therapies, use of haematological growth factors with chemotherapy or modalities interfering with growth factor-receptor interaction may have reached a more mature stage by then and possibly change treatment policies and results. The current increase in the use of adjuvant Tamoxifen in many countries, resulting in a possible decrease of the number of women suffering from advanced disease (Early Breast Cancer Trialists' Collaborative Group, 1992) is a good example. Criticism may also be expressed on the calculated savings due to mass

screening, as women whose advanced breast cancer has been prevented will have a chance of getting other diseases and use other medical resources.

Furthermore, our direct scaling method should in the future be adapted for being able to do measurements in patients. We have not used different utilities for different age groups. The relationship between age at time of disease and utility is not clear cut. Some studies do find a relationship, others do not. Furthermore, if changes are found, they may be in different directions for each dimension (Cassileth et al., 1984; Nerenz et al., 1986; de Haes et al., 1990). Nevertheless, our study does provide insight in the changes to be expected, at present, due to nationwide screening.

Whether our results will be applicable to other countries will depend on four important factors. The number of in-patient days and the percentage of women with radiotherapy are the most influential two factors for total costs. There may be differences between countries on these points. For instance, terminally ill breast cancer patients in Finland appeared to have been treated significantly less by radiation (Holli and Hakama, 1989). Recording the total durations of hormonal treatment and chemotherapy will be important for quality of life. Finally, the fourth and most important factor is the breast cancer mortality reduction that can be expected in different countries when implementing nationwide screening, since this is implicitly related to the prevention of advanced disease and its treatment, the savings and the gain in quality of life.

In conclusion, the impact of breast cancer screening with respect to the prevention of advanced disease and its morbidity is evident. In the long term, high quality screening will result in an important benefit for the women involved and in a considerable reduction of the amount of effort and money being spent to treat women with advanced disease. As long as there is no treatment highly effective against advanced disease, breast cancer screening will play an important role.

PART II

4 Impact on clinical medicine

Introduction

Several trials have shown that mammographic screening of post-menopausal women reduces breast cancer mortality. The introduction of a national programme in the United Kingdom, offering 3-yearly screening to women aged 50-65, would result in a mortality reduction of 8% (Knox, 1988).

Although breast cancer mortality reduction is the fundamental effect (Day et al., 1989b), there is much debate about other desirable and undesirable consequences of breast cancer screening (Warren, 1988; Skrabanek, 1988). However, publications which quantify these consequences are lacking. Starting to screen will result in a temporary increase in the number of women with newly diagnosed breast cancer. Detecting these cancers at an earlier stage will affect the type of assessment and treatment. Mass screening will also generate referrals of women, who appear to have no breast cancer (false positives).

This chapter presents the impact of nationwide screening on three issues: change in referral patterns and assessment procedures; change in breast cancer incidence; and change in stage distribution and subsequent primary treatment. An overview is given of the consequences for health care. The computations are made for the Dutch population, but the conclusions are relevant to other countries too. Predictions are made for both the first years of implementing the programme and for the stable situation to be reached after some years. Medical activities outside the programme are included.

This chapter explains systematically the different steps in the approach and results of integrating assessment and treatment changes in screening analyses. Part of the results (in absolute terms) has been updated in the other chapters of this thesis.

Chapter based on de Koning et al., Br J Cancer, 1990c

Material and methods

Generally accepted medical practice
A flow chart was made of the different assessment procedures that are used when breast cancer is suspected. Another flow chart was made of the main primary therapeutic procedures, representing generally accepted medical practice in the Netherlands. The criteria for the choice between available medical procedures were defined on the basis of protocols of the Dutch Comprehensive Cancer Centres, on literature and on interviews with expert clinicians.

Assessment and excision biopsy
In assessing breast abnormalities, at least three successive steps were distinguished: physical examination, clinical mammography, and excision biopsy (with possible specimen radiography). Additional possibilities which were taken into account for palpable lesions only were fine needle aspiration (cytology) and ultrasound. Estimates on the number and type of these procedures used in the situation without screening were first based on the following sources: general practitioners' registry (Trouw, 1986); the radioactivity and radiation application division of the Ministry of Welfare, Health and Cultural Affairs; the Central Office for the Administration of Specialists' Fees (de Waard et al., 1986) and the Steering Committee on Future Health Scenarios (1988).

Estimates on the change in numbers and types of assessment when introducing mass screening were derived using the referral patterns from the Dutch screening trials in the first 10 years. The Dutch results were adjusted for the fact that these were achieved in trial projects, in which quality standards will probably have reached higher levels now than at the beginning of a nationwide programme. The positive predictive value of a screening mammogram for women aged 50-70 is estimated to be 40% for the first screen and 60% for subsequent screens in the Netherlands. A distinction is also made between the predictive values for biopsies of non-palpable and of palpable lesions. False positives are assumed not to be treated.

When implementing screening, a decrease can be expected in the number of mammograms for women aged 50-70 resulting from visiting the general practitioner. This decrease in the number of mammograms and possibly successive assessment is assumed to be proportional to the decline in clinically diagnosed breast cancers.

Therapy
All women with breast cancer but without signs of distant metastases are assumed to receive primary treatment independently of the way it is diagnosed. The type of primary treatment presently used was determined by analyzing data on hospital admissions and breast surgery in the Netherlands

from 1983 till 1985 (unpublished data), data from clinicians and treatment protocols. The following types of treatment with curative intent were distinguished: total mastectomy (DCIS); total mastectomy with axillary dissection; total mastectomy with axillary dissection plus postoperative radiotherapy; and breast conserving therapy.

Estimates on future developments were based on expert interviews, international recommendations and literature (Harris et al., 1985; Holland et al., 1985; Hayward and Rubens, 1987; Steering Committee, 1988). This resulted in the assumption that women in the Netherlands with invasive breast carcinoma up to 3 cm in diameter, without fixed axillary lymph nodes, will undergo breast conserving therapy.

Women with axillary lymph node metastases are assumed to receive adjuvant systemic therapy: CMF for premenopausal and Tamoxifen for postmenopausal women (Early Breast Cancer Trialists' Collaborative Group, 1988). Irradiation on regional lymph nodes is assumed to depend on both lymph node metastases and tumour size. Women with distant metastases present at first detection of breast cancer or diagnosed in the course of time are assumed to receive treatment for advanced disease.

Simulation model of mass screening

The simulation package MISCAN was used to predict the effects of screening (Habbema et al., 1987). The model is based on the development of invasive breast cancer in three stages, reflecting its size ($<$ 10 mm, 10-19 mm, $\geq$ 20 mm). Initially, five per cent of the invasive cancers were assumed to be preceded by a screen-detectable ductal carcinoma in situ, for which 100% progression was assumed. Key parameters of the model are the mean duration of pre-clinical screen-detectable disease, sensitivity of mammography and the improvement in prognosis for screen-detected cases. These parameters were derived from results of the HIP-analysis (Habbema et al., 1986), and from a new analysis of the results from the first 10 years of the Dutch trials (van der Maas et al., 1989). Results from the randomized Kopparberg/Östergötland trial were used for estimating the improvement in prognosis in screen-detected patients in this chapter (Tabár et al., 1985).

In this chapter the simulated screening policy consists of 2-yearly mammography for women aged 50-70, starting in 1988 and ending in 2015, but effects emerging after 2015 are also calculated. The attendance rate was based on the Dutch trials and assumed to be 70% on average. The build-up period is considered to last 7 years.

Table 4.1 Key data on the diagnosis and treatment of breast cancer

	Referrals for clinical mammography[1]	*Excision biopsies*[1]	*Breast cancers*	*Primary treatment*	*Treatment of advanced disease*[2]
1988					
No screen, all ages	71,000	23,500	7,300	6,850	2,600
No screen, age 50-70	30,300 (43%)	10,000 (43%)	3,100 (43%)	2,900 (43%)	1,000 (38%)
1994					
No screen, all ages	73,500	24,500	7,750	7,250	2,850
With screen, all ages	66,000	23,800	8,750	8,275	2,750
1998					
No screen, all ages	77,000	25,500	8,000	7,450	3,050
With screen, all ages	64,500	23,000	8,225	7,700	2,900

Predicted annual numbers for the Netherlands, with and without mass screening. Years 1988, 1994 (network complete) and 1998 (steady situation). 2-yearly screening of women aged 50-70. Female inhabitants 40-84 years: 3 million. For the year 1988 numbers are also given for the 50-70 age group.
[1]. To confirm or exclude breast cancer. [2]. New cases.

The model predicts the yearly number of women with newly diagnosed breast cancer in the situation with screening and if no screening is carried out. Cancers are classified according to size, invasiveness and way of detection. The simulated data was combined with data on lymph node metastases and on distant metastases from all breast cancer cases in the projects at Utrecht and Nijmegen (screen-detected and clinically diagnosed). This resulted in predictions on the distribution of cancers according to size, lymph node metastases and distant metastases. By using the assumptions on predictive values, assessment and treatment flow charts, these were translated into outcomes concerning national assessment and treatment procedures. It is assumed that all women who had (or would have) died from breast cancer had or would have been treated for advanced disease.

Results

Assessment procedures for breast cancer

In the Netherlands, it is the general practitioner who is consulted first in case of breast symptoms or complaints. Dutch registries revealed that in the course of one year 9% of the female population older than 40 visits the physician for breast assessment. About 30% of these consulting women are referred for clinical mammography, which in turn will lead to an excision biopsy in about one third of the referrals.

From other sources we have estimated the annual total number of clinical mammographies to be about 120,000 (year: 1987). Adjusting this figure for mammograms due to other reasons (reconstructive surgery, follow up, metastases of unknown primary), or in other age groups, a mammogram is made each year in 2% of women older than 40 years to confirm or exclude the possibility of breast cancer. This estimate corresponds well with that from the general practitioners' registries.

Table 4.1 shows the main steps in the diagnosis and treatment of women in the Netherlands with a (possible) breast cancer, and the predicted numbers with or without mass screening. In 1988, before national screening began, 71,000 women were referred for clinical mammography and 23,500 women successively underwent an excision biopsy to exclude or confirm breast cancer. Related to the annual incidence rate, which amounted to about 7,300, it can be concluded that the ratio between the number of diagnosed cancers and the number of biopsies is rather low. Malignancy is only confirmed in 30% of these biopsies, thus in 10% of women as a result of a clinical mammogram. Table 4.1 also shows that the age group 50-70, the target population of the future Dutch screening programme, accounts for 43% of all breast cancer cases.

Given the assumptions on screening, in 1994 when the screening network is planned to be completed, a 10% decrease may be expected in the

total number of women referred for clinical mammography. The number of women with newly diagnosed breast cancer will be higher, mainly because of the prevalence load of pre-clinical cancer in the screened population. In the first years of the programme, the decrease in biopsies is not as evident as the decrease in mammograms. This is due to the prevalence load, to the relatively large number of non-palpable lesions detected by screening and to a lower predictive value at the first screen.

Especially in the build-up period, there will be an increase in the number of biopsies for non-palpable lesions of approximately 12% per year. After this period, the number of biopsies for non-palpable lesions will remain slightly higher. On the other hand, a strong decrease of 2,700 biopsies of palpable lesions per year can be expected in the long run. The decrease in assessment is explained by the fact that the positive (mammographic) predictive value of 40-60% in a screening programme seems higher than the predictive value of a (diagnostic) referral by general practitioners. If one million women aged 50-70 are screened at 2-yearly intervals, some 2,100 women will subsequently undergo an excision biopsy, without breast cancer being confirmed (false positives). But there will be a reduction in the number of biopsies with a benign histological result outside the programme, larger than the increase in false positive biopsies from screen referrals. The preceding number of clinical mammograms will decrease too, and may be related to the increase in clinical mammograms made for screened women.

Women with newly diagnosed breast cancer

At the start of screening an initial increase is expected in the number of women with newly diagnosed breast cancer. In 1994 1,000 (=13%) extra cases will be diagnosed (Figure 4.1). The number of women with breast cancer will remain slightly higher (3%) in the steady state of screening. The latter is caused by the (earlier) detection of breast cancer in women who would have died from other causes (in the absence of screening), before breast cancer had been diagnosed.

Screen-detected cancers tend to be smaller and tend to have a more favourable lymph node status, than clinically diagnosed cancers (Figures 4.2a and b). Figure 4.2 applies to the years following the build-up period, when 26% of all women with breast cancer will be detected through screening. The programme will also detect some 125 women each year with ductal carcinoma in situ (DCIS), which represents 1.5% of all newly diagnosed breast cancers. This carcinoma is diagnosed only occasionally in clinical setting (Rosen, 1979; Rosner et al., 1980). This change is important to realise because of the existing uncertainty about the natural history of DCIS (Schnitt, et al., 1988).

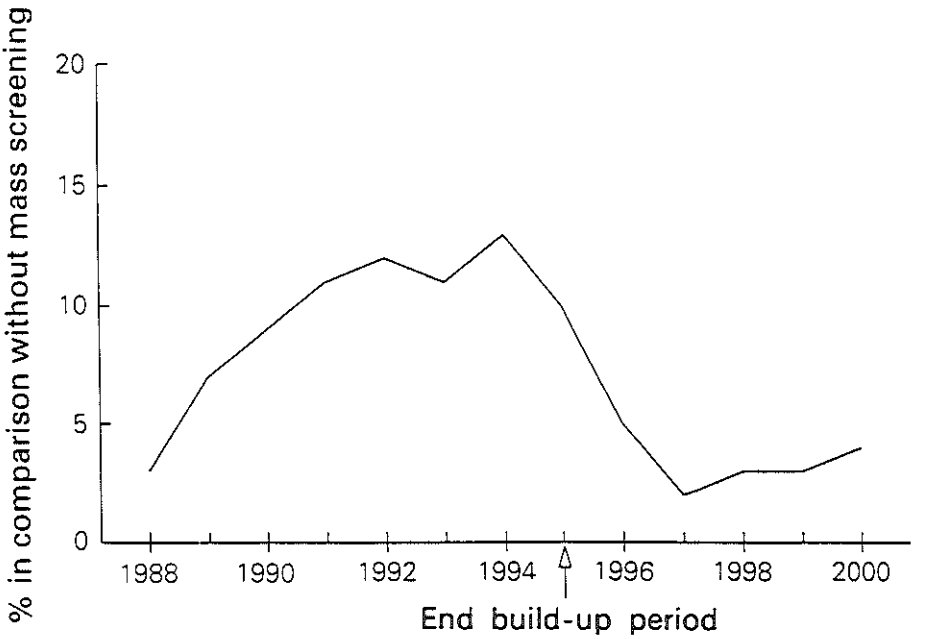

Figure 4.1 Predicted increase (%) in the yearly number of women with newly diagnosed breast cancer, when implementing nationwide screening. Percentage compared to the situation without mass screening in that year. Screening 2-yearly, women aged 50-70. Build-up period 1988-1995: 1988, 8% of the screening network functioning; 1991, 52% of the screening network functioning; 1995, screening network planned to be complete.

Changes in treatment

The total number of women to be treated with curative intent will increase as a result of screening temporarily. The more favourable stage distribution will also influence the modality used, in particular breast conserving therapy. Our analysis of hospital registries shows a moderate increase in breast conserving therapy in the years 1983-1985 (from 9% in 1983 to 12% in 1985). However, in recent years this development has continued. Considering the present treatment protocols we have concluded that in the near future approximately 40% of women with breast cancer would undergo breast conserving therapy, if no screening is carried out.

Figure 4.3 shows the predicted changes in primary treatment as a result of mass screening, compared to the situation without screening in 1994 (first column). The second column of each modality applies to the build-up period with a large number of prevalent cases (1994), the third column to the steady state (1998). In the long run the expected increase in breast conserving therapy will be 25% per year, compared to the situation without screening. More tumours are detected with a size under 3 cm and without node metastases.

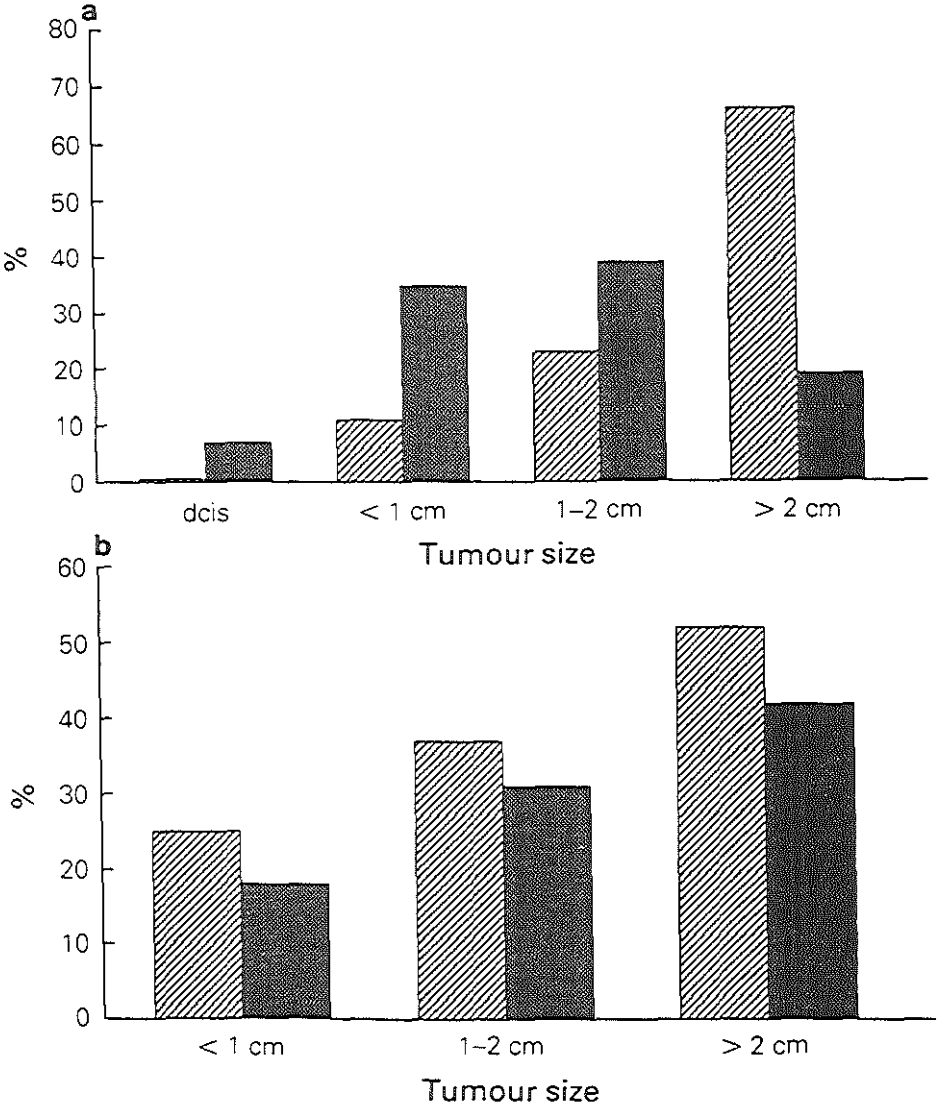

Figure 4.2 Predicted distribution of two parameters in screen-detected (right bars) and in clinically diagnosed (left bars) breast cancers. The distribution applies to the stable situation 2000-2015. Screening women aged 50-70 every 2 years. a. distribution by size of tumour. b. percentage of axillary lymph node metastases by size of tumour (women of all ages).

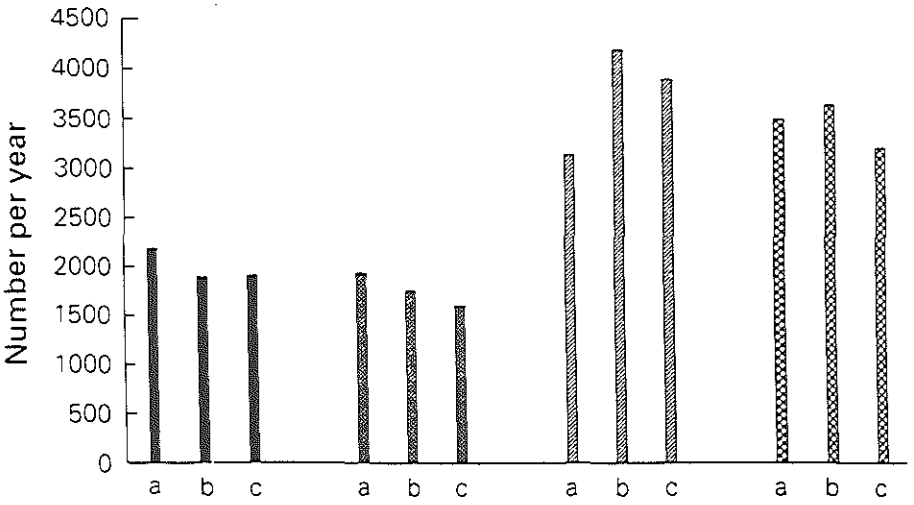

Figure 4.3 Yearly number of primary treatment modalities for breast cancer and predicted changes as a result of nationwide screening of women aged 50-70 every 2 years. Build-up period 1988-1995. Demographic change between 1994 and 1998 excluded. a. Year 1994 without mass screening. b. Year 1994 with mass screening (network being completed). c. Year 1998 with mass screening (stable situation).
From left to right: total mastectomy; mastectomy + radiotherapy; breast conserving therapy; adjuvant systemic treatment (both hormonal and chemotherapy).

Consequently, the number of mastectomies (plus axillary dissection) without postoperative radiotherapy will decrease by 13%. The number of women treated by mastectomy, axillary dissection and postoperative radiotherapy will decrease by 17% per year, as a result of the more favourable lymph node stage, when assuming no differences in treatment criteria in the near future and assuming that the more favourable lymph node status of screening is equally divided for patients undergoing breast conserving therapy or mastectomy. The 8% decrease in adjuvant systemic treatment also reflects the difference in node involvement. The decrease in mastectomies does not equal the increase in breast conserving therapy, as a result of the higher incidence and the detection of non-invasive lesions due to screening.

In the first years of screening the change in breast conserving therapy is even 35%, as a result of the more pronounced increase in newly diagnosed cancers. This may be a relatively small increase on the total number of surgical procedures, but it will cause a strong increase for radiotherapy sessions. At present one-quarter of all new patients who need radiotherapy are breast cancer patients. On the other hand, the stage distribution of breast cancers detected at the first screening round is still less favourable compared to cancers detected in the stable situation some years after. Therefore, in the first years of the programme, the number of mastectomies will only gradually decrease.

Finally, the influence on treatment of advanced disease is important (Table 4.1). For each patient, this involves a combination of treatments, spread out over a longer period. In the absence of screening each year 2,600 breast cancer patients undergo palliative treatment for the first time, due to diagnosing breast cancer at an advanced stage, or due to recurrences. Mass screening will reduce the number by 3.5% in 1994, and by 5% in 1998. By the year 2015 the reduction amounts to 12% (van der Maas et al., 1989). It will take 25 years for this change to be reached, which is quite different from the immediate changes in primary treatment.

Discussion

The estimates about the impact on clinical medicine are calculated for the Dutch population and screening policy. A relevant question is whether the results would also apply to other countries where centrally organized programmes are being implemented, and what the main uncertainties are.

Predictive values

One important aspect in nationwide screening is the predictive value of a positive screening mammogram. A relatively low value (considering age group) will strongly increase workload for clinical medicine and may result in a less favourable balance between positive and negative effects. Our assumptions are based on results obtained in the Dutch trials over 10 years. The predictive values of biopsies (52% at first and 70% at subsequent screens) correspond well with those of 50% and 75% respectively, reported from Kopparberg/Östergötland. The policy not to take a biopsy of mammographically benign lesions and the systematic consultation between different specialists may contribute to these good results. In some countries less favourable figures are reported (Day and Miller, 1988).

The question remains whether good results from screening projects will also be achieved in a national setting (Day et al., 1989b). Quality control will be an important part of the programme in the Netherlands. Training of radiographers, radiologists and pathologists before the implementation of screening, and periodic evaluation of diagnostic and technical performance will be tasks of a National Expert and Training Centre.

Detecting more breast cancers and non-invasive cancers

Theoretically, overdiagnosis and subsequent overtreatment are important risks involved in mass screening. In practice it is difficult to assess whether overdiagnosis is actually taking place. Screening always leads to a (temporary) increase in the number of women with breast cancer, which is only a desirable effect and should not be confused with overdiagnosis. After termination of the programme one would expect a relative fall in this number.

Inevitably, some women will be detected who would, in the absence of screening, have died from other causes before the cancer had become manifest. However, this percentage is rather small; we predicted an increase in the incidence of at most 3% because of this phenomenon.

Overdiagnosis may also occur if some of the tumours would, in the absence of screening, never have progressed, or only very slowly, to a stage in which symptoms would lead to a clinical diagnosis. This may apply to ductal carcinoma in situ. In this chapter all screen-detected cancers are assumed to be progressive. The percentages of DCIS are very similar in the major screening trials, and vary between 9% and 15% of screen-detected cancers (Hendriks, 1982; Tabár et al., 1985; Andersson et al., 1988; UK Trial, 1988). In this respect our breast cancer (screening) model fits with international data. When starting nationwide screening, this type of cancer will still remain a small fraction of the total number of diagnosed breast cancers. In view of these modest percentages, it can be concluded that even if some of these tumours would not progress or even regress, it would only result in a very small amount of overdiagnosis (see chapter 8). For the women invited, the fact that non-invasive carcinoma, until recently treated by mastectomy, may be a non-progressive lesion is of more (psychological) importance. Knowledge that less radical treatment for early lesions is advocated could have a positive effect on the attendance rate.

Detection rates and stage distribution

The predicted numbers of screen-detected and interval cancers are based on assumptions on pre-clinical duration and sensitivity, and are compatible with data from screening projects in Nijmegen, Utrecht and New York. These data were also found to correspond closely, adjusting for screening interval, with published data from the randomized Kopparberg-Östergötland trial (Figure 4.4).

Moreover, the distribution of tumour size and lymph node metastases of the detected cancers influences the predictions on assessment and treatment. The data on women with clinically diagnosed breast cancer were comparable with data from large non-screened patient groups. In screen-detected cases, we expect 30% of the cancers to be over 20 mm at the first screen, and 15% at subsequent screens, which results in an average of 20% in the stable period (Figure 4.2a). Tabár reported 26% of the cancers to have a diameter of 20 mm or more up to 1986 (Day et al., 1989b). Within each tumour size class, women with screen-detected cancer had a lower percentage of axillary lymph node metastases in the Kopparberg/Östergötland trial, although not statistically significant (Tabár et al., 1987a). However, the same difference was found in the Dutch trials and is used in the treatment predictions.

Impact on primary treatment

Publications on changes in treatment as a result of mass screening are scarce

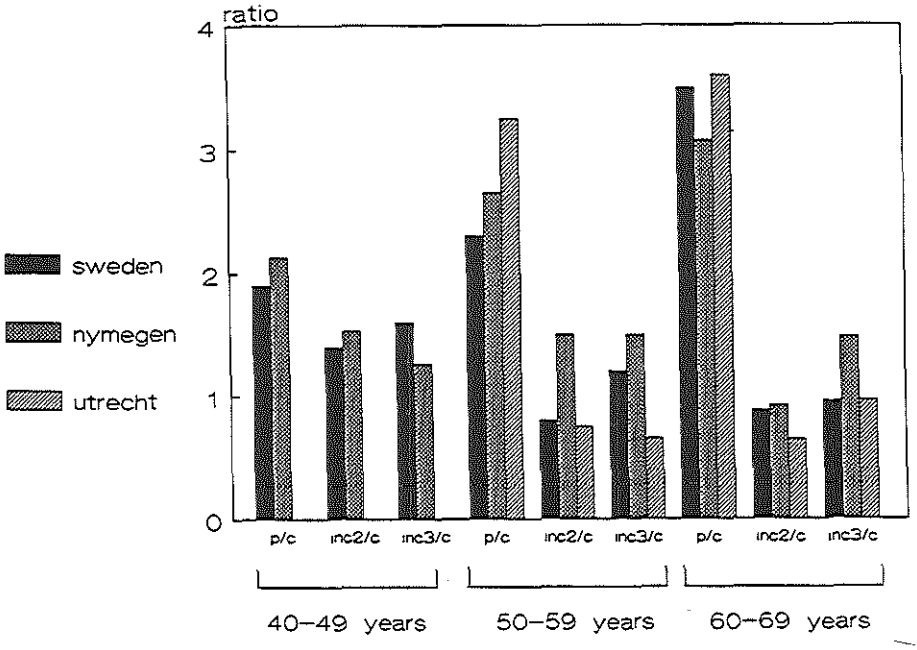

Figure 4.4 Comparison of the results of the trials in Sweden (Tabár et al., 1988), Utrecht and Nijmegen. The numbers give the ratio between the number of cases detected in a given age group by screening and the incidence in a comparable control group. The detection rate in the first screening round (p) and the 'incidence' rate of the second (inc 2) and the third screening round (inc 3) are divided to the incidence rate in the control group (c). 'Incidence' (2 and 3) includes interval cancers; some adjustments have been made to be able to compare the three trials (different intervals).

(Holmberg et al., 1986a; Andersson et al., 1988). According to Andersson, no important differences were found in the treatment of women with breast cancer between study and control group of the Malmö trial. However, Malmö figures do show that less women were given hormone therapy or chemotherapy in the study group. It is unclear whether the therapy trial, which was operational, contributed to this difference.

More importantly, in the control group in Malmö 20% of the women with breast cancer underwent breast preserving therapy. In the study group, this figure was 25%, if including stage 0 tumours, and 23% if not. This means a relative increase of 15-25% compared to the control group. Because of small numbers, the difference was not statistically significant and did not seem important. But for a nationwide programme this increase in breast conserving therapy by 15-25%, when compared to the situation without screening, will have a major impact on health care services. The Malmö trial seems to support our 25% increase in this type of treatment.

At least three important developments may influence the predictions on primary treatment. The first is the tendency to apply breast conserving therapy even in tumours over 3 cm (van Dongen et al., 1987). Secondly, there is

discussion whether radiotherapy may be omitted in some cases, e.g. in screen-detected small or in situ lesions. Finally there is increasing evidence about the benefit of adjuvant systemic treatment.

Assessment outside the screening programme

Our results show that centrally organized screening will result in less biopsies than in the situation without screening. This beneficial effect will depend to a large extent on the expected change in the number of preventive mammograms after the implementation.

Many women visit their doctor for breast problems nowadays. Gravelle et al. (1982) already found that 4,7% of women above 40 years of age were examined at the hospital (or undergoing a biopsy) because of breast symptoms or complaints, but no cancer was found ("worried well").

In the Malmö trial, the control group had free access to mammography equipment outside the programme. Some 24% of women in the control group (aged 45 and more) have actually had a mammogram in a mean period of 8.8 years, most only once (Andersson et al., 1988). A percentage of 13% over a period of maximum 7 years is mentioned in the Kopparberg-Östergötland study (Tabár et al., 1985). These figures are comparable with our estimate of 2% in the age group over 40.

However, it is difficult to predict the possible decrease in 'preventive' mammograms outside the programme, when a national screening programme is being implemented. Some 500,000 screening mammograms per year in the Netherlands will certainly cause a strong decrease in mammograms outside the programme in the age group 50-70. However, it also remains uncertain how much the demand for women in other age groups will increase at the same time.

Overview

The increase in referrals and women diagnosed having breast cancer will have its impact on workload in the first years of screening. Breast conserving therapy requires more time and effort from surgical and radiotherapy departments, and assessment of non-palpable lesions with specimen radiography and paraffin section increases workload for pathologists. The decrease in women with advanced disease will have the opposite effect on workload only after several years. Uncertainties remain about the attitude of women towards breast assessment outside a screening programme. The present results may enable a more balanced judgement to be made on the value of breast cancer screening and its impact on clinical medicine. To be able to ensure these advantages, assessment and treatment should also be monitored and evaluated in a national system.

5 Impact on quality-adjusted life-years

Introduction

Mammographic screening reduces breast cancer mortality in women aged 50 years and over. Moreover, it affects the quality of life of women participating. Therefore, when studying the effectiveness of a screening programme, preferably both endpoints should be taken into account. This was done in an evaluation of breast cancer screening in the United Kingdom (Forrest, 1986). Effects on life-years were adjusted by multiplying these years by values for quality of life which were between 0.92 and 0.99. However, it is not evident whether such similar values can be used for all phases in the screening and disease process.

A different approach has been chosen here. Breast cancer screening may have an impact on quality of life in the short run: short-term emotional effects can be expected in women participating (Dean et al., 1986; Ellman et al., 1989; Eardley and Elkind, 1990). It also has an impact in the long run as a shift is to be expected in the number of women experiencing early and advanced phases of the disease (de Koning et al., 1990c). Quality of life is affected in these phases and there will therefore be a concurrent shift in quality of life.

To determine the quality of life adjustment, 3 steps have been distinguished. (1) The consequences of disease and treatment in the different phases had to be mapped. (2) An overall value was assigned to these phases. (3) By multiplying this value by the duration of the different phases and the number of women to be expected in these phases, quality-adjusted life-years after the implementation of a breast cancer screening programme and without such a programme could be computed, and the difference calculated. In this chapter the effect of 2-yearly screening for women aged 50-70 on quality-adjusted life-years is investigated in detail and the question posed whether taking into account quality of life alters the results of the effectiveness analysis.

Chapter based on de Haes, de Koning et al., Int J Cancer, 1991
Acknowledgements are given to Hanneke de Haes for allowing me to publish this chapter in my thesis.

Table 5.1 Content of selected papers dealing with quality of life and breast cancer (screening)

	Total (N=176)		Empirical (N=139)	
	N	%	N	%
Screening attendance[1]	6	3	3	2
Diagnostic phase	25	14	13	9
Initial surgery	47	27	30	22
Initial chemotherapy	31	18	22	16
Initial radiotherapy	13	7	3	2
Initial hormonal therapy	7	4	5	4
Breast reconstruction	15	9	11	8
Disease free (d.f.)3 mths-1 yr after mastectomy	47	27	38	27
D.f. > 1 yr after mastectomy	55	31	46	33
D.f. 3 mths-1 yr after br.conserving therapy	10	6	8	6
D.f. > 1 yr after br.conserving therapy	17	10	12	9
Palliative treatment (p.t.) + surgery	11	6	8	6
P.t. + chemotherapy	33	19	30	22
P.t. + hormonal therapy	19	11	17	12
P.t. + radiotherapy	10	6	8	6
Terminal illness	5	3	2	1
General	8	5	4	3
Physical[2]	132	75	104	75
Psychological	117	65	93	67
Social	67	38	49	35
Global	11	6	11	8
Utility	2	1	1	1
Comparative[3]	88	50	84	60
Descriptive	141	80	111	80
Subjective[4]	---	—	39	28
Subj. + objective	---	—	59	42
Objective	—	—	41	29

notes 1-4, see appendix 5.1

The consequences of breast cancer (screening) and value assignment; a literature review

An empirical study of all relevant disease and treatment phases involved, with regard to their impact on quality of life, was not feasible in the context of this study. Therefore, the consequences of breast cancer and its treatment and the values of the health states described were, preferably, to be derived from published studies. A Medline computer search was done covering the English literature from January 1982 until October 1988. The key words 'health status', 'psycho-social', 'psychological', 'psychiatric', 'performance' and 'quality of life' were combined with the key word 'breast cancer'. Reports on the toxicity of the main hormonal and cytostatic treatment regimens (Tamoxifen and CMF) were sought for as well. From the reference lists in selected studies, other and older relevant titles were searched for. Finally, 252 articles were selected, of which 176 turned out to be useful. The others were related to benign breast disease, did not contain specific data with regard to breast cancer or were related to breast cancer (screening) but not to its impact on quality of life (e.g., breast self examination, prediction of the course of the illness as a result of psychosocial variables or the prediction of quality of life on the basis of other factors than the disease and treatment).

The selected articles were classified, as shown in Table 5.1. Of the 176 studies, 139 reported empirical data. This distinction is made because empirical papers can more directly be translated into health state descriptions. Firstly, it can be seen from Table 5.1 that the amount of attention paid to the different phases in the disease and treatment process varies, in the literature in general as well as in empirical work. Ample attention is paid to primary surgery and chemotherapy, the disease-free period after mastectomy and palliative chemotherapy. Little attention is paid to hormonal treatment, radiotherapy and terminal illness. The same holds true for effects of screening attendance: 2 empirical studies are related to long-term effects and one to screening attendance outside a programme organized for the population at large. However, the main effects one would expect are short-term ones: immediately before or after screening participation. Secondly, it is evident that more attention is paid to the physical and to the psychological domain than to social consequences. A global evaluation is seldom reported, and neither are utilities. Only Buxton et al. (1987) described utility scores (see definition on page 24) for the situation 1 year after surgery. These values lie between 0.22 and 0.76. Thirdly, several studies report on quality of life in a descriptive manner. Moreover, the subjective and objective approach are often combined.

We concluded from Table 5.1 that the literature provides insufficient data from which to derive values suitable for direct insertion into the cost-effectiveness (CE) analysis. Some phases in the disease and treatment process have not extensively been described and no global evaluation or utility score is given for most of the relevant health states. We, therefore, decided to construct a questionnaire describing the patients' situation in the distinct phases and approach

a sample of respondents to evaluate these health states and subsequently determine a quality adjustment factor.

Material and methods

Construction of the questionnaire

Starting from the physical, the psychological and the social dimensions, the symptoms and functional levels reported in the literature were summarized for each disease and treatment phase. The prevalences may vary in different studies. Therefore, ranges were added unless a phenomenon was reported in less than 25% of the patients in all studies or if no empirical data were available. When data were available on comparison with other groups, this was reported in brackets (e.g. breast conserving therapy > mastectomy). As for some phases more findings are reported in the literature, some health state descriptions became more extensive than others. The states listed in Table 5.1 were described. Breast reconstruction has been omitted as it turned out to be applied in only 1% of the cases (van Dam and Bergman, 1988).

Thus, 15 health state descriptions were constructed. The states were presented in random order to prevent bias as a result of the order given. A visual analogue scale was given underneath the state descriptions. Respondents were asked to give a mark reflecting their evaluation of the state on a 10-cm line, on which the anchors were the worst (score = 0) and the best imaginable (score = 100) quality of life. The score of the respondent was determined by measuring the distance between 0 and the mark given. An example is given in Figure 5.1.

For quality-adjusted life-year analysis, the values or utilities should reflect the subject's willingness to trade off quality against length of life. This is the case when using the 'standard gamble' or the 'time trade-off' technique. The respondent is asked whether (s)he would opt for a better quality of life and a (risk of) shortened survival or for a worse quality of life but longer survival and in which circumstances both situations are balanced (Torrance, 1976). However, some problems exist (de Haes and van Knippenberg, 1989): only a limited number of phenomena can be inserted into these health state descriptions and data are necessarily collected orally. Therefore, in this study, a direct scaling method (the visual analogue scale) has been used.

Scores on direct scaling methods are systematically lower, but are related by a power function to the scores of the 'time trade-off' or 'standard gamble' method. In breast cancer screening, attenders are asked to trade off quality of life in the short run against duration of life in the long run. Therefore, the 'time trade-off' method has been taken as a starting point. Bombardier et al. (cited by Loomes, 1988) have described the relation between the 'time trade-off' (TTO), and 'visual analogue scales' (VAS): $TTO = 1-(1-VAS)^{1.82}$. Using this function, the 'time trade-off values' for the different states could be computed from the scores of our direct scaling method.

DIAGNOSTIC PHASE (duration 1-2 months)

physical:	sexual dysfunction/decreased libido (0-29%) lethargy/fatigue (41-47%) not able to work (16-26%)
psychological:	psychological dysfunction (66%) psychiatric morbidity (31-46%) fear (e.g. of death or aesthetic consequences) (5-83%) depression (2-61%) emotionally most difficult period (42%) irritability (30-53%) helplessness
social:	deteriorated social function (18-29%) problems discussing emotions (89%) unchanged satisfaction with social relations

worst imaginable quality of life — 0——————100 — best imaginable quality of life

Figure 5.1 Example of phase description and method of evaluation in the questionnaire

Study samples

Ideally, one should consider 3 groups to evaluate the states: the normal population, breast cancer patients, and experts. However, our method of establishing utilities was novel and expected to be too complicated to approach either patients or the population at large. To test the feasibility of the method, 18 employees of the Department of Public Health were first asked to participate. Two found the questionnaire too difficult and finally 15 (83%) completed the questionnaire. As the response in this group had turned out to be satisfactory, 13 experts, key figures in breast cancer treatment and epidemiology in the Netherlands, were approached next. Three surgical oncologists, 3 medical oncologists, 2 radiotherapists, 1 radiologist, 2 epidemiologists and 1 psychologist returned the questionnaire (92%).

Duration of health states
The duration of the different states had to be established as to indicate the length of loss in utility associated with each phase. The duration of possible morbidity associated with the screening examination itself is difficult to determine. The literature has not shown any (negative) long-term effects of attending the screening (Dean et al., 1986; Ellman et al., 1989; Gram et al., 1990). We have assumed a period of 1 week applicable for all women who attend. In the Netherlands, this is the maximum period between the mammographic examination and receipt of the screening result by letter. The total diagnostic phase from referral by the general practitioner up to an excision biopsy takes 5 weeks on average.

The effects of surgery and/or radiotherapy for women in whom breast cancer is diagnosed have been divided into short- and long-term morbidity. A period of 3 months for the initial effect of surgery was mentioned in the questionnaire. However, this would imply some overlap with the radiotherapy period for those women who need radiotherapy. A period of 2 months, covering hospitalization and waiting time until start of radiotherapy, was considered in the final analysis for all women who are undergoing surgery. The initial morbidity of radiotherapy will primarily depend on the total number of sessions. A mean period of 2 months has been taken into account. The long-term effects of breast-conserving therapy or mastectomy become manifest after these initial periods in the so-called disease-free interval (2-12 months; > 1 year). If women received adjuvant chemotherapy (6 months CMF) or hormonal treatment (2 years in age group 50-70 and 5 years at ages 70+), losses in utility due to adjuvant treatment were considered in these years. The different types and durations of treatment for women with advanced disease have been based on a file study of 68 women who had died from breast cancer (de Koning et al., 1992a).

Effects of the breast cancer screening programme
The MISCAN breast cancer model was used to predict the number of women in the relevant disease and treatment phases with and without a breast cancer screening programme (de Koning et al., 1990c; van Oortmarssen et al., 1990b). The number of women in whom breast cancer will be diagnosed and the stage in which these cancers will be diagnosed after or without the initiation of screening are predicted until the year 2088, based on data from the Dutch screening trials. The improvement in prognosis is based on the reduction achieved in the randomized trials in Kopparberg/Östergötland (Tabár et al., 1989) and Malmö (Andersson et al., 1988). Results in this chapter are based on nationwide screening organized for women aged 50-70 at 2-yearly intervals (Dutch policy). The attendance rate is assumed to be 70% on average.

Results

Reliability

As a new method had been developed, we investigated whether the collected data were reliable and not attributable to random variation. Therefore, the ranking of the states by the 2 groups of respondents was compared first. For each individual respondent, the ranking of the different states was established. Next, the mean ranking of the states in the 2 groups was computed (Table 5.2). As can be seen from this Table, in only 4 out of 15 states a substantial difference (>2) between the groups can be noted: hormonal treatment was ranked more favourably by the experts, whereas they ranked the diagnostic phase and palliative radiotherapy lower (less favourably) than the employees of the Department of Public Health. Other disease/treatment phases were ranked almost equally in both groups. These rankings are, therefore, not attributed to random assignment.

Table 5.2 The rank order of the disease/treatment phases within the two groups of respondents

		Dept. of Public Health (N=15)	Experts in breast cancer (N=12)
1[1]	3 mths-1 year after mastectomy	9.5[2]	9
2	Palliative + surgery	5	4
3	Palliative + chemotherapy	2	2
4	Initial surgery	8	8
5	Palliative + hormonal therapy	3	6
6	Initial radiotherapy	9.5	7
7	Initial hormonal therapy	6	10.5
8	Initial chemotherapy	4	5
9	3 mths-1 year after breast-conserving therapy	11	12
10	Palliative + radiotherapy	7	3
11	Terminal illness	1	1
12	Screening attendance	15	15
13	Diagnostic phase	13	10.5
14	Disease-free > 1 year after mastectomy	12	13
15	Disease-free > 1 year after breast-conserving therapy	14	14

[1] the order of the states in the table is the order in the questionnaire
[2] the highest ranking (15) is attributed to the state valued most favourably

Table 5.3 Three analyses of variance (ANOVA) to test whether evaluations can be predicted on the basis of the difference between disease/treatment states, inter-respondent variation and inter-sample differences

	SS	df	F	R^2
No difference	229918	404		
Difference between states	92072	390	41.71*	59.95
Difference between respondents	177251	378	4.32*	22.91
Difference between samples	228363	403	2.74**	.67

* the model assuming a difference was supported
** not significant

Secondly, 2 analyses of variance (ANOVA) were done to establish whether the evaluation scores could be predicted on the basis of random variation or on the basis of the difference between disease/treatment phases, or the difference between respondents. If the first were the case, the reliability of the values would not be substantiated. In both analyses the model 'there is' and the model 'there is no' prediction on the basis of random variation was tested. Results are given in Table 5.3. As is shown, the model attributing scores to the difference between states is supported; 59.95% of the variance in the scores can be explained by the variation between states (F-statistic 41.71; critical value 95% one-sided = 1.69). In other words, the answers obtained do reflect an evaluation of different disease/treatment phases. In the second analysis the prediction of scores on the basis of inter-respondent variation was tested. The results of the analysis indicated that 22.91% of the variance in scores can be attributed to the inter-respondent variation (F statistic 4.32; critical value 95% one-sided = 1.56). Based on these analyses it was concluded that the evaluation scores cannot be attributed to random error. A third analysis of variance was done to test whether the scores of the Public Health employees and the experts were different. From the results it could be concluded that there was no systematic difference between the groups of respondents (F-statistic 2.74, not significantly different from zero). The same could be concluded from analyses done to test these differences in evaluation scores for individual health states. Therefore, we have combined the scores of both samples.

Quality adjustment of life-years

The median evaluations from the visual analogue scale have been taken as a starting point. The deviation in scores is considerable (Table 5.4) and medians are less sensitive to extreme scores than averages. Secondly, these values have been transformed into a utility score (see Material and Methods). These are given in Table 5.5. Thirdly, these utilities have been subtracted from 1 so as to establish the loss in utility in each phase. Fourthly, the losses in utility have been multiplied by the relevant durations to establish the loss in quality-adjusted time. The final results are presented in Table 5.6. As seen in Table 5.6, this breast cancer screening programme results in 259,704 life-years gained. The number of quality-adjusted life-years (QALYs) gained is 251,474. Thus, the effect 'diminishes' a bit if quality of life is taken into account.

This result is attained due to different factors. Screening attendance has a limited impact on quality of life for a large number of women during a short period (-1,790 QALYs). More important is the negative effect on women who have been diagnosed earlier as a result of the screening. This so-called 'lead-time effect' by itself is the main negative result of the programme. On the other hand, quality adjustment leads to a gain as a result of the decreasing number of women experiencing advanced phases of the disease. The screening programme would prevent chemotherapy, radiotherapy, surgery, hormonal treatment and terminal

Table 5.4 The mean and median evaluations of the disease/treatment phases (both groups are taken together), and either the mean or the median in one of the groups (whichever is the more extreme) in a favourable or unfavourable direction

		Both groups (N=27)			Variant	
		Mean	Median	s.d.	Favourable	Unfavourable
1	3 mths-1 year after mastectomy *	65	64	16.5	61	66
2	Palliative + surgery	46	41	16.4	40	49
3	Palliative + chemotherapy	36	34	17.0	34	37
4	Initial surgery	62	67	15.5	69	59
5	Palliative + hormonal therapy	47	45	19.1	37	59
6	Initial radiotherapy	60	59	15.2	60	57
7	Initial hormonal therapy	63	61	16.8	57	74
8	Initial chemotherapy	50	50	17.6	46	55
9	3 mths-1 year after breast-conserving therapy	71	74	15.5	77	68
10	Palliative + radiotherapy	43	38	16.4	38	46
11	Terminal illness	19	17	15.3	16	20
12	Screening attendance	92	94	6.9	97	91
13	Diagnostic phase	71	71	14.2	68	79
14	Disease-free > 1 year after mastectomy	77	80	12.7	82	74
15	Disease free > 1 year after breast-conserving therapy	82	83	11.5	89	77

* the order of the states in the table is the order used in the questionnaire

Table 5.5 Medians, corresponding utilities and duration of the phases

	Median	Utility	1-utility	Duration
Terminal illness[1]	17	0.288	0.712	1 month
Palliative + chemotherapy	34	0.531	0.469	4 months
Palliative + radiotherapy	38	0.591	0.419	1 month
Palliative + surgery	41	0.617	0.383	5 weeks
Palliative + hormonal therapy	45	0.663	0.337	14 months
Initial chemotherapy	50	0.717	0.283	6 months
Initial radiotherapy	59	0.803	0.197	2 months
Initial hormonal therapy	61	0.820	0.180	2 years
2 mths-1 year after mastectomy	64	0.844	0.156	10 months[2]
Initial surgery	67	0.867	0.133	2 months[2]
Diagnostic phase	71	0.895	0.105	5 weeks
2 mths-1 year after breast-conserving therapy	74	0.914	0.086	10 months[2]
Disease-free > 1 year after mastectomy	80	0.947	0.053	Life expectancy
Disease-free > 1 year after breast-conserving therapy	83	0.960	0.040	Life expectancy
Screening attendance	94	0.994	0.006	1 week

1. the order is according to the utility calculated, 2. see material and methods

Table 5.6 Number of women, multiplied by the duration of states (woman-years), in the situation with (S^+) or without (S^-) breast cancer screening and the correction as a result of quality adjusting the effect. No discounting

	With breast cancer screening		Without breast cancer screening			
	Number of women	Woman years (S^+)	Number of women	Woman years (S^-)	Difference in years (S^+-S^-)	Quality adjustment[2]
Terminal illness	271,815	22,561	288,862	23,976	- 1,415	1,008
Palliative + chemotherapy	271,815	90,514	288,862	96,191	- 5,677	2,665
Palliative + radiotherapy	271,815	22,561	288,862	23,976	- 1,415	593
Palliative + surgery	271,815	27,725	288,862	29,464	- 1,739	666
Palliative + hormonal therapy	271,815	317,118	288,862	337,005	- 19,888	6,700
Initial chemotherapy	31,934	15,967	31,934	15,967	0[1]	0
Initial radiotherapy	388,937	64,952	378,634	63,232	1,721	- 340
Initial surgery	580,847	97,002	571,620	95,461	1,541	- 205
Initial hormonal therapy	188,742	377,484	198,313	396,625	- 19,141	3,449
2 mths-1 yr after mastectomy	309,271	257,623	320,188	266,717	- 9,094	1,416
Diagnostic phase	1,244,643	119,486	1,257,049	120,677	- 1,191	125
2 mths-1 yr after breast cons. therapy	255,687	212,987	234,521	195,356	17,631	- 1,519
Disease-free > 1 year after mastectomy	3,203,247	3,203,247	3,169,573	3,169,573	33,673	- 1,800
Disease-free > 1 year after breast-cons. therapy	3,372,980	3,372,980	2,907,354	2,907,354	465,626	- 18,512
Screening attendance	15,768,572	299,603	0	0	299,603	1,790
Correction for double counting with respect to hormonal treatment and disease-free years						- 686
Total correction for quality of life (C)						- 8,230
Effect of breast cancer screening programme on life-years gained						259,704
Effect of breast cancer screening programme on quality-adjusted life-years gained						251,474

[1] initial chemotherapy is, usually, not given to postmenopausal women. We have, therefore, predicted no changes for the screening programme for women aged 50-70

[2] see Appendix 5.2 for computation

Table 5.7 Variants in the computation of the impact of the breast cancer screening programme on quality-adjusted life-years (sensitivity analysis) and the % difference between life-years gained and QALYs gained

	Life-years gained	Quality adjustment	QALYs gained	Difference (%)
Median/0% discounting	259,704	- 8,230	251,474	- 3.2%
Median/5% discounting	60,205	- 3,361	56,844	- 5.6%
Mean/5% discounting	60,205	- 5,600	54,605	- 9.3%
Favourable evaluation[1]/5% discounting	60,205	1,950	62,155	3.2%
Unfavourable evaluation[1]/5% discounting	60,205	- 11,837	48,368	- 19.7%

[1] see values Table 5.4

illness for 17,000 women. As a result of this, 11,630 QALYs are gained. The effectiveness of the screening programme 'decreases' with 3.2% when taking quality of life into account. We have considered a small loss in utility in all disease-free years. If one assumed that, for instance, 5 years after initial treatment quality of life would equal the quality of life of the normal female population again, the number of QALYs gained would probably be more than the number of life-years gained.

Variants have been computed to establish the sensitivity of this result to the values included in the analysis (Table 5.7). If a 5% discount rate for effects in the future were included, 60,205 life-years and 56,844 quality-adjusted life-years would be gained. In this variant the effect is lowered by 5.6% if quality of life is taken into account. A 5% discount rate is used in all other variants presented. In the second variant, the mean value is inserted into the model instead of the median value. In this case, the screening programme would be slightly less effective: 54,605 quality-adjusted life-years would be gained; quality adjustment reduces the effectiveness by 9.3%.

Thirdly, sensitivity was examined by inserting an extreme favourable or an extreme unfavourable value for each phase. These were selected by taking the most favourable and most unfavourable mean or median value within the 2 groups (as shown in Table 5.4). In one variant, we inserted the high value for all those phases that are expected to be prevented by screening and the low value for those phases that are expected to occur. For example, we used the value 74 for initial hormonal therapy and 20 for terminal illness (screening will prevent both), but 91 for screening attendance and 59 for initial surgery (screening will increase the number of women in these phases). This is a so-called pessimistic or unfavourable variant. In the other variant, we used the opposite values, which would lead to a favourable variant. In this favourable variant the programme leads to a gain of 62,155 quality-adjusted life-years (3.2% more than without quality adjustment). In the unfavourable variant the effectiveness is lower; 48,368 quality-adjusted life-years are gained by the screening programme. This is 19.7% lower than without taking quality adjustment into account.

Discussion

In this study, the effectiveness of a nationwide breast cancer screening programme adjusted for quality of life is investigated. The programme is predicted to be effective in gaining life-years or saving breast cancer mortality: the screening programme in the Netherlands organized for women in between 50 and 70 years of age at 2-yearly intervals would in the long run result in a gain of 260,000 woman-years. The question is whether taking into account the impact of the programme on quality of life would alter the results. In the variant without discounting, the number of quality-adjusted life-years gained is 251,000. The difference with the estimation without quality adjustment is 9,000. In the least

quality-effective variant the number of QALYs gained is 48,000, which implies 12,000 'less effective' than without quality adjustment. The reduction to be attributed to quality adjustment lies in between +3.2% and -19.7%. Even inserting extreme values for quality of life does not lead to large differences in the predictions. All things taken together, the most plausible result is that this screening programme is about 8% (19.7 + 3.2/2 = 11.45; 3.2 - 11.45 = -8.25) less effective with quality of life adjustment than without it. From these findings it may be concluded that the predicted effectiveness of a breast cancer screening programme is not substantially influenced by the changes in quality of life of screening participants and patients. This result is very similar to the one suggested by Forrest (1986). Quality of life is, therefore, not a major issue in the decision to undertake a large-scale breast cancer screening programme.

For us, this conclusion has been unexpected. Articles emphasizing possible negative consequences of breast cancer screening have been published (Skrabanek, 1988; Roberts, 1989). Ellman et al. (1989) found a slight increase in worries and anxieties after participation. Gram et al. (1990) described negative effects in some, but not in most women with a false positive result. These reports are most often limited to the disadvantages of screening participation in the short run. However, as seen in Table 5.6, effectiveness is influenced in more ways in the different phases described. In our study, too, attendance is considered to have a negative impact on the quality of life.

Moreover, as breast cancer is detected earlier, more women experience the disease and suffer from the effects of diagnosis and initial treatment during the so-called disease-free periods. However, less patients have to undergo palliative treatment as a result of the improvement in prognosis. This results in a gain in quality of life for women who would, without a screening programme, have lived with advanced disease. The expected decrease in hormonal therapy, as adjuvant treatment for post-menopausal women or as initial treatment in the 70+ age group, is responsible for 17% of the positive quality of life adjustment. Other changes appear to be of minor importance. The increase in the total number of primary operations or radiation treatments reflects only 2% of the negative adjustment. Although the taking of biopsies in women who appear to have a benign lesion has a negative effect on quality of life, it has no important effect in our analysis. The increase in biopsies due to false-positive referrals is often regarded as a serious problem of mass screening. One has to realize, however, that the number of referrals of women outside the programme is expected to diminish after initiation of the programme. Our analysis shows that this results in a small positive effect on quality of life.

Finally, there has been much discussion on the different impact of breast-conserving therapy and mastectomy on the quality of life (Levy et al., 1989; Fallowfield et al., 1990; Kiebert et al., 1991). In our study, Public Health employees and breast cancer experts consider women who underwent breast-conserving therapy to have a 7% higher utility in months 3-12 than women who underwent mastectomy and a 1% higher utility after the first year. Therefore, the

shift from mastectomies towards breast conserving therapy implies a small gain in QALYs. On the other hand, the additional number of breast-conserving therapies (lead time) results in a loss in QALYs, as compared to the situation of no screening. Although our method has been complex and laborious, the results illustrate the necessity of taking into account all disease and treatment phases when discussing the impact of screening.

The method used, describing health states on the basis of aggregated published data, is logical if no empirical study can be done, but unusual. Usually, health states descriptions are less elaborate (Llewellyn-Thomas et al., 1984). A simple framework could not be used in our study as the literature is not organized in this way and because important side-effects of treatment in breast cancer would not be covered. With our method, almost all respondents could complete the questionnaire and reliability was satisfactory. However, the method merits some discussion. Most experts in breast cancer did not forward comments, whereas the employees of the Department of Public Health made a number of remarks. This might imply either that they were more critical, or that the method is more feasible for respondents familiar with the situation to be evaluated. As it may be necessary to approach non-experts for other cost-utility analyses, the comments given may be useful for future studies.

Some respondents found completing the questionnaire difficult for several reasons: (1) the difficulty of integrating 3 dimensions into 1 global evaluation; (2) the broad range of prevalences (it was suggested to present the median of the different studies); (3) the large number of concomitants given (which might be due to the specific treatment described or to the extensiveness of the literature); (4) the incomprehensibility of certain side-effects for lay judges. It was suggested that the more complex the descriptions were, the more one would judge the state on the basis of his/her intuition. These comments imply that the questionnaire would have to be simplified if given to patients or a sample from the lay population. The use of medical terminology would have to be avoided and the number of categories used within each dimension reduced. For example, for the physical dimension one might present the deterioration in daily activities and the main side-effects, even without specified information on prevalence. This might lead to an oversimplified picture of the breast cancer patients' situation but might also make a questionnaire based on the aggregation of data from the literature feasible for other study samples. The latter would be important as it has been shown (Slevin et al., 1990) that patients and experts have different preferences when evaluating cancer and cancer treatment.

6 Changes in use of breast conserving therapy in years 1983-2000

Introduction

Treating breast cancer patients has changed rapidly in the last decade. Several trials have led to a better understanding of the possible results of different types of primary treatment. Breast conserving therapy, limited surgery followed by high dose radiation, has shown to be as effective as modified radical mastectomy for most operable patients (Veronesi et al., 1981). For mastectomy patients, routine postoperative radiotherapy now usually is considered unnecessary treatment (Edland, 1988; Harris and Hellman, 1988). One would assume that treatment practice is influenced by the information from these trials, but empirical data at (sub)national level is scarce in most countries (Farrow et al., 1992).

At the same time, countries have started to implement nationwide breast cancer screening, as trials have shown that a reduction in mortality from breast cancer can be expected for screened women aged 50 and over. The impact on the number and type of surgical procedures and radiotherapy will be large, both due to the temporary increase in the number of women detected with breast cancer and due to the increase in early cancers (de Koning et al., 1990c). Bottlenecks in capacity are expected, but there is lack of empirical data to support this hypothesis. Waiting lists for treatment might endanger the effectiveness and organisation of a screening programme.

Having been uncertain especially about the development of breast conserving therapy (outside a screening programme), we have analyzed data on the actual number and type of primary treatments for breast cancer in the Netherlands in the years before the start of nationwide screening. In this chapter, we present (1) the trend in primary treatment for the period 1983-1990, (2) predictions for the near future and (3) the expected influence of breast cancer screening. Emphasis is being put on surgery and radiotherapy, and especially on breast conserving therapy versus mastectomy, and on postoperative radiotherapy

---

de Koning HJ, van Dongen JA and van der Maas PJ. Submitted

after mastectomy. The findings are compared to the scarce international data from the United States and may serve for planning and evaluation of both breast cancer treatment and screening.

Material and methods

Two independent sources provided data on the present use of surgery and/or radiotherapy for women with breast cancer. The number of different types of surgical procedures was determined by analyzing 160,000 records on hospital admissions in the Netherlands for women with breast cancer and/or breast surgery (Centre for Health Care Information). The available information contained all data for the years 1983-1984 (coverage 97-99%) and 83% for the years 1985-1988 due to a clinicians' embargo (see Appendix 6.1). The records included the patient's age, rank of admission, types of surgical procedures, number of nursing days, diagnoses at discharge and residence after discharge. Data on radiation treatment are not centrally registered in the Netherlands. Therefore, a questionnaire was sent to all 20 Dutch radiotherapy departments, concerning the number of female breast cancer patients having had radiotherapy in the years 1986-1988. Sixteen departments answered, and additional information on the total patient population of all departments enabled us to extrapolate figures for the whole population.

The clinical (age specific) breast cancer incidence was based on the national registry of first hospital admissions and on an additional 15% of breast cancers in the 70+ group, assumed to be underregistered in incidence figures based on hospital admissions (e.g. due to primary hormonal treatment). Except for one regional cancer registry, there were no incidence data yet available from a complete national cancer registry before the year 1992, but the recent data confirm the 15% assumption by showing a lack of pathological confirmation of breast cancer in the 75+ group of 14% (Netherlands Cancer Registry, 1992).

All primary breast cancer treatment could be divided into treatment for women with ductal carcinoma in situ (DCIS) (treatment by either local excision, local excision plus radiotherapy, or total mastectomy) and for women with invasive carcinoma. The latter group is treated by either breast conserving therapy with an external booster, breast conserving therapy with an iridium implant, total mastectomy, total mastectomy with postoperative radiotherapy (all usually with axillary dissection), a combination of treatment modalities for stage IIIB tumours, or by primary hormonal treatment (Tamoxifen); see also Table 6.1. Usually, breast conserving therapy consists of local excision, axillary dissection, postoperative radiotherapy on the breast and/or regional lymph nodes and a booster dose on the original tumour site.

For these treatment options, and for adjuvant systemic treatment, we estimated the number of women treated in the Netherlands in 1990. The year 1990 was chosen for practical reasons, since it coincided with the start of the

screening programme in the Netherlands. Combining these numbers with the clinical incidence and stage distribution of breast cancer, and the treatment guidelines, we predicted the chance of being treated by each modality for 4 groups of women: those with clinically diagnosed ductal carcinoma in situ, those with invasive carcinoma smaller than 10 mm in diameter, those larger than or equal to 10 mm but smaller than 20 mm, and invasive carcinomas of 20 mm and more. The axillary lymph node status corresponding with these stages was taken into account (see Appendix 6.2).

Additionally, the existing treatment guidelines, international literature and a 10 questions-survey amongst 40 Dutch breast cancer experts were used to estimate the possible trend in primary treatment in the nineties if there would be no screening programme. Three variants were distinguished in the predictions for primary treatment in 1990-2000 without screening: one in which the number of women treated by breast conserving therapy will still follow the analyzed rise up to 1995. In this scenario, we assume -from the data analyzed- that there is still diffusion of technology in these years. A second variant in which the proportion will remain at the 1990-level (treatment criteria are stable) and will only be influenced by the increase in breast cancer incidence, and a third intermediate variant.

All variants are also analyzed in the situation with screening. In that situation, there is a shift towards smaller sizes and towards a more favourable lymph node status (Tabár et al., 1992b). Even within each tumour size category, women with screen-detected cancer have a lower percentage of axillary lymph node metastases (de Koning et al., 1990c). Lymph node and distant metastases status for screen-detected cancers was determined from the Utrecht and Nijmegen projects; for clinically diagnosed cancers from the Utrecht "non screened" group (Collette et al., 1984; Verbeek et al., 1984). Chances of treatment modality per tumour size category were predicted too for women with screen-detected breast cancer. The MISCAN breast cancer model was used to predict the number of women in the 4 tumour categories with and without breast cancer screening (van Oortmarssen et al., 1990b). Key-parameters of the model were derived from results of the HIP-analysis (van Oortmarssen et al., 1990a), and from a new analysis of all results from the Dutch screening trials in Nijmegen and Utrecht (de Koning et al., 1991). In this chapter, the influence of screening on the change in treatments will be shown for the actual Dutch screening policy of 2-yearly screening for women aged 50-70 and an attendance rate of 70% on average, both at national and local level. The programme started in 1990 and is planned to be fully implemented around 1994.

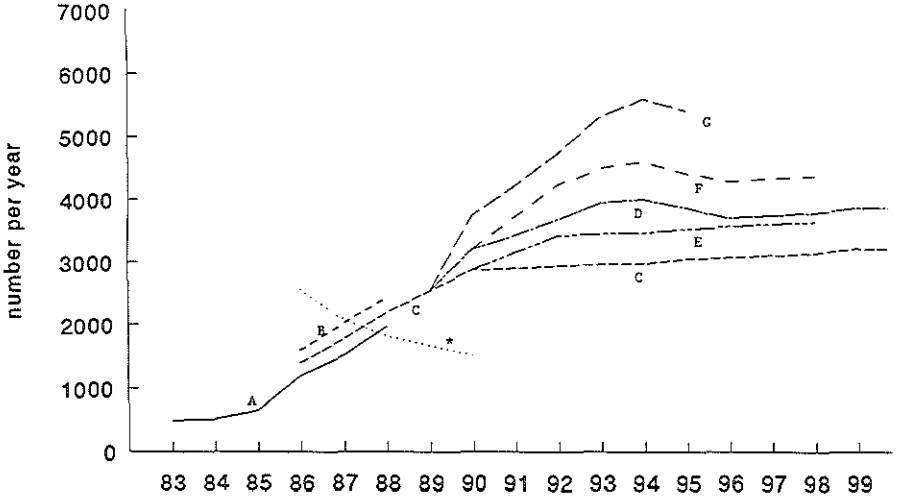

Figure 6.1 Yearly number of patients with invasive breast cancer treated by limited surgery followed by radiotherapy (breast conserving therapy) in the Netherlands, 1983-2000.
a) according to national hospitals' registry b) according to radiotherapy departments
c) numbers without screening; 1990-level (no trend)
d) numbers with screening (base level c)
e) numbers without screening; 1992-level (including time trend)
f) numbers with screening (base level e)
g) maximum number with screening (including time trend and expert opinion)
* number of patients treated by radiation after mastectomy

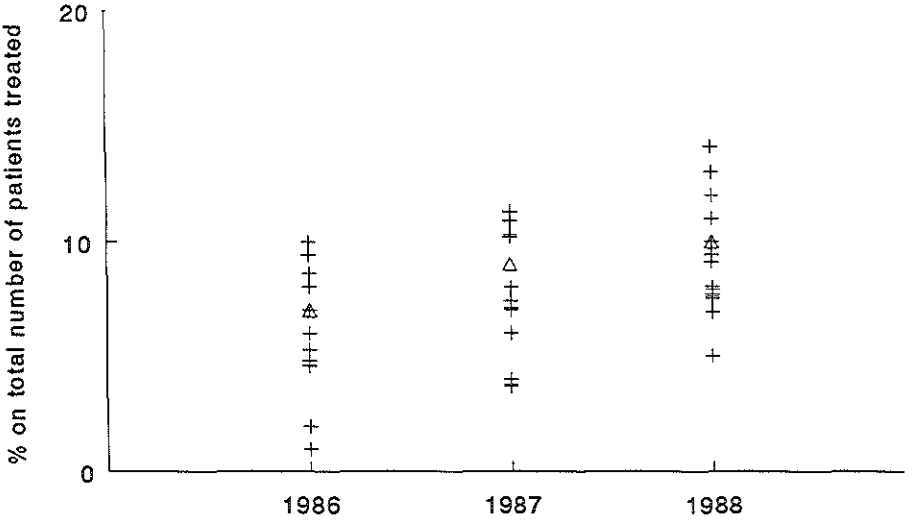

Figure 6.2 Breast cancer patients treated by radiation after limited breast surgery (breast conserving therapy) as registered in each radiotherapy department. The Netherlands 1986-1988. Triangle = mean % of all departments. Percentage expressed in annual number of women with breast conserving therapy on total annual number of all patients radiated (not only breast cancer). Total on average N=23,350; total numbers in 1987 or 1988 differ only 1% with 1986.

Results

Primary surgery and/or radiotherapy 1983-1990

The main treatment choice for women with operable invasive breast cancer is between conservative or more radical treatment. There is a relatively small absolute increase in breast conserving therapy in the years 1983-1985, caused primarily by the yearly increase in breast cancer incidence of 0.5-1%. However, a rapid and steady increase from 1985 onwards is visible, ultimately leading to a number of almost 2,000 registered hospital admissions for limited breast cancer surgery in 1988 (Figure 6.1, line a).

The independent survey amongst radiotherapy departments confirms the rapid and steady increase in the number of radiation treatments as part of breast conserving therapy for the years 1986, 1987, 1988 (Figure 6.2). In sixteen departments 1330, respectively 1679 and 2015 treatments were recorded. If these centres are assumed to be a representative sample of all 20 centres, the estimation for the whole of the Netherlands would be 2400 treated women in 1988 (Figure 6.1, line b). About 15% of these treatments is followed by an iridium booster, the others with an external booster. Since there is not a uniform radiotherapy register throughout the country, and the fact that carcinoma in situ was not separately mentioned in the survey (in contrast to the hospital registration), we may conclude that both sources on surgery and radiotherapy are in agreement with each other. We used the average estimation for the present situation (line c).

Assuming this trend has continued in 1989 and 1990, we expect that approximately 2900 women had been treated conservatively in 1990, which is 36% of all women with newly diagnosed invasive breast carcinoma in the Netherlands. The experts thought that 35% of the women were treated conservatively, so it seems reasonable indeed to assume this trend up to 1990.

Except for a small proportion of women treated by primary hormonal treatment, or treated for locally advanced disease by e.g. primary radiotherapy, the remaining women are treated with breast ablation, mostly by modified radical mastectomy. Whether this should be followed by radiotherapy on a routine basis is a hot dispute. Fifteen radiotherapy departments register a clear decrease in the number of women treated by radiation after a mastectomy from 2126 in 1986 to 1733 in 1987 and to 1518 representing respectively 11% of all women treated by radiation, and in the years after respectively 9% and 8%.

The decrease in 1987-1988 is less evident than in the previous 2 years, and it remains difficult to establish a certain trend. Nevertheless on the basis of these data, we assume a total of 1525 postmastectomy irradiations in 1990, which corresponds to 39% of the breast cancer patients treated by mastectomy (Figure 6.1, line *). Table 6.1 summarizes the estimated numbers of treatment modalities applied in the Netherlands to newly diagnosed female breast cancer patients in 1990.

Table 6.1 Estimated numbers and percentages of new breast cancer treatments in the Netherlands in 1990 without mass screening

	Treatment	numbers	%	main sources
1.	Primary treatment DCIS			
	- local excision	115	30%	hosp.; trial
	- local excision + radiotherapy	115	30%	hosp.; trial
	- total mastectomy	150	40%	hosp.
		380	100%	
2.	Primary treatment invasive carcinoma			
	- breast conserving therapy, external booster	2400	30%	hosp.; radioth.
	- breast conserving therapy, iridium implant	475	6%	hosp.; radioth.
	- mastectomy, no radiotherapy	2400	30%	hosp.; radioth.
	- mastectomy and radiotherapy	1525	19%	hosp.; radioth.
	- primary Tamoxifen	375	5%	model
	- treatment locally advanced disease (IIIB)	400	5%	registry
	- stage IV treatment	450	5%	registry
		8025	100%	
Total female breast cancer incidence		8405		
3.	Adjuvant Tamoxifen	1750	26%x	survey
	Adjuvant chemotherapy	750	11%x	survey
		2500	37%x	

x % related to invasive carcinomas, except stages III_B and IV and already primary Tamoxifen

Possible developments 1990-2000 without mass screening

Further important developments are to be expected in the near future on at least the criteria for performing breast conserving therapy, and criteria for radiotherapy after mastectomy. The experts envisage an increase in the proportion of female breast cancer patients treated by breast conserving therapy in the next 5 to 10 years from the present 35% towards 50% of all female breast cancer patients, if there would be no screening programme. This would mean a continuation of the line as seen in 1985-1988. They foresee room for further diffusion of technology throughout the country, due to more and better multidisciplinary treatment protocols, change of treatment criteria, stronger demand of women, and/or a general earlier diagnosis of breast cancer.

Surely, this expected increase will have to level off, as there will be a moment that all patients who are eligible for and want to have conservative surgery are actually treated in this fashion. In practice, the upper limit of tumour size tends to decrease due to unsatisfactory cosmetic results in larger tumours.

When taking into account the presence of extensive ductal carcinoma in situ around the invasive part, refusal of women or doctors and tumour size/breast volume-ratio, we assume that 50% of all women with operable invasive breast cancer (but excluding stage IIIb, not stage IV, not primary Tamoxifen) might be treated conservatively. In fact, it represents approximately 42% of all newly diagnosed invasive breast cancer patients per year, which would thus be reached around 1992 (medium variant, Figure 6.2, line e).

The routinely radiation of the internal lymph node chain after a mastectomy for centra-medial tumours is perhaps one of the most discussed breast cancer treatments at present. Almost all experts thought it very likely that the percentage of women that would undergo radiation treatment after mastectomy would decrease in the next 5 to 10 years (without considering a screening programme). If we would assume that only the presence of axillary lymph node metastases remains the criterium, and not the localisation of the tumour itself, the number of women radiated postoperatively is expected to decrease in 1994 to 37% of all women treated by mastectomy.

Influence mass screening 1990-2000

These estimates could be strongly influenced by the introduction of mass screening. Implementation will always result in a temporary increase in the number of persons detected with the disease. Most outspoken is the increase in detected cancers at the first (prevalent) screening. The start of the Dutch breast cancer screening programme in 1990 will result in a rise in the number of newly diagnosed cases, with a maximum 17% (=1450 cancers) increase in the year 1993. From 1996 onwards, the yearly total number of diagnosed breast cancers will be 3.5% higher compared to the situation without screening (de Koning et al., 1991).

The main consequences of earlier diagnosis for treatment practice of these women are shown in Figure 6.3. In these calculations, we assumed that the relative proportions of treatment as shown in Table 6.1 would also apply in the years after to clinically diagnosed cancers (no increasing or decreasing trend). Since almost all treatments depend at least on tumour type, size and/or lymph node metastases, we may expect differences to occur on all aspects. Most important is the shift from mastectomy to breast conserving therapy due to the detection of cancer in earlier stages. At national level, we expect a steady increase in breast conserving therapy up to a 34% increase (+ 1025) in 1994. Hereafter, the increase stabilizes at approximately 21% (+ 640 per year), compared to the number of expected treatments without screening each year (see also Figure 6.2, line d). The number of women treated by mastectomy will ultimately be decreased by 9% (- 370 per year). The decrease in mastectomies does not equal the increase in breast conserving therapy, as a result of the higher incidence and the detection of early lesions due to screening. All primary treatments that will eventually show a decrease due to the more favourable stage distribution of screen-detected cancers, do still show an increase in the build-up

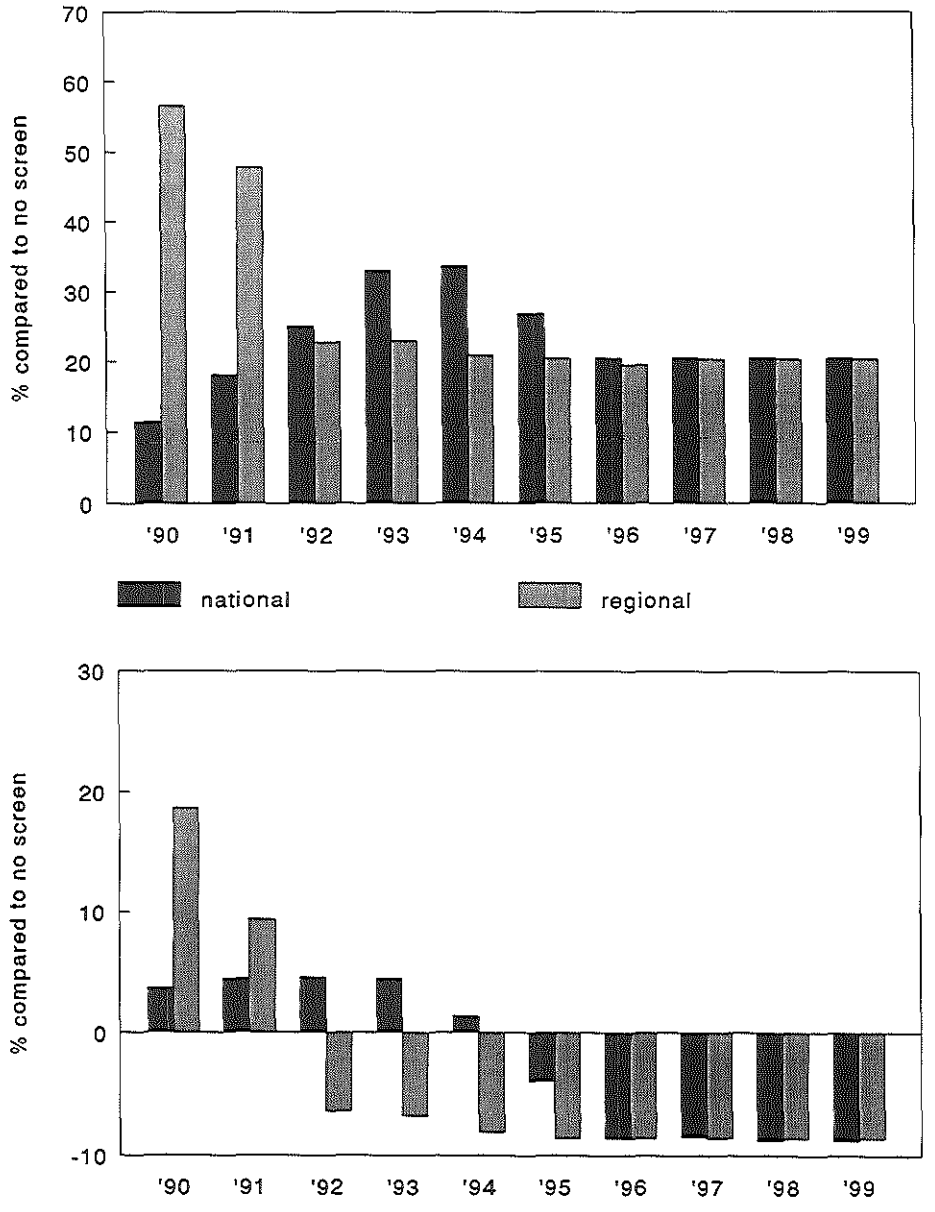

Figure 6.3 Yearly change (%) in breast conserving therapy (top) and breast ablation (bottom) at local and nationwide level (years 1990-1999), due to screening of women aged 50-69 every 2 years. % compared to the situation without mass screening in that year.
National level: gradual build-up during 5 years.
Regional level: consequences surrounding 1 screening unit with approximately 10,000 screens per year (2-year build-up).

period of screening. This is a result of both the temporary increase in women diagnosed with cancer and the less favourable stage distribution in which these cancers are detected at the first screening round, compared to subsequent rounds.

At the local level, the temporary increase in the number of breast conserving therapies may be much more dramatic, since the regional build-up periods of the screening are even shorter than the nationwide build-up. The figures also show the percentual changes for the region surrounding a centre with 10-12,000 screens a year. If not adjusting for specific regional differences in treatment, the number of women who undergo breast conserving therapy increases by 50% in the beginning, only it lasts less longer.

Discussion

Increase in breast conserving therapy

The number of breast cancer patients treated conservatively has rapidly increased in the Netherlands since the mid-eighties. It is clear that earlier estimates based on 1983-1985 admission data only are outdated due to this change (de Koning et al., 1990c). The increase is, however, not never-ending, as trials have shown breast conserving therapy to be as effective as modified radical mastectomy only for localized small cancers, and several studies have shown definitive contra-indications (Bartelink et al., 1988; Holland et al., 1990). Data from the Netherlands Cancer Institute suggest that approximately 55% of all operable cases are appropriate candidates for breast conserving therapy (Hooning et al., 1991). The present analysis shows two important aspects on quality of breast cancer treatment in the Netherlands: breast conserving surgery is always followed by radiation treatment, and the actual number of women treated conservatively is rapidly approaching the number one would forecast on the basis of oncologic data and protocols. Still, there is room for diffusion of the technology on the basis of oncologic criteria and the published treatment trial results.

The extent to which breast conserving therapy has been adopted in other countries is rarely described. Only recently, information came available from parts of the United States (Farrow et al., 1992). In 9 areas, the percentage of white women (18,399 in 4 years) with localized invasive breast cancer treated by breast conserving surgery increased from approximately 22% in the period 1983-1984 to 33% in 1985-1986, which seems an extremely early increase in this technology. It is known that the publication in 1985 of the results from the US randomised clinical trial of breast conserving surgery led to a temporary peak in 1985 in some districts, followed by a return (Lazovich et al., 1991). A publication on a much larger group of 41,680 breast cancer patients treated in nine (other) areas in the United States of America in 1988 - 31% of all cases in that year - reveals only 25% of these patients being treated by partial mastectomy (Osteen et al., 1992). It is apparent that there are strong regional differences in the United States, and that the often heard proposition that Europe follows some

time after the developments seen in the United States seems not sustained by the data on breast conserving therapy in the Netherlands. The percentage is higher in the latter country in recent years. Secondly, although there was a nationwide temporal trend towards the increasing use of radiation after breast conserving surgery in the United States, it was quite different from the Dutch situation where in principle all women received radiation treatment after limited surgery.

Future treatment changes

Uncertainties focus more or less on the upper ranges of our results. Of course the future practice in radiation treatment after mastectomy seems the most unknown factor (Haybittle et al., 1989; Levitt, 1988). One may say that we are unable to predict these numbers in the situation with screening in the next few years. It is likely that even less radiation treatment after radical mastectomy will be given, but that it will follow a decreasing line in both women with screen-detected breast cancer as in clinically diagnosed cancers. Screening might accelerate the discussion on this type of treatment. The influence of new treatment trials or new results of ongoing trials can of course not yet be taken into account. Most of these apply to treatment schedules in which a less major intervention is compared to the present treatments described in our results: e.g. omission of a booster dose (EORTC, 1989), reduction in radiotherapy sessions and dosage (The Uppsala-Örebro Breast Cancer Study Group, 1990) and DCIS-trial (van Dongen, 1989). Even the influence of mass screening on breast conserving therapy becomes different if we assume that the results of treatment trials had not yet influenced some clinicians' and/or women's decisions. If we assume that the increase in breast conservation would continue until 1995, mainly as a result of a general broadening of criteria, we might expect that both women with breast cancer diagnosed clinically and women with screen-detected breast cancer will be treated more frequently by conservation in the nearby future (Figure 6.1, line g). New insights in treatment criteria are therefore especially important for countries with mass screening, and will show the difficulties in predicting treatment changes due to screening.

There has been no literature on the predicted treatment consequences of implementing screening, which seems a logic result of the lack of national data on applied treatment modalities for breast cancer in most countries in the situation without screening. A very important asset of this analysis are the predictions of treatment changes before a screening programme has started, to be used to estimate the needed capacity. The initial increases in breast conserving surgery, mastectomy and DCIS-treatment with 1500 in 1993 as a result of screening may be relatively small on the total number of surgical procedures, but represents more than 7.5% of breast surgery. Together with the additional increase in biopsies, the early years of screening will result in an additional average 15-20% increase in breast surgery at national level.

Changes for radiotherapy services seem even more important to forecast, as the investments for possible new machinery take 5 to 7 years. Considering

both the increase in primary radiation treatment due to screening and the decreases in the treatment for advanced disease (de Koning et al., 1992a), the total number of radiotherapeutic sessions increases by 5% in 1994, which is 22% of all sessions for breast cancer. In the steady state, only 2.5% is needed additionally, which means 2 more machineries at a total of 76. In the early years, this increase in workload has to be dealt with within the existing facilities.

We will follow the actual changes in the next years, both at national and regional level. Important regional differences when screening is implemented should be monitored and findings, together with the stage distribution, may function as a guide on whether screened women who appear to have breast cancer are treated adequately. The first published results on changes in breast conserving therapy due to screening in the Enschede-region support our analysis at the local level (Boekema et al., 1992). Although numbers are quite small, breast conserving therapy increased by 50-70% compared to the situation without screening, if the observed national trend is to be applied in this region too, and when adjusting for a 15% higher detection rate than average. 61% of operable screen-detected (invasive) breast cancer patients underwent conservative surgery, compared to 43% of patients diagnosed outside the screening programme, all treated primarily in one hospital. Recent data from another large region, the IKA-region, shows approximately 40% of operable (invasive) screen-detected patients to undergo conservative surgery (1990-1992), but there is a wide variation between hospitals (17-71%). In half of the hospitals there are relatively more women undergoing mastectomy than women treated by breast-conserving therapy (van Velthoven, personal communication 1993). But again, the difference induced by screening is striking, as in 1989 approximately only 30% of the breast cancer patients was treated conservatively (without a screening programme). Detailed information on treatment changes in respect to mass screening are also crucial for estimating the influence on quality of life, and for adjusting the expected future breast cancer mortality reduction, if necessary. Especially the increasing use of adjuvant systemic treatment and consequently mortality reduction may interact with the achievable reduction due to earlier detection of cancer.

In overview, the rapidity of change towards the use of breast conserving surgery comparable to the change of radical mastectomy to modified radical mastectomy during the 1970s, will be enhanced by the additional increase due to mass screening.

We would especially like to thank the Dutch Society for Radiotherapy for their voluntary contribution to the study, and in particular Prof.dr. G.M.M. Bartelink and Prof. J.W.H. Leer. All radiotherapy departments and persons that completed and returned the questionnaires are thanked. Dr. P.C.M. van Velthoven, medical auditor IKA-region, is thanked for supplying regional data. Acknowledgments are given to Mrs A.E. de Bruyn and G.J. van Oortmarssen.

PART III

7 The cost-effectiveness

Introduction

Breast cancer is a major cause of death among women in Western countries. Mammography can detect this disease at an early stage. Large trials in different countries have shown that mammographic screening of women aged 50 years and over reduces breast cancer mortality. Screening of women under 50 has considerably less impact.

Governments in several countries are considering the introduction of nationwide breast cancer screening. Implementation is already being carried out in Sweden, Finland and the United Kingdom (UK Working Party on Breast Cancer Screening, 1987; Day and Miller, 1988). In 1986, the Dutch Ministry of Welfare, Health and Cultural Affairs invited a research group to perform a cost-effectiveness (CE) analysis, before deciding on the implementation of such a programme. Likewise, in the United Kingdom the Forrest Committee was requested to perform a CE analysis (Forrest, 1986).

The reasons for performing a CE analysis before introducing mass screening are more or less self-evident. Under all circumstances mass screening will be a costly service. Small changes in the organization or in the invitation schedule may have considerable consequences, for both costs and effects. The purpose of the study is to estimate the costs of mass screening, assessment and treatment, and the health effects of mass screening (primarily mortality reduction). Costs and effects are related by computing CE ratios. This chapter presents the main results of the first preliminary analyses on cost-effectiveness ratios. More recent research has updated the different absolute values, also on expected breast cancer mortality reduction (see chapter 8). In this chapter, emphasis has been put on the structure of CE analysis and the influence of screening interval in the age group 50-70.

Chapter based on van der Maas, de Koning et al., Int J Cancer, 1989

Material and methods

The costs of mass screening were estimated by using detailed organization charts. Cost accounts of the Dutch breast cancer screening trials formed the empirical basis for the required man-power and facilities for a national programme. In order to estimate costs of assessment of suspected breast cancer cases and costs of treatment, "national average schedules" for assessment and treatment were compiled. These schedules were based on data from central registries and on interviews with specialists in the field of breast cancer. The cost estimates for the medical procedures are based on tariffs and have been derived from file studies. The main constituent of the direct costs of treatment is the cost of hospital nursing. These costs are based on an analysis of data concerning hospital admissions for breast cancer and/or operations on the breast during 1983-1985, supplied by the Dutch Centre for Health Care Information. The cost for treatment of advanced disease was based in this chapter on a reported file study of 52 patients with breast cancer metastases (de Waard et al., 1986). The costs are expressed in US dollars, using purchasing-power parities (OECD, 1987).

The impact of mass screening on incidence, stage-distribution and survival has been estimated here from the results of trials in New York (Shapiro et al., 1982), Sweden (Tabár et al., 1985), Nijmegen (Verbeek et al., 1984) and Utrecht (Collette et al., 1984). A detailed re-analysis of the Dutch data (Nijmegen and Utrecht) provided the basis for more refined calculations.

All computations were done using the actual Dutch population structure and its predicted future development. In this way, realistic estimates were obtained for the expected number of women screened, the required diagnostic and therapeutic capacities, costs, and mortality reduction. However, the effect estimates were also produced for a birth cohort, in order to obtain results which do not depend on a specific dynamic population and that will be comparable to studies from other countries. Data on breast cancer incidence and survival were derived from the results of the Dutch trials, from hospital statistics and from the international data (Adami et al., 1986). Data on mortality were obtained from the Dutch Central Bureau of Statistics. The age-specific incidence of pre-clinical, but screen-detectable breast cancer is based on incidence data for clinical breast cancer, corrected for the mean duration of preclinical disease.

Four alternative screening programmes are investigated, as shown in Table 7.1. The programmes are all restricted to the age-group 50-70 years, in accordance with the recommendations of the Health Council (Gezondheidsraad, 1987). The main difference between the programmes is the number of invitations per woman and thus the length of the screening interval. Results will mainly be presented for alternative II which involves 10 screens per woman with intervals of 2 years. This alternative has been proposed by 2 Dutch working parties and will be applied in the Netherlands. Based on the experience from the Dutch trials, the attendance rate for all alternatives was set at an average of 70%, decreasing from 75% at age 50 to 65% at age 70.

Table 7.1 Characteristics of 4 alternative programmes (I-IV) of mass screening for breast cancer in the Netherlands

Characteristics	Alternative			
	I	II	III	IV
Age group	52-68	51-69	50-69	50-69
Interval	4 yr	2 yr	1 1/3 yr	1 yr
Number of invitations per woman	5	10	15	20
Network completed	1994	1995	1997	1999
Attendance rate	70%	70%	70%	70%
Number of screens (x 1,000) in year of completion of network	260	525	825	1,150
Number of screens in 2015 (x 1,000)	385	770	1,150	1,535

Due to limitations in the personnel training capacity and for reasons of organization, it will take several years to complete a nationwide screening network. Furthermore, the required expertise will have to be spread progressively from existing units, in order to guarantee a high quality standard. The size of the network, and thus the year in which it will be completed, depends on the annual number of screens. After the completion of the network, the total number of women screened will continue to increase due to the aging of the population during the next decades.

All computations were carried out by using the computer simulation package MISCAN (Habbema et al., 1987). This package has been developed for analyzing and reproducing the observed results of screening projects and for predicting the future effects of alternative screening programmes. The disease model is based on a 3-stage division of the development of invasive breast cancer (stages reflect the size of the tumour). Five percent of the invasive breast cancers were initially assumed to be preceded by a screen-detectable ductal carcinoma in situ (DCIS), for which 100% progression was assumed. These and other parameters for the disease model are estimated from the results of the Dutch projects. The mean duration of preclinical disease appears to be age-dependent, and increases from less than 2 years for women under age 45 to more than 4 years for women over age 65 (see appendix 7.1 and van Oortmarssen et al., 1990b). The sensitivity of mammography is high (95%) for tumours with a diameter of more than 1 cm and is estimated to be 70% for smaller tumours. Sensitivity for ductal carcinoma in situ is not known. We assumed, rather arbitrarily, a value of 70% in this chapter. Survival is stage- and age-dependent (Adami et al., 1986; de Waard et al., 1986). The model assumptions for the improvement of prognosis after early detection are based in this chapter on the

initial 29% overall mortality reduction recorded in the Kopparberg/ Östergötland trial for women aged 40-74 (Tabár et al., 1987b).

It was assumed that the screening programme will have started in 1988 and that it will stop in 2015. For a complete evaluation of costs and effects a finite period for running the programme has to be assumed. The year 2015 was chosen because 27 years is sufficiently long for evaluation purposes, and also in order to make a comparison with the Forrest study possible. The costs and effects of mass screening that occur after 2015 are computed until all women who may have participated in screening will have died. During these periods no important changes in the incidence of breast cancer and the effectiveness of its therapy are assumed to occur.

A discount rate of 5% is applied in calculation of the present value of social costs. This rate has been stipulated by the Dutch government. In the "Results" section, the effects are presented without discounting. In the computation of CE ratios, however, equal discount rates have been applied to both costs and effects (Keeler and Cretin, 1983; Ludbrook, 1987; Russell, 1987).

Results

Screening programme

Nationwide screening in the Netherlands will be organized at 4 levels: the screening unit, where X-ray images are taken and films developed; the central unit, where mammograms are read by radiologists; the regional joint management board, responsible for sending out invitations to women and for regional coordination; the national organization, which is responsible for training medical staff, national quality control and evaluation. According to Dutch recommendations, the screening test considered here consists only of a single oblique mammographic view. Only in women under age 56 are 2 images taken at the initial screen (Health Council, 1987). The films are developed at the screening unit, so that the radiographer can immediately judge its quality and take another image if needed for better quality.

There are 3 types of screening unit: the static unit which is located in a building, and has an expected capacity of 12,000 screens a year; the mobile unit (truck and trailer), with a capacity of 10,000; and the semi-mobile unit (Portacabin), with a capacity of 11,000. In calculating the required number of screening units, the admissible average travelling distance for the women was set at 8 kilometres. This means that more mobile and semi-mobile units will have to be used in rural areas.

The annual financial costs of mass screening for alternative II are summarized in Table 7.2(a). During the first years, when the whole network is being built up, the cost per screen is high, but after completion of the build-up period it decreases sharply to 35.5 US dollars per person screened. When the network is completed, more than 80% of the costs concern activities at the level

Table 7.2 Costs of mass screening for breast cancer. The Netherlands. Costs in US dollars x 10^6

	(a) Financial costs[1] per year for alternative II			
	(Start) 1988	1990	1995[2]	2010
Screening units and central units	2.15	6.95	16.85	22.70
Regional joint management boards	0.70	1.50	2.20	2.45
National organization	1.20	1.20	1.20	1.20
Total	4.05	9.65	20.25	26.35
Costs per screen	US$ 86.	US$ 53.50	US$ 36.50	US$ 35.50

	(b) Present value of social costs for the period 1988-2015. 5% discount rate. Four alternatives			
	Alternative			
	I	II	III	IV
Screening programme	158.0	274.5	367.5	456.5
Time and travel costs of women screened	15.0	29.5	41.0	52.0
Total	173.0	304.0	408.5	508.5

[1] excluding publicity which is financed by private organizations
[2] year in which network will be completed

of the screening units and the central units.

The Table also shows the present values of social costs of the total programme (1988-2015) for each alternative, social costs being the costs for society as a whole.

Incidence and mortality

The expected incidence of breast cancer cases is given in Figure 7.1. Incidence is defined here as "the number of newly diagnosed cases of breast cancer requiring treatment". In the absence of mass screening, this incidence increases steadily due to the aging of the Dutch population. The incidence is somewhat higher in the case of mass screening. This is due to the fact that in the absence of screening some women will die from other causes while the breast cancer is still in a pre-clinical but already screen-detectable stage.

About 30% of all breast cancers will be detected by screening. The remaining 70% occur outside the 50-70 age range, or they concern non-partici-

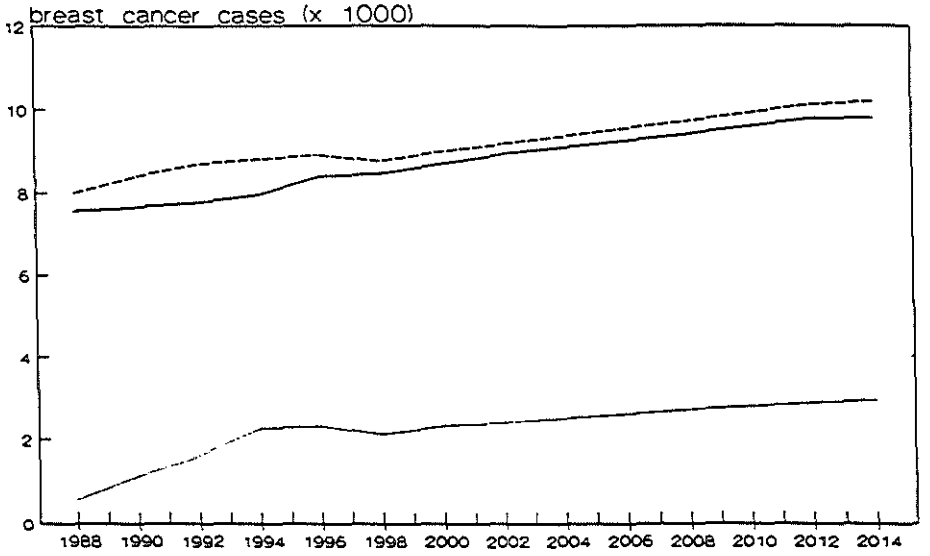

Figure 7.1 Expected yearly number of newly diagnosed cases of breast cancer, the Netherlands, 1988-2015, both with and without mass screening. Screening policy: ten 2-yearly invitations between ages 50 and 70. _____, No mass screening; - - - - -, mass screening, total incidence; bottom line, mass screening, screen-detected cancer.

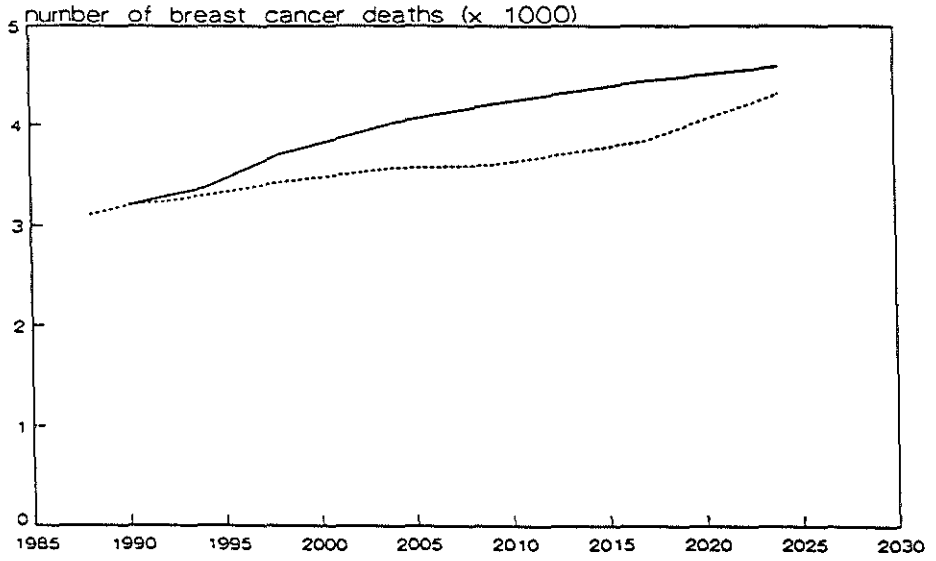

Figure 7.2 Expected yearly breast cancer mortality, the Netherlands, 1988-2025, both with and without screening. Screening policy: ten 2-yearly invitations between ages 50 and 70. The screening programme is assumed to be terminated in 2015. _____, No mass screening; - - - - -, screening. Improvement in prognosis based on Tabár et al., 1987b (40-74) only.

pants, fast-growing tumours and false-negative test results. The peaking of newly diagnosed cases during the first-"prevalence''screen is diluted over time because of the 7-year build-up period. It takes considerable time after the start ofscreening for a significant reduction in mortality to be reached, but it is important to notice that this reduction will disappear only gradually after termination of the programme (Figure 7.2). When the effect is at its maximum, mass screening will prevent at least 500 breast cancer deaths per year. This constitutes about 12% of the Dutch breast cancer mortality. Without screening, mortality will increase from about 3,000 deaths in 1985 to about 4,500 in 2020, due to aging of the population.

Assessment

At least 4 topics concerning the effects of breast cancer screening on the demand for assessment and therapy need to be considered: an increase in the number of impalpable lesions that require further assessment; possible changes in the use of different treatment modalities; a reduction of the need for treatment of advanced disease; the costs of these changes.

When a woman is found to have a mammographic abnormality, further assessment is required. Assessment procedures for palpable lesions differ considerably from those for impalpable lesions. Therefore, two sequential diagnostic schedules have been developed, one for each type of lesion. The sequential nature of the assessment should protect women from unnecessary interventional procedures and unnecessary treatment. The Dutch trials have shown that the predictive value of a positive screening result can reach 40% in the first round and 60% in subsequent rounds (these figures are more unfavourable for younger women). About half of the false-positive referrals will have had an excision biopsy. In the other women malignancy is excluded earlier, due to magnification-view mammography or fine-needle aspiration cytology. In the Netherlands the costs for assessment of impalpable lesions are high.

Excision biopsy and the (eventual) primary treatment are not done under the same anaesthetic, as is usually the case with palpable malignant lesions. Impalpable lesions, which prove to be malignant, require either an additional admission or a longer stay in hospital before treatment is given, due to the more laborious histological examination. The upper part of Table 7.3 shows the financial costs of assessment for abnormalities in which an excision biopsy is performed. The changes in the demand for these procedures resulting from the introduction of mammographic screening are also presented. The 13% reduction in the costs of assessment of palpable breast abnormalities and the 3% increase in the costs of impalpable abnormalities reflect the shift towards smaller tumour size due to screening and the changes in the number of false positives. Furthermore, these changes take into account the decrease in the number of women aged 50-70 who will consult their general physician with suspected breast lesions. Most of these lesions will have been detected at screening.

Treatment
The main types of breast cancer treatment, the number of patients expected to undergo these treatments, and the expected changes in these numbers resulting from screening are given in the lower part of Table 7.3. For instance, breast-conserving therapy will be applied in approximately 40% of women with clinically diagnosed breast cancer, compared to 70% for screen-detected cancers. There will be a concomitant decrease in the demand for other modes of primary treatment. Parallel to the achieved reduction in breast cancer mortality there will be a fall in the number of women treated for advanced disease.

Total costs of mass screening
The right-hand column of Table 7.4 shows the total costs associated with breast cancer diagnosis and treatment for the whole period in the Netherlands, in the absence of screening. Costs are in 1988 values, using a 5% discount rate. Assessment, primary treatment and treatment of advanced disease all take a nearly equal share. The 4 central columns present the differences in cost which arise from

Table 7.3 Financial costs of various types of assessment and treatment for breast cancer, and changes (%) in use of these types after completion of the screening network (year: 2000); alternative II

	Costs per unit		Numbers per year	
			Without screening (absolute)	Change by screening (%)
Assessment of palpable lesions:			20,000	-13
breast cancer	US $	275		
benign lesions	US $	1,525		
Assessment impalpable breast abnormality	US $	1,450	6,900	+ 3
Breast-conserving therapy[1]	US $	5,500	3,300	+ 26
Mastectomy	US $	4,500	2,400	- 10
Mastectomy and radiotherapy	US $	6,000	2,100	- 18
Adjuvant systemic therapy	US $	875	3,800	- 8
Treatment of advanced disease	US $	14,500 [2]	3,200	- 8

[1] in this table breast-conserving therapy includes radiotherapy and an external booster
[2] new calculations presented in chapter 3

Table 7.4 Social costs of breast cancer assessment and treatment if no screening is carried out, and the difference in cost for 4 alternative screening policies, 1988-2015, when compared to the situation without screening in the Netherlands. Costs in millions US dollars, 1988 value, 5% discount rate

	Screening alternative				No screening
	I	II	III	IV	
	Number of invitations per woman				
	5	10	15	20	0
Screening programme	+ 173.5	+ 304.0	+ 408.5	+ 508.5	0
Assessment after screening	+ 34.0	+ 49.0	+ 55.0	+ 61.0	0
Assessment, not referred by screening	- 74.0	- 109.5	- 121.5	- 133.5	753.0
Primary treatment	+ 19.5	+ 28.5	+ 31.5	+ 35.5	864.5
Treatment of advanced disease	- 39.5	- 60.0	- 70.0	- 78.0	850.5
Total costs	+ 113.5	+ 212.0	+ 304.0	+ 393.5	2,468.0

Table 7.5 Mortality effects and cost-effectiveness indices for breast cancer screening in the Netherlands, 1988-2015. Discount rate 5%

	Alternative				
	I	II	III	IV	Forrest[1]
Deaths prevented	2,700	4,100	4,800	5,400	2,500
Life-years gained	28,400	43,800	51,500	57,500	27,900
Costs (x 10^6)	113.5	212.0	304.0	393.5	122.0
Costs per death prevented	42,050	50,300	58,100	73,140	48,500
Costs per life-year gained (US $)	4,050	4,850	5,900	6,840	4,400
Marginal costs per life-year gained (US $)	4,050	6,050	12,900	14,800	4,400

[1] 3-yearly screening, 50-65

having a screening programme running during this period. The costs of screening are approximately proportional to the total number of screens in each alternative.

The costs of assessment after screening have been obtained by applying the "national assessment schedules" to the number of true and false-positives (palpable and non-palpable) detected at screening. The costs of assessment for women who are not referred by screening decrease, because considerably less women aged 50-70 will consult their general practitioners for breast lesions. Moreover, in the latter situation the diagnostic process is likely to be less efficient than in a screening setting. Costs of primary treatment increase by a few percent, because a somewhat greater number of women are treated earlier (Figure 7.1), and also more often by the slightly more expensive breast-conserving therapy. The saving of 40 to 80 million dollars from treatment of advanced disease reflects the reduction in breast cancer mortality. Altogether, about one-third of the costs of the screening programme are expected to be compensated for by savings in assessment and treatment.

Cost-effectiveness ratios

Costs per death prevented and costs per life-year gained are the 2 cost-effectiveness indices considered. As can be seen from Table 7.5, the costs per death prevented are about 50,000 US dollars and the costs per life-year gained about 5,000 US dollars for alternative II. More frequent screening leads to far less favourable values for the CE indices. The bottom row of Table 7.5 gives the marginal CE with increasing screening intensity. Thus, the extra costs per extra life-year gained when comparing 10 invitations (II) with 5 invitations (I) are 6,050 US dollars. The marginal costs per life-year gained are 12,900 US dollars when having 15 invitations instead of 10, and 14,800 US dollars when having 20 instead of 15. The table also shows the CE of applying to the Dutch situation the invitation schedule that was proposed by the Forrest Committee for the United Kingdom: 3-yearly screening between ages 50-65 (UK Working Party on Breast Cancer Screening, 1987). The schedule is suboptimal when compared to alternative I. This result is comparable to the analysis by Knox (1988), who concluded that optimization of the investment proposed by Forrest, in terms of deaths saved, results in a somewhat wider age range.

Discussion

Mortality reduction

With ten 2-yearly invitations for screening between ages 50 and 70, breast cancer mortality will be reduced by 12% in the total population according to the known trial results up to 1988. Why is this reduction not more pronounced? In the first place, more than 40% of all breast cancers will surface clinically before age 50 or will only become screen-detectable after age 70. From the remaining almost

60%, at least half will be missed because of non-attendance, false-negatives, or fast-growing tumours. Thus, approximately 26% of all breast cancers will be screen-detected. These screen-detected cases however, represent less than 26% of the breast cancer mortality, due to the relatively higher breast cancer lethality in the older age groups where no screening takes place. About 5 out of 10 breast cancer deaths that would have occurred in this group will not be prevented, in spite of their earlier detection.

This leaves a 12% reduction of breast cancer mortality due to screening. Twelve percent may not sound impressive, but it represents 500 deaths per year, more than the total mortality from cervical cancer in the Netherlands.

Screening for women under 50

For the 40-49 age group, the results of breast cancer screening trials have been rather disappointing in terms of mortality reduction, sensitivity of the mammography, and predictive value of an abnormal mammogram (Shapiro et al., 1982; Verbeek et al., 1984; Tabár et al., 1985; Andersson et al., 1988). Only the HIP results showed a mortality reduction in this age group (Chu et al., 1988). Moreover, the incidence is lower than in the over-50 age-group.

Instead of making assumptions on the effectiveness of screening women under 50, we took an indirect and rough approach by studying the question of how effective screening under 50 would have to be compared to screening over 50, in order to make it equally cost-effective. It appeared that in order to be equally cost-effective, the percentage mortality reduction under 50 should amount to 70%-90% of the mortality reduction over 50. For example, to obtain a cost-effectiveness ratio of 4,850 US dollars per life-year gained with a policy involving 10 invitations between ages 46 and 70, the mortality reduction under age 50 should at least be 75%. If the mortality reduction due to the screening of women under 50 falls below this figure, it is more cost-effective to invite women over 50 more frequently than to invite women under 50. This lower bound of 70% to 90% for the relative effectiveness is undoubtedly higher than current estimates. This strengthens the conclusion that screening women under 50 is less cost-effective. Of course, there may be other reasons for nevertheless screening this age group, but it shows the relative ineffectiveness of screening this age group.

Comparison with other cost-effectiveness studies of breast cancer screening

Other estimates about CE of breast cancer screening have been published (Gravelle et al., 1982; Forrest, 1986; Knox, 1988). The CE ratio in the Forrest report is £3,044 (at the 1989 exchange rate: 5,323 US dollars). With the same schedule of 5 screens between 50 and 65 we found a ratio of 4,400 US dollars (Table 7.5). When taking health-cost parity into account, the ratio should be lower for the United Kingdom instead of higher. So in fact the Forrest ratio is even more unfavourable than the ratio in our study.

In our opinion, the reasons for this rather large difference are 2-fold: in the Forrest report, the mortality effects are underestimated, and the costs are overestimated. The underestimation of the effects occurs as the period of analysis is the same for costs and effects in the Forrest study, thus leaving out the large part of effects that will emerge after the screening programme is stopped. This underestimation is only partly compensated for by the use of a cohort instead of the real population for the effect estimates, and by not taking a build-up period into account. The costs are overestimated, because the reduction in treatment of advanced disease which is directly linked to the mortality reduction is not taken into account in the Forrest report. These are also omitted in the studies by Gravelle (1982) and Knox (1988). In our study, the total costs of the programme are reduced by at least 20% as a result of this reduction in treatment costs.

The model that was used by Knox is a steady-state cohort model, without a build-up period for screening. This is clearly different from our approach, which includes realistic demographic development (aging population) and a build-up period.

Comparison with other cost-effectiveness ratios

Up to 1989, only a few Dutch studies were available which had calculated CE ratios for health care services of other types. It appears that breast cancer screening at 2-yearly intervals between ages 50-70 is very cost-effective when compared to, e.g., screening for cervical cancer, the treatment of endstage renal disease, and liver-transplantation (Habbema et al., 1988; de Charro, 1988; Habbema and Bonsel, 1988). In particular, the costs per life-year gained for the best schedule of cervical cancer screening with 10 invitations are about twice as high as those for breast cancer screening in our alternative II. International comparisons of CE ratios are more difficult to make, but nevertheless it can be concluded that breast cancer screening compares favourably with many other health care services (Williams, 1985; Torrance, 1986). Of course, such comparisons should ideally take the completeness, the assumptions and the quality of the studies concerned into account.

Uncertainties to be resolved

Each of the factors in the CE calculations carries some degree of uncertainty. The influence of these uncertainties on the final result can be assessed by performing sensitivity analyses. The following factors seem among the most influential ones. First, the current amount of diagnostic and treatment procedures resulting from visits to general practitioners because of possible lesions will be modified after the introduction of mass screening. This modification also includes a possible increase in the demand for mammograms in the younger age group, which is not invited to the screening programme. Second, the implications of more frequent detection of ductal carcinoma in situ (DCIS) on the costs and benefits of screening are not yet clear. Third, the mortality reduction estimates have a certain range of uncertainty.

These factors may have considerable impact on the CE ratios, and are discussed in the next chapter. Still, it seemed justified to build up a nationwide screening programme for breast cancer in the Netherlands for women aged 50-70, with a 2-year interval between each screening. Quality control, training, registration and evaluation should be arranged and controlled at a national level, directly from the beginning. There should be room to carry out well-planned trials, for example for the under-50 age group.

8 Cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors

Introduction

Several European countries have started to implement nationwide breast cancer screening. On the basis of results from randomized trials, a reduction in mortality from breast cancer can be expected for women aged 50 and over (Miller et al., 1990). Breast cancer screening seems cost-effective in countries with relatively high breast cancer incidence and mortality rates, if measures are taken to guarantee a high mammographic quality and participation rate.

The impact of such national programmes on quality of life has been the subject of much discussion. The potential negative effects of the screening examination itself (Ellman et al., 1989), the referral of a significant number of women with benign lesions (Gram et al., 1990) and the consequences of earlier and often more intensive treatment can not be ignored. And although the evidence for mortality reduction is substantial, screening trials report wide ranges (Andersson et al., 1988; Roberts et al., 1990). At first sight, the discussion whether to expand screening programmes to younger (Rutqvist et al., 1990) or older women (Hobbs et al., 1990) seems a contradiction.

A large number of factors may influence both favourable and unfavourable effects of screening and possibly its cost-effectiveness. Important variables are the improvement in prognosis for screen-detected cases, the predictive value of the screening test and the chance that screen-detected ductal carcinoma in situ would have progressed to invasive carcinoma. Due to the long time interval between screening and the realization of favourable effects, developments irrespective of the screening programme itself -such as a possible improvement in survival of breast cancer patients- could affect the impact of today's screening.

Chapter based on de Koning et al., Int J Cancer, 1991

This chapter presents results on quality of life and cost-effectiveness for 5 screening variants, each differing in age group or screening interval. It also gives a systematic overview of the influence of various factors, concerning screening or clinical practice, that are uncertain or may become important in the future. Special attention has been given to the uncertainties on the attainable mortality reduction in a real population and to the extension of the programme to younger or older women. We calculate the impact of breast cancer screening for the Dutch population and compare this to the results achieved for the first 7,500 women screened in the national programme in the Netherlands in 1989.

Material and methods

Cost-effectiveness: analysis and principal screening variant
Costs and effects of different breast cancer screening policies are compared to the situation where mass screening is not applied. The description of the latter situation is based on current national data on assessment and treatment. Cost and effect estimates for screening are based on results from the screening trials in the Netherlands, but the assumption of the improvement in prognosis after early detection is based on the randomized screening trials. For each policy, costs and effects are related by computing CE ratios, expressing cost per life-year gained (5% discount rate both), using the MISCAN model for breast cancer screening (van Oortmarssen et al., 1990b; de Koning et al., 1990c) and the Dutch population structure. Screening programmes are assumed to run from 1990 until 2017. The corresponding effects and costs are computed until all women who may have participated in screening will have died.

The programme of 2-yearly mammographic screening for women aged 50-70 is used as reference policy for all other variants. The attendance rate is 70% on average, decreasing from 75% at age 50 to 65% at age 70, based on the data from the Dutch trials. In general, we have used actual data as far as possible. Different sources were used to obtain the most recent and accurate estimates (de Koning et al., 1990b). We mention only the most important sources. We analyzed age-specific data from 76,020 general practitioners' consultations (National Study of Diseases and Procedures in General Practice) and from mammography registries (Continuous Morbidity Registration Sentinel Stations in the Netherlands and in Amsterdam) to estimate the number of women who ask for (or need) assessment outside or without a breast cancer screening programme. The change in numbers and types of assessment when introducing screening was estimated using the referral patterns from the Dutch trials and by analyzing all procedures carried out for referred women in Nijmegen. A survey amongst all Dutch radiotherapy departments (1986-1988), an analysis of all hospital admissions for breast cancer (1983-1988) and a questionnaire survey amongst 40 experts provided data on primary treatment.

Actual costs have been estimated for screening, radiation treatment and

pathological procedures concerning impalpable tumours. For all other medical procedures, tariffs have been used as approximations for actual costs. We analyzed 68 patient files of women who had died from breast cancer in 1985-1989 in order to estimate the cost and types of treatment for women with advanced disease (de Koning et al., 1992a), including an analysis of extramural cost (Koopmanschap et al., 1992). Costs are presented in US dollars (= 2 Dfl).

Effects on breast cancer mortality

In all variants, the base-line assumption of improvement in prognosis after early (mammographic) detection is based on the reported results from the randomized trials in Kopparberg/Östergötland (Tabár et al., 1989) and in Malmö (Andersson et al., 1988). Simulating the 2 Swedish programmes separately, taking into account the specific attendance rate, screening interval, age distribution at entry and follow-up period for each and assuming the same improvement in prognosis per shift in stage, the expected effects on mortality reduction for both trials were the same as observed. These factors could, therefore, not explain the different published point estimates on mortality reduction, in other words we could not find reason to assume a difference in efficacy between both trials. Assuming equal mammographic quality in both projects, we combined the results and calculated a weighted (on size of confidence intervals) average: a 32% reduction in breast cancer mortality in the study group aged 50-70 after 8 to 9 years for trial situations like these two. The boundaries of the 95% confidence intervals of the combined projects (14% and 46% reduction) were used in sensitivity analyses. Point estimates of mortality reduction from other European trials (Collette et al., 1984; Verbeek et al., 1984; Palli et al., 1989; Roberts et al., 1990; Frisell et al., 1991) are within this range, or even more favourable.

Impact on quality of life

In investigating the impact of screening on women's quality of life, we first defined the possible phases she may pass through: screening examination, assessment and biopsy, primary surgery, primary radiotherapy, adjuvant systemic therapy, first year after mastectomy, first year after breast-conserving therapy, disease-free interval (DFI) after mastectomy, DFI after breast-conserving therapy, and advanced disease (divided into treatment episodes and a terminal phase). The literature on quality of life and breast cancer (screening) was used to construct a questionnaire summarizing the concomitants of disease for each phase. These ques-tionnaires were sent to 31 breast cancer clinicians or public health experts, who were asked to evaluate the described phases on a visual analogue scale ranging from 0 to 100 (de Haes et al., 1991). Utilities were calculated by transforming the median values of 27 respondents (Bombardier et al., 1982) and using assumptions on the durations of each phase. These utilities were combined with the MISCAN predictions on assessment and treatment procedures, number of screened women and life-years in the situation with or without screening to estimate the overall gain in QALYs.

Screening variants and sensitivity analyses of variables
We simulated 3 alternative screening variants for women in the age group 50-70 with screening intervals of 1, 1.3 and 4 years. Other simulated screening policies are: the United Kingdom policy, 3-yearly screening of women aged 50-65; 2-yearly screening of women aged 40-70; and 2-yearly screening of women aged 50-75 (Table 8.1 shows the 5 variants). Specific differences in the younger or older age groups, compared to the 50-70 group, have been taken into account. The non-significant 8% breast cancer mortality reduction published for screening women under age 50 in Kopparberg/Östergötland (Tabár et al., 1989) has been used as an assumption for this age group in the 40-70 policy. The data on the effect of screening at ages 70 and more is scarce. The improvement in prognosis in the over-70 age group is assumed to be equal to that of the 50-70 group. Other differences were based on actual data: the attendance rate has been set at 75% in the 40-49 age-group and at 45% in the 70-75 group. Further adjustments were made for the positive predictive value of screening (30% in young age-group instead of 41% at first, and 43% instead of 57% at subsequent screens) and for the amount of assessment outside a screening programme in these groups.

The main variables used in the cost-effectiveness (CE) analysis, applicable to the screening situation or clinical practice or both, were varied to study the possible impact on cost-effectiveness and quality of life. Estimates for these alternative assumptions were based on our analyses of data, literature or discussions with experts. We also analyzed the influence of including travel, time and other direct (non-medical) costs incurred by the women involved. The impact on the results of including indirect costs, such as medical expenses for non-related diseases in the life-years gained, was also analyzed.

Results

Effects on mortality and costs of the principal variant
The 2-yearly screening programme for women aged 50-70 is predicted to detect 26% of all diagnosed breast cancers in the population. The average size of the screen-detected tumours is small: after the build-up period of screening with a relatively large number of prevalent (large) cancers, 80% of all screen-detected cancers are smaller than 20 mm or are non-invasive cancers. This would be 37% in the non-screening situation. The proportion of women with axillary lymph node metastases is approximately 60% of the proportion in clinically diagnosed cases, given the same tumour size. The start of the screening programme will initially result in a sharp rise in the number of newly diagnosed cases, with a maximum 17% (= 1,450 cancers) increase in the year 1993. From 1996 onwards, the total number of diagnosed breast cancers will be 3.5% higher each year, compared to the expected situation without screening.

Earlier diagnosis will gradually make its impact on mortality reduction. After 10 years, the number of women who will die of breast cancer may have

fallen from 3,675 to 3,325. A maximum breast cancer mortality reduction of 16% (= 700 women) annually in the total population is attainable from the year 2015 onwards. On average, each year of screening will prevent 630 breast cancer deaths. This leads to a total of 17,000 with a 27-year programme (Table 8.1). The total costs for screening are 300 million US dollars and the additional costs of treating and following up more women earlier are 72 million (5% discount rate). In the opposite direction, a large decrease in the cost for women with advanced breast cancer is expected (-128 mln). The net additional cost of 233 million US dollars divided by 61,000 life-years gained results in a CE ratio of 3,825 US dollars per life-year gained (5% discount rate).

Quality of life and cost-utility

Although mortality reduction is the fundamental effect, there is much debate about other desirable and undesirable consequences of screening that may influence women's quality of life. We therefore combined the changes in life-expectancy with the expected changes in morbidity. Figure 8.1 summarizes the most important favourable and unfavourable effects of the principal screening programme (per million screens), other than mortality reduction or gain in crude number of life-years. The scale represents the relative weights to be given to different types of morbidity in the phases a woman may be in; 100 means a perfect quality of life and 0 represents the worst possible state. The value 82 for adjuvant hormonal treatment implies an estimated 18% loss in utility during this phase as compared to the situation of perfect health. Screening 1 million women will make adjuvant hormonal treatment unnecessary in 525 women, due to the smaller number of women with lymph-node metastases. Therefore, this effect leads to an increase of (525 x 0.18 x 2 years =) 189 QALYs.

The screening examination itself is estimated to have only a slight negative and short-lasting impact, resulting in a decrease of (1 million x 0.006 x 1/52 year =) 115 QALYs. Despite the large number of women screened (15.8 million) during the period 1990-2017, only 7% of the total negative quality-adjustment is incurred by these examinations. More important, per every million screens, approximately 4,500 women will be diagnosed as having breast cancer on average 4 years earlier (lead time) and more than 1,000 patients will experience a longer disease-free interval (16,500 life-years gained). In Figure 8.1, both types of additional years in follow-up appear on the negative side of the balance, as we are trying to correct the total gain in life-years for quality of life. A small loss in utility during years in follow-up is justified since women will have to undergo medical follow-up proce-dures, and breast surgery and the knowledge that she has had breast cancer will have a negative impact on her quality of life. The large increase in woman-years in follow-up is almost entirely responsible for the negative quality of life adjustment, whereas the decrease in the number of advanced breast cancer patients as a result of screening is responsible for 70% of the positive quality of life adjustment.

Table 8.1 Effects on mortality, costs, cost-effectiveness and cost-utility for different breast cancer screening policies (1990-2017) in the Netherlands. 5% discount rate and costs in millions US dollars (unless stated). Cost amounts are expected differences between situation with and without screening

Age group Screening interval	50-70 2 yr	40-70 2 yr	50-70 1.3 yr	50-75 2 yr	50-65 3 yr
Breast cancer deaths prevented[1]	17,000	17,800	19,800	19,450	10,800
Life-years gained[1]	260,000	290,000	310,000	275,000	180,000
Cost of screening	300	457	405	310	185
Cost of assessment/biopsy	- 10	- 62	- 12	2	- 12
Cost of primary treatment	50	57[2]	55	71	26
Cost of follow-up	22	25	25	27	14
Cost of advanced disease	- 128	- 131[2]	- 145	- 145	- 80
Difference in costs	233	346	328	265	133
Breast cancer deaths prevented	6,000	6,115	6,780	6,790	3,770
Life-years gained	61,000	64,000	70,000	64,500	41,000
Quality-adjusted life-years gained (QALYs)	57,500	59,500[2]	66,000	59,500	39,300
Cost (US $) per life-year gained (CE-ratio)	3,825	5,385	4,670	4,100	3,235
Cost (US $) per QALY gained	4,050	5,815[2]	5,000	4,450	3,400

[1] not discounted
[2] no age-specific data for treating women < 50

In the 27-year programme, a total of 252,000 QALYs are gained, which is a very small decrease compared to non-adjusted life-years gained (260,000). Since the more favourable effects are preceded by unfavourable effects, this difference gets larger if effects and costs are discounted. The cost-utility ratio is 5.6% higher than the cost-effectiveness ratio: 4,050 US dollars per QALY gained (Table 8.1).

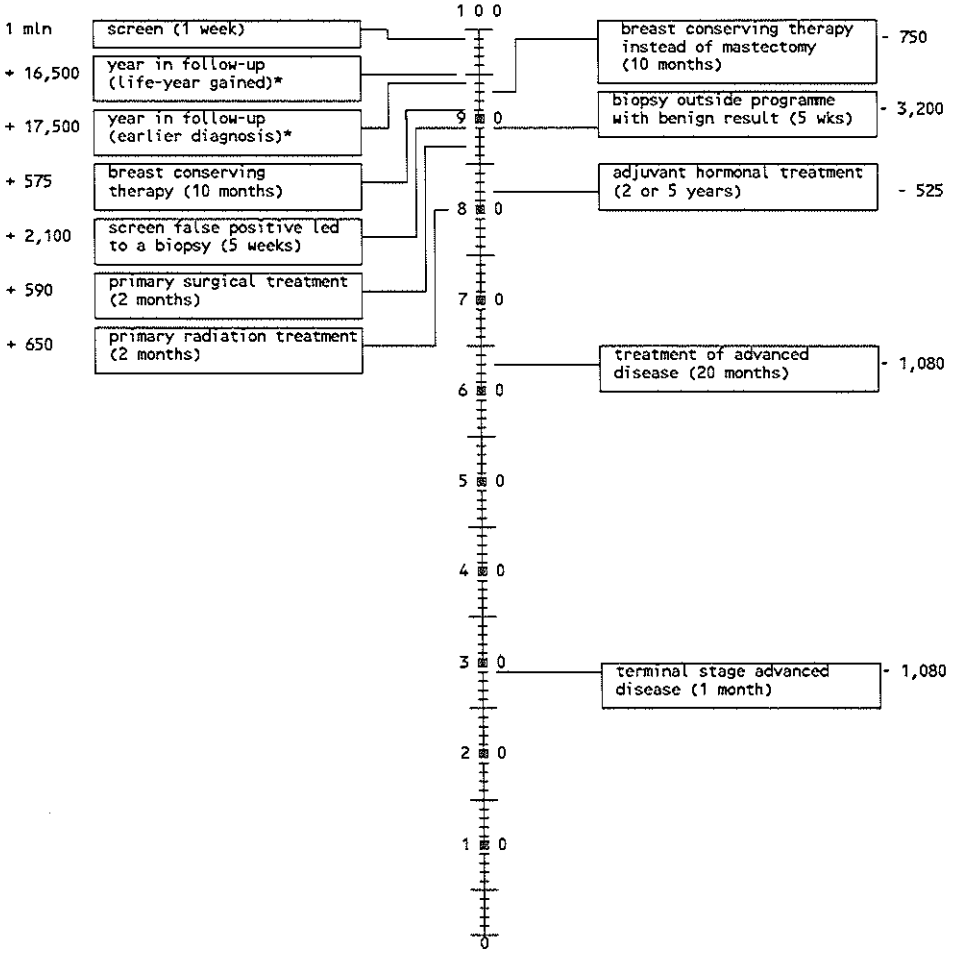

Figure 8.1 Favourable and unfavourable effects of 2-yearly mammographic screening in age group 50-70, other than mortality reduction. Changes in numbers are given per one million screens (not discounted). Effects on morbidity have been divided into short-term (treatment phase; 2 months), intermediate (first year after treatment; 10 months) and long term effects (life-years). Scale represents relative weights to be given to different types of morbidity. Decrease in quality of life due to unfavourable changes on the left and increase due to favourable changes on the right. * Reduced quality of life (see text).

Alternative screening policies

Two important questions regarding screening policies are frequently discussed. The first is whether or not to extend the programmes to younger age groups. In Table 8.1, the second column shows a 50% increase in the cost of 2-yearly screening for women aged 40-70, compared to that for women aged 50-70. Using the Kopparberg/Östergötland-trial estimate for the 40-50 group, the total number of prevented breast cancer deaths increases by 800, which makes only a 5% increase in the discounted number of life-years gained. The difference in costs is strongly influenced by our assumption that assessment procedures outside the programme will diminish in proportion to the decrease in clinically diagnosed breast cancer cases as a result of screening. Applying this assumption to the younger screening group also, the relatively high number of medical procedures carried out at present in the under-fifties results in 6-fold higher savings in cost of assessment than in the principal policy despite a lower positive predictive value. The CE ratio is 40% higher than with 2-yearly screening starting at age 50.

The additional cost per additional life-year gained, marginal CE ratio, comparing 2-yearly screening in age group 40-70 with 2-yearly screening in age group 50-70, is 35,000 US dollars. This may be compared with the option of more intensive screening within the age group 50-70 (15 invitations) with an expected marginal CE ratio of only 10,550 US dollars per additional life-year gained (Figure 8.2).

The second question is whether screening should continue in women over 70. In the age-group 70-75 both the attendance rate for screening and women's life-expectancy are decreasing, but the incidence and mortality rates for breast cancer and the sensitivity of mammography are relatively high. Two additional screens up to the age of 75 would increase both the number of prevented breast cancer deaths and costs by approximately 15%. The CE ratio is only 7% higher compared to the principal policy and the cost-utility ratio 10% (Table 8.1). The relatively lower number of life-years gained per death prevented in the 70+ group, if compared to the 50-70 group, results in a marginal CE of 8,000 US dollars per additional life-year gained (Figure 8.2). However, this is definitely more favourable than extension to younger age groups and compares well with intensifying the programme within the age group 50-70.

The final column in Table 8.1 shows that a longer screening interval, as in the United Kingdom policy, is also relatively cost-effective: it would result in a CE ratio of 3,235 US dollars per life-year gained in the Dutch situation. The additional costs are 57% and the life-years gained amount to 68% of those in the present Dutch programme.

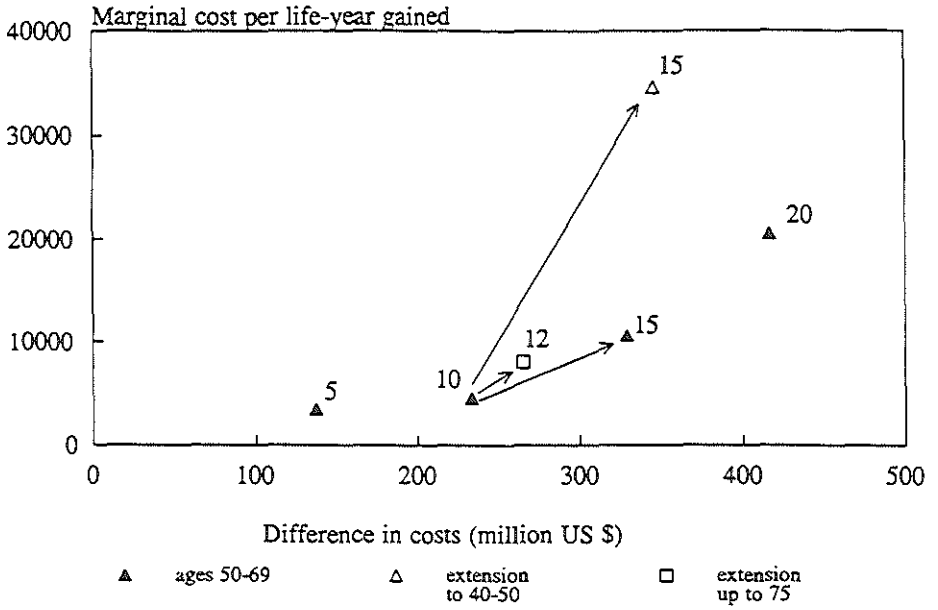

Figure 8.2 Marginal cost-effectiveness (additional US dollars per additional life-year gained) of 6 breast cancer screening policies: 5, 10, 15 or 20 invitations in the age group 50-70, 12 invitations in the age group 50-75, and 5 invitations during ages 40-49 followed by 10 invitations during ages 50-70. The corresponding differences in cost for each screening policy have been put at the horizontal axis. 5% discount rate.

The improvement in prognosis for screen-detected cases

The actual attainable reduction in breast cancer mortality at the population level is a matter of discussion for several reasons. The estimates of cost per life-year gained are especially sensitive to variation in the benefit estimates. We analyzed this influence on CE for different assumptions on the improvement of prognosis after early detection for women aged 50 and older. The triangles in Figure 8.3 correspond to the upper- and lower limits of the (weighted) combined 95%-confidence interval of the 2 Swedish trials. Applying the more unfavourable estimate (based on a 14% reduction) to the Dutch situation would substantially influence the number of breast cancer deaths prevented. The breast cancer mortality would be lowered by only 7% (instead of 16%) in 2015. The discounted number of life-years gained shows a 58% decrease from 61,000 to 25,900, whereas the costs rise by 27% to 294 million US dollars. The cost of additional follow-up procedures would be lower, but the savings in treating less women for advanced disease would also be reduced to 55 million US dollars. Therefore, the CE ratio increases 3-fold to 11,300 US dollars per life-year gained. The favourable 46% assumption would lead to a CE ratio of 2,250 US dollars per life-year gained. Figure 8.3 also shows the use of other point

estimates on mortality reduction in this age group, considering 70% attendance, of the other European trials. The "Swedish average" appears a realistic one for the CE analysis.

The difference in mortality is a reflection of both the situation without screening and the improvement due to mammographic screening. The published trend in earlier diagnosis in different countries (Bennett et al., 1990) or in improved survival for breast cancer patients in recent years (Adami et al., 1986) may in part be a result of the higher awareness of women and physicians of breast problems, resulting in an earlier diagnosis irrespective of nationwide screening. In one variant we assumed that this general shortening of the delay between onset of disease and diagnosis would lead to a rather high (and immediate from 1990 on) 25% decrease in the chance of a breast cancer patient, not diagnosed via screening, dying of her disease. If so, the CE ratio of 2-yearly screening would increase to 6,500 US dollars per life-year gained. However, we did not take into account possible additional costs for mammograms, etc. to obtain such an earlier diagnosis and possible survival improvement for a group of women. This, of course, would lower the CE ratio again.

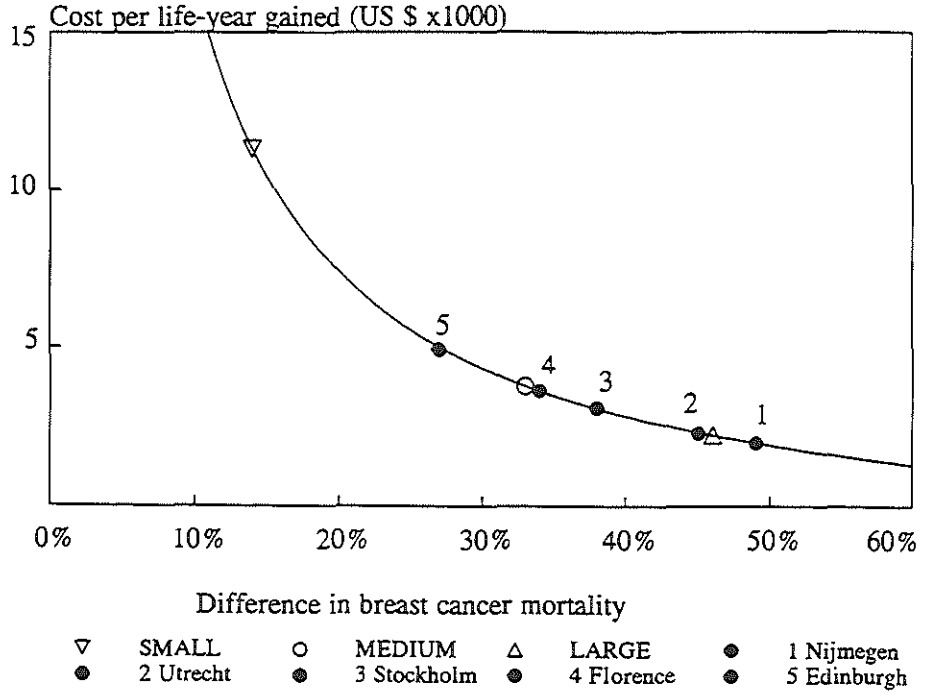

Figure 8.3 Cost (in 1,000 US dollars) per life-year gained for 2-yearly screening in age group 50-70 (5% discount rate), with changing assumptions on difference in breast cancer mortality in this age group. The "Medium" assumption corresponds to the weighted average from Kopparberg/Östergötland and Malmö. The boundaries of 95% confidence intervals correspond to "Small" and "Large" (separate trial results adjusted for 70% attendance).

Influence of other variables

The accuracy of our CE ratios may be influenced by a large number of factors, of which the possible improvement in prognosis is certainly the most crucial. Table 8.2 summarizes 8 other factors, which appear to have been responsible for a more than 5% difference in CE ratio, if varied within plausible ranges. The only variable for which our estimate may well have resulted in too high a predicted CE ratio for the principal policy is the cost for women with advanced breast cancer. We have used a moderate average cost estimate per treatment.The positive predictive value of the mammographic screening test and the possible change in attitude of women in the screening interval appear to be important factors to be monitored already during the first years of implementation. If we assume that mass screening would only reduce the present 33% of assessment procedures that is known or registered to be "for preventive reasons", the CE ratio increases by 18% to 4,465 US dollars per life-year gained. If implementing a programme for women aged 50 and over led to a significant increase in the demand for mammograms in women under 50, CE could also deteriorate strongly.

The most influential cost components are the cost of follow-up procedures, as there is a great variety in the protocols for treated breast cancer patients, and the cost of screening. The latter may, for various reasons, become higher than expected when the national programme is under way. Uncertainties would forecast a 10% increase in costs and CE in the Netherlands. Including travel, time and other non-medical expenses for the women involved increases the CE for 2-yearly screening in the age group 50-70 by 17%. In particular, the inclusion of indirect costs (production losses and additional costs due to other diseases in the life-years gained) would double the CE ratio. However, these variants are only informative if cost-effectiveness ratios of other health care programmes are calculated in the same way.

Factors that might have been expected to be important, but that were found to have only limited impact on the cost-effectiveness of screening, are trends in primary treatment (e.g. significant increase in breast-conserving therapy irrespective of screening), average attendance rate for screening or the chance of ductal carcinoma in situ (DCIS) progressing to invasive carcinoma or clinical DCIS. For the latter chance, we worked with a 87.5% at long-term estimate based on the scarce literature, but varied this between 75% and 100%. Due to the relatively small proportion of these lesions, even with screening, a less progressive DCIS appears relatively unimportant in respect of costs and effects. The impact of future changes in adjuvant systemic treatment has not been analyzed. It was not yet clear whether the different trial results for adjuvant systemic treatment known up to 1991 would result in rapid changes in the criteria for this treatment and influence the treatment of screen-detected breast cancer patients.

Table 8.2 Alternative assumptions, other than mortality reduction, that influence the cost-effectiveness of 2-yearly mammographic screening of women aged 50-70 (CE in US dollars per life-year gained)

	Actual data or assumption[1] in principal variant	Alternative assumption	CE-ratio and % difference with principal variant
Cost of treatment of advanced breast cancer	US $ 21,000 per woman	25% higher costs	3,300 (- 14%)
Capacity screening units	12,000 screens per year	10,000 per year per unit	4,100 (+ 7%)
Positive predictive value mammographic screening test	51% on average over all rounds	43%	4,130 (+ 8%)
Follow-up examinations of treated women	Every 3 months in first 2 years	Twice as frequent	4,190 (+ 9%)
Total costs screening	US $ 40 per screen	US $ 43	4,225 (+ 10%)
Non-medical direct costs	Not included	Include travel, time and out-of-pocket costs to women	4,460 (+ 17%)
Demand for mammograms outside screening programme	Decrease in assessment proportional to decrease in clinical breast cancers[1]	Only decrease in assessment for preventive reasons	4,465 (+ 18%)
Indirect costs	Not included	Include medical costs for other diseases in life-years gained	7,250 (+ 90%)

[1] Assumption used in the analysis, as there is no data yet available on the magnitude of this change.

Discussion

Two-yearly mammographic screening for women aged 50-70 is expected to reduce breast cancer mortality by 16% in the Dutch population. At the same time, this attainable reduction is also the main uncertainty in calculating the balance between effects and costs for breast cancer screening. After the first significant evidence from the HIP-study and the Kopparberg/Östergötland-trial, less favourable results have been reported from other randomized trials. We have considered the HIP-results to be too outdated for predicting the effects of the present Dutch or European programmes. Our simulation has shown that the 2 Swedish studies in Kopparberg/Östergötland and Malmö are comparable. The combined analysis of all Swedish trials (see also chapter 9) has to indicate whether the assumption of 32% mortality reduction is indeed the best one at the moment, but the results of the Stockholm trial seem promising in this respect (Frisell et al., 1991). The reported United Kingdom results are strongly influenced by the low attendance rate and a less than optimal mammographic quality in the first years (Roberts et al., 1990). Correcting for 70% attendance alone already changes the attainable reduction to a great extent.

The extreme values on improvement in prognosis, based on the 95% confidence intervals in 2 Swedish trials, could alter the computed CE ratio 2- or 3-fold, but even the pessimistic assumption leads to a CE ratio of breast cancer screening that compares well with many other health care programmes. Variation of other assumptions generally results in CE ratios ranging from US dollars 3,000 to 5,000 per life-year, or 3,200 to 5,300 per QALY gained. To our knowledge, no superior cost-utility ratios have yet been reported for other screening programmes or cancer programmes. This is true even if an immediate 25% improvement in survival of breast cancer patients is assumed irrespective of screening.

Differences in QALYs is a more preferable measure than crude life-years gained alone. When taking into account associated morbidity for all possible phases, other favourable and unfavourable effects besides mortality reduction have only limited impact and appear to cancel each other out. More extreme assumptions on the expected utilities in the different phases result in an adjustment between -19% (most unfavourable) and +3% (most favourable variant) on life-years gained for 2-yearly screening women aged 50-70. Our evaluation of the effect on quality of life strongly supports the decision to introduce mammographic screening for women aged 50 and over.

Programmes with a screening interval of 2 or 3 years are both relatively cost-effective. If the budget is restricted, a 3-year interval may be appropriate. Quality of life appears to be of minor influence. Recently, CE ratios have been calculated for breast cancer screening in Australia based on a 2-yearly policy for women aged 50-79 (Australian Institute of Health, 1990). As the attendance rate is assumed to be 70% in the group 50-70 and only 15% for ages 70-79, the reported ratio of 9,340 Australian dollars (6,540 US dollars) per life-year gained

may be compared to our principal policy or our policy with 2 additional screens up to 75. Excluding travel and time cost results in a ratio of 4,670 US dollars per life-year gained, whereas our estimate for the Netherlands would be 4,800 to 4,900 (impact on treatment costs disregarded). However, the apparent similarity obscures 2 important differences. The breast cancer incidence and mortality rates are about 25% lower in Australia, and the cost for screening is estimated to be approximately US dollars 56, compared to 40 in the Netherlands. We would, therefore, have expected the CE ratio in Australia to be almost twice as high. However, the used model neglects the time-lag between screening and the realization of effects (life-years gained), which could easily mean a 2-fold overestimation of the gain in life-years in case of discounting.

At present, nationwide screening for women under 50 is not to be recommended. Controversy on the effectiveness of screening in this group is strong, but all recent publications and trials suggest that there is no short-term benefit in terms of mortality reduction. Our analysis shows that the assumption of a possible small reduction still leads to a very unfavourable balance in terms of marginal cost-effectiveness and cost-utility. A 1.5-year screening interval is often advocated in this age group in order to minimize the relatively high number of interval cancers (Tabár et al., 1987b). From a CE point of view, considering the present estimates on mortality reduction, it would further deteriorate the unfavourable balance between life-years gained and costs. Except for the HIP-trial, there are no other data yet on which to base a more favourable long-term

Table 8.3 Predicted results for the nationwide breast cancer screening programme compared to actual findings in first 7,500 women screened (50-70) in the Netherlands in 1989

	Predicted	In practice*
Attendance rate	70%	72% of invited populations
Screen positives	1.6%	1.6 % of screened women
Biopsy	75%	72% of referred women
Breast cancer detection rate	6.4	6.4 per 1,000 women screened
Positive predictive value screening test (referral)	41%	39% at first screen
Biopsies with malignant diagnosis	54%	55% of all biopsies
Lymph-node metastases	26%	28% of women with screen-detected breast cancer

* Source: IKO Comprehensive Cancer Centre and Dronkers (1990)

estimate. Besides, it appeared that this younger age group is, as in other countries, already undergoing mammographic examinations too often. Implementation of screening for women aged 50 and over may catalyse even further the demand for "screening." in younger age groups (Ashby et al., 1990), which would be regarded as a negative side-effect of screening women aged 50 and over. Finally, we have compared the actual findings of the first 7,500 women screened in the Arnhem region in 1989 (IKO Comprehensive Cancer Centre and Dronkers, 1990) to our predictions (Table 8.3). The similarity, e.g. in the attendance rate and the ratio between benign and malignant biopsies is quite striking. Comparing the predictions in our CE analysis with these (very) early outcome measures will have to serve as an early check on the effectiveness of the Dutch nationwide programme. In the Netherlands, 3 centres are now cooperating to function as such a National Evaluation Team for Breast cancer screening and a National Expert and Training Centre is responsible for quality control regarding mammography and pathology. Possible unfavourable side-effects of the programme, like regional differences in treatment or the mentioned increase in mammographies for younger women, have to be excluded as much as possible. The present analysis will to a large extent form the basis for national screening evaluations. It has also pointed at some aspects that are not directly related to the screening performances, but that may strongly influence the CE of the programme or women's quality of life. In conclusion, the implementation of breast cancer screening programmes for women aged 50 and older with a 2- or 3-year interval should be further stimulated, provided that a high quality of mammography and staff is guaranteed and that a high level of national evaluation is realized.

9 Discussion

Introduction

In this chapter, an update is given of the breast cancer mortality reductions achieved in screening trials, as new data from the Swedish randomized trials and from the United Kingdom has recently become available. This chapter also gives an overview of research priorities, identified on the basis of the remaining uncertainty about some issues. Finally, the research and the project as described in this thesis is put into the context of other similar research and into the political context.

Update screening trials

Combined analysis of Swedish randomized screening trials

Follow up of the existing breast cancer screening trials is still needed for some years, although it will become more and more difficult to interpret the results while nationwide screening programmes are being implemented around these areas or while the control group is being invited for screening. The very good results from the screening trials published in the mid-eighties were followed by seemingly less good results from some trials published in the late-eighties or early-nineties. As explained already in chapter 2, various factors like a lower attendance rate, worse mammographic quality in the early years of these trials (compared to the quality in the beginning of the earlier published trials) and more screening in the control group must have contributed to these lower reductions in breast cancer mortality. A somewhat longer period of follow up could probably show that the curves on cumulative breast cancer mortality in control group versus study group did need some more time to diverge, given these factors.

In fact, the data on mortality reduction that became available with longer follow up from trials since 1990, only strengthen this assumption. The randomized Stockholm trial showed a 37% breast cancer mortality reduction for women aged 50 and over after 7.4 years (Frisell et al., 1991), the good Kopparbèrg/Östergötland results remained roughly the same (Tabár et al., 1992b), whereas the Malmö results improved with time (Andersson, 1992).

A combined analysis of all (5) randomized Swedish breast cancer screening trials has recently been completed and presented (Nyström, 1993), in which causes of death were assessed of all deceased breast cancer patients

in the different trials (either participants of screening, non-participants or from the control group). This was done by a blinded review and by an independent end-point committee, which members were not involved in the screening trials. One of the main goals of the analysis was to estimate whether there had been any important misclassification which could have led to overoptimistic results. The analysis shows that the mortality reduction was similar in the trials as a whole, irrespective of the end-point used: "breast cancer as underlying cause of death" or "breast cancer present at death". There was no evidence of any detrimental effect of screening in terms of breast cancer mortality reduction in any age group (no significant increase in breast cancer mortality in the screened population). Again, the different point estimates between the different trials for all age groups varied non-significantly. The analysis showed a 29% reduction in mortality from breast cancer observed among women aged 50 to 69 at randomization for the 5 trials together (95% confidence intervals: 10% - 43%). Most of these trials have (had) longer screening intervals than 2 years, higher attendance rates than 70% and a mean period of follow up of 9 years.

Unfortunately, the new results are too new to have been able to perform new analyses on cost-effectiveness, embedding the specific characteristics of the different trials in the breast cancer model MISCAN and to assess whether the new results do or do not show any true variation in the efficacy of the screening between the individual centres. In chapters 2 and 8, we explained that the first published results from Kopparberg/Östergötland and Malmö did not provide evidence to assume such a difference.

Table 9.1 compares the earlier published results, on which our analyses have partly been based, with the data from the new combined analysis. Three specific differences have to be mentioned before comparing: in the new analysis actual age at randomization has been used, all women with breast cancer diagnosed before the start of the programmes have been excluded from the cohort, and an independent (and sometimes different) ascertainment of breast cancer death was applied. For some of the trials, follow up is longer in the new analysis. Except for the Östergötland trial, the new analysis definitively supports the earlier knowledge and evidence. The difference for this specific trial may come from the difference in the 70-74 age group, where the recent assessment by the independent committee of whether the woman did or did not die of breast cancer has resulted in a strikingly different point estimate for this age group (see Table 9.1). At the same time, the number of breast cancer deaths in the younger age group 40-49 increased relatively more in the study group with longer follow up than it did in the control group and has negatively influenced the point estimate including all ages.

The publication of the combined analysis from Sweden will certainly close the debate on mortality reduction for women aged 50 to 69 in the setting of screening trials. Especially, since these trials have been randomized trials with relatively high quality standards and high participation for screening.

Table 9.1 Comparison of published relative risks for dying from breast cancer in 5 Swedish randomized trials (point estimates), compared to new combined analysis with independent and blind assessment of all causes of deaths in deceased breast cancer patients, with longer follow up (all ages)

Trial	Ages	Earlier published	New analysis[1]
Malmö	45-69	0.96[2]	0.81
Kopparberg	40-74	0.64[3]	0.68
Östergötland	40-74	0.74[3]	0.84
Stockholm	40-64	0.71[4]	0.72
Göteborg	40-59	-	0.84
Kopp./Öst.	70-74	0.77[3]	0.94

Sources: [1] Nyström, 1993; [2] Andersson et al., 1988; [3] Tabár et al., 1989; [4] Frisell et al., 1991. Note: actual age at randomisation used, women excluded from cohort with breast cancer diagnosed before start, independent assessment of breast cancer deaths (new analysis)

Screening trials in the United Kingdom

Table 9.2 shows the results with longer follow up from the United Kingdom trial and the Edinburgh trial, in age-specific groups (Chamberlain, 1993). If any adjustments in predicting the breast cancer mortality reduction for screening in the Netherlands have to be made on the basis of the most recent results, they will be marginal considering the consistent evidence for the 50-69 group. The results confirm the assumptions on mortality reduction that were used in our analyses. The general conclusions from the thesis will not change.

Priorities for future research

Screening under age 50

Maybe the largest dilemma in setting priorities for (research in) breast cancer screening is the screening for women under age 50. The evidence on mortality reduction due to screening in the age group 40-49 has shown "waves" from favourable results in the classic HIP-trial started in the sixties (Shapiro et al., 1982), unfavourable results in case-control studies from trials started in the seventies (Collette et al., 1984; Verbeek et al., 1984; Palli et al., 1986), more recently a slightly favourable (but far from statistically significant) result in the large randomized Kopparberg/Östergötland trial (Tabár et al., 1992b) to unfavourable information from the Canadian trial last year (Miller et al., 1992).

Table 9.2 Relative risks of dying from breast cancer in the screening trial in Edinburgh and the United Kingdom at 10-year follow up. Columns respectively identify age group (a), woman-years in follow up in invited (b) versus control group (c), breast cancer deaths in invited (d) versus control group (e) and relative risk (95% confidence intervals) (f)

Edinburgh*					
(a)	(b)	(c)	(d)	(e)	(f)
45-49	56,000	54,000	18	20	0.86(0.41-1.8)
50-64	160,000	146,000	86	93	0.84(0.63-1.14)**
Total	216,000	200,000	104	113	0.84(0.63-1.12)
United Kingdom trial (screening centres)					
45-49					0.74(0.54-1.01)
50-54					0.94(0.69-1.3)
55-59					0.77(0.57-1.07)
60-64					0.78(0.56-1.09)
Total ***			189	227	0.80(0.67-0.95)

Source: Chamberlain, 1993

* corrected for cluster randomisation

** 7-year follow up RR=0.80 (0.54-1.17), unadjusted for cluster randomisation (Roberts et al., 1990)

*** 7-year follow up RR=0.80 (0.64-1.01) (UK Trial, 1988)

Enough information not to advise a nationwide screening programme for women under age 50.

We know that sensitivity is lower in this age group and the mean duration of detectable preclinical phase is shorter (van Oortmarssen et al., 1990b; Brekelmans et al., 1992). The existing trials may therefore not have been equipped well enough to show a mortality reduction yet (optimist) or will never show one (pessimist), since screening intervals have been too long or mammographic quality has been too poor. The dilemma has been between starting a new screening trial in the age group 40-49, or not. Other points in favour of starting an international trial is the demand or pressure from women who are not yet in the age group to be invited for the present national programme and from doctors who claim an increase in incidence in this age

group. Another point is the present "laissez-faire" attitude in which more and more screening is being applied in this group via other channels than an official screening programme which is certainly a relatively cost-ineffective situation. The United Kingdom has started a trial in which women of 40-41 years of age will be invited for annual mammographic screening during seven years. In a randomized trial, 65,000 women will be screened and compared to a control population of 130,000 women. Up to 1993, 10-15,000 women have entered the trial.

It could be useful to do a "mathematical modelling meta-analysis" of all trial results, all over the world, of breast cancer screening in this age group: starting from the favourable trials, describing the natural history of breast cancer and cancer screening in a mathematic model specifically for this age group by analyzing empirical data, and try to predict the results for the other trials. If these contradict with the published outcomes from the trials, one should look at differences in the specific characteristics that explain a less favourable result. If this is not the case, one should adapt the earlier model to a less favourable model, since apparently the organization of the second trial does not explain the difference. Such an approach would especially give a clear sight on whether the different trials after adjustment for differences in design, lead to different point estimates of mortality reduction (contradict each other) or show the same thing. Table 9.3 shows the new data from the combined analysis of the Swedish results: a non-significant mortality reduction of 13% is found after follow up of the trials for 5-13 years for women aged 40-49 at randomisation (Rutqvist, 1993). A small part of this reduction is attributable to women screened after the age of 49. Whereas in the older age groups, the difference between invited and control group is already apparent after 4 to 5 years in the Swedish trials, the (non-significant) difference in the 40-49 group becomes visible only after 10-12 years.

Evaluation and monitoring

Both pessimists and optimists are aware that the favourable results in screening trials are to a large extent (pessimists) or may be at least partly (optimists) due to the very selective high quality personnel and researchers that initiated and have worked in these trials. The start of nationwide screening is therefore no guarantee for establishing the effects expected on the basis of trials. Evaluation and monitoring of all basic national and regional information is needed to forecast and monitor mortality reduction and needed to prevent an imbalance between favourable and unfavourable effects. Especially in the first 10 years, including the build-up period of the programme, it is the only way to ascertain that screening meets its demands in reaching earlier diagnoses than without screening, eventually lowering the breast cancer mortality, and doing as less harm as possible.

Table 9.3 Relative risks of dying from breast cancer in 5 randomized Swedish trials for women aged 40-49 at entry*. Uniform combined analysis of breast cancer deaths (Rutqvist, 1993)

Trial	Woman-years		Breast cancer deaths		RR (95%)
	study	control	study	control	
	(x 1,000)				
Malmö*	46	47	8	16	0.51(0.22-1.17)
Kopparberg	107	56	26	18	0.75(0.41-1.36)
Östergötland	104	106	24	19	1.28(0.70-2.33)
Stockholm**	107	64	20	12	1.04(0.53-2.05)
Göteborg	64	77	6	10	0.73(0.27-1.97)
Total					0.87(0.63-1.20)

Note: part of relative risk due to actual screening examinations above age 49
* Malmö age 45-49
** 6-year follow up RR=1.09 (0.4-3.0) (Frisell et al., 1991)

Since screening is a service newly offered to healthy people, especially preventing an imbalance induced by screening is of the utmost importance too (the second evaluation focus). This implies evaluation of assessment and treatment at national level. The complexity of the disease and the screening, the inevitable regional differences, and the impact to be evaluated on either public health, quality of life or health care makes a difficult multidisciplinary and modelling approach, as described in this thesis, indispensable. Estimating the period that the diagnosis is brought forward (lead time), and forecasting when the increase in breast cancers has to be levelled off (if not speaking of overdiagnosis) are two additional examples, for which this approach is to be preferred.

Both in the Netherlands and in the United Kingdom, first results of nationwide screening are already available and are compared to the expected values published prior to its introduction (see Table 9.4). The early outcomes in the nationwide Dutch programme compare favourably to the expectations. The regions without a history of experimental breast cancer screening (newly screened women) also show good results: high attendance rates, high positive predictive values without negatively influencing the detection rates and a significant earlier detection of breast cancers (NETB, 1992). Women seem to appreciate the invitation for screening, and unfavourable consequences are being limited due to strict criteria for referral. In the first 3 years of the programme, the efforts of regional and national evaluation boards can already lead to the most important conclusion that the expertise from the trial projects

Table 9.4 Results from the nationwide breast cancer screening programme compared to predicted results, in first 33,000 women screened (50-70) in new regions in the Netherlands in 1990, and in first 1 million women screened in the United Kingdom (1991-1992)

	NL 50-69		UK 50-64
	Predicted	Observed[1]	Observed[2]
Attendance rate	70%	77%	70%*
Screen positives	1.6%	1.4%	6.3%
Breast cancer detection rate	6.4	7.0	6.3
Positive predictive value screening test (referral)	41%	51%	10%
Biopsies (per thousand)	11.0	9.0	8.9
Biopsies with malignant diagnoses	54%	72%	70%
Tumour size $T_{is,1}$	79%	84%	
Lymph-node metastases	26%	29%	

[1] Source: National Evaluation Team for Breast cancer screening (Rotterdam, Utrecht, Nijmegen, 1992)
[2] Chamberlain, 1993 * approximation

is transferable to a national screening programme. Also the high quality achieved supports the expectations made on breast cancer mortality reduction in the future. The same applies to the United Kingdom, but in the Netherlands the positive predictive value of referral is strikingly higher due to stricter selection criteria for further assessment and probably partially due to the different design of the programme and assessment or the 3-year interval. The biopsy rate and the detection rates in the United Kingdom and in the Netherlands are similar.

Regional evaluations are to be stimulated and may be important, e.g., in identifying possible reasons for non-attendance, the causes of interval cancers, or the specific differences compared to the national average. Interval cancers are a relative minority, but both psychologically and technologically (screening sensitivity) of utmost importance. Regional evaluations are also important in discussing or influencing the predictions at national level. E.g., the increase in assessment of non-palpable lesions with specimen radiography and paraffin section increases the workload for pathologists. The predictions about which lesions are palpable or non-palpable were difficult to make on the basis of literature. We made the assumption that lesions less than 10 mm and lesions that appear to be ductal carcinoma in situ were likely to be non-palpable, although we were aware that part of the latter lesions appear to be palpable. Analysis from the first results from the Enschede region may seem to indicate that relatively more lesions referred after screening seem to be im-

palpable than initially assumed (Boekema et al., 1992). Most importantly, regional evaluations are reflections upon their own regional achievements.

Quality of screening
Directly related to the breast cancer mortality reduction is the quality of mammographic screening. Especially the Canadian trial, hampered by relatively low and diverse quality in the beginning of the trial must, however, be praised for actually publishing on the quality (Baines et al., 1990). Quality assurance is essential for probably almost all the outcomes discussed in this thesis, and must have priority before deciding on any screening programme and priority in monitoring an on-going programme: regional differences in quality must be kept to a minimum, the effect of one-view mammography versus two-view mammography (UK Trial) should be resolved, and the minimum quality of a nationwide breast cancer screening programme should be stated more precisely on the basis of the first results from the nationwide screening programmes in the United Kingdom and the Netherlands. Several European countries are eager to offer screening to their population, but research on the favourable and unfavourable effects expected and the possibility of maintaining high quality standards of screening in countries with a more decentralised screening system is strongly needed. The differences in outcomes from different screening trials should more and more be related to the technical quality and observers' quality and such findings should be published.

Screening versus earlier clinical diagnosis without screening
The number of mammographies, outside a screening programme, has increased rapidly in all developed countries and seems unfortunately to reflect the pattern known from cervical cancer screening: many mammograms made for relatively young women, probably reflecting an increase in mammograms made without medical reasons. Until our analysis started, almost no information had been collected regularly or at subnational level on the number of mammograms made per age group and/or for what reasons. For research on this aspect at least 3 points are important: a) to investigate whether the increase in mammograms outside a screening programme does also result in earlier diagnoses of breast cancer; b) to investigate whether this is doing more or less harm to women than a nationwide breast cancer screening programme, and at which cost; and c) to analyze whether the implementation of screening for women in the age group 50-70 will induce differences in assessment in- or outside this age group. The answers on the above questions could be useful in trying to help or stimulate the medical profession and the women involved in continuation or stopping this service. Part of this research, especially on the 3rd aspect, has been initiated (de Bruin et al., 1993).

The trend towards earlier clinical diagnosis in different countries has also been used in postulating that the effect of breast cancer screening now to be introduced is less than seen in the trials, since stage distribution would have

changed (or is changing) much already. Earlier clinical diagnosis would result in a reduction of the lead time (due to mammographic screening), hence in a reduction of effectiveness. Chapter 8 showed this to be an important factor, but even with a hypothetic strong 25% improvement in prognosis for breast cancer patients (without screening), screening is assumed to be relatively cost-effective. In fact, the recently published national data on stage distribution show a 52-60% stage I and IIa distribution in 1989 (Netherlands Cancer Registry, 1992), which is to be compared to the 48% of breast cancers smaller than or equal to 2 cm in diameter in the clinical situation assumed in our analysis (based on the historic Nijmegen and Utrecht data). Nevertheless, this topic requires further analysis. The trend described in the SOOZ-region (Coebergh et al., 1990) should be simulated to show the possible reduction through such a "non-national screened" approach. This research is planned to be initiated. Other regions should try to look for these trends too, if possible.

Primary treatment

In relation to the possible side-effects of screening, annual analyses of primary treatment in the Netherlands will be more important than analysis of assessment changes due to screening: up to the year 2000 breast conserving therapy, radiation treatment and treatment of DCIS will change considerably. Chapter 6 has shown that annual analysis of hospital records, periodically supported by a survey amongst radiotherapy departments is useful. Regional differences in treatment are likely to occur after implementation of screening, as a result of different policies, different regional incidences or different preferences. It should however be monitored whether screen-detected breast cancer patients that are (oncological) alike will be treated significantly differently in different regions. Therefore, one has to be able to interpret regional differences according to oncologic status.

To obtain (annual) information on the number of breast cancer patients, both screen-detected and not, receiving adjuvant systemic treatment in the Netherlands is crucial, in order to interpret any possible breast cancer mortality reduction in the future. The increase in this treatment may have a significant independent effect on mortality reduction. Systemic adjuvant hormonal treatment in lymph node positive patients may improve the survival at 10 years from 42.2% to 50.4%. In women aged 50-69 years, Tamoxifen may lead to a reduction in the annual odds of recurrence of breast cancer of 28-33%, and polychemotherapy of 20-29% (estimated 30-40% proportional risk reduc-tions by combined chemo-endocrine therapy in middle age) (Early Breast Cancer Trialists' Collaborative Group, 1992). It is probably the only improvement in prognosis to be reached due to treatment in the intermediate future. On the other hand, in women with screen-detected cancer, it is difficult to say whether the earlier detection means diagnosing breast cancer before important micrometastases have occurred (adjuvant treatment no additional benefit) or at a time that micrometastases can still be cured (additional

benefit). Research from a random sample of patient records in different areas in the Netherlands in the years 1985 up to now would give important numerical information.

Decisions for Europe

The good results from breast cancer screening trials together with the favourable CE-analyses from the United Kingdom and the Netherlands, have led to discussions on starting nationwide breast cancer screening in almost all countries in Western-Europe. The question on cost-effectiveness: 1. estimate the effect on breast cancer mortality (reduction), 2. present an overview of other favourable and unfavourable effects, and 3. include estimations of cost, especially related to different screening policies, has been analyzed by us also for Australia (Carter et al., 1993), Italy (Paci et al., submitted) and is in progress for Germany, using the same methodology as described in this thesis.

Such country-specific studies can be important. The striking differences in breast cancer incidence and mortality even within Europe seem to make one uniform policy recommendation for all countries of the European Community inappropriate. Rough analyses have shown that the number of breast cancer deaths prevented per 1,000 screens compared to the Netherlands are much lower in, e.g., France and Spain, only due to the different epidemiology (incidence, stage distribution and mortality) in these countries (68% and 53% respectively) (van Ineveld et al., 1993; WHO, 1987; WHO, 1989; UN, 1987). Furthermore, the cost per screen appears to be higher in countries with more decentralized screening systems (Lancry and Fagnani, 1989; Australian Institute of Health, 1990) and have to be adjusted for differences in health care prices. The cost-effectiveness of the same breast cancer screening schedule is then expected to be 4.6 times higher in Spain than in the Netherlands. These analyses did not take into account possible differences to be expected in quality or attendance rates. Such rough calculations seem to indicate that priorities might have to be set differently for different countries and probably be made country-specific. The decision to first start pilot projects on breast cancer screening in the different countries of the European Community, seems to have been a good tool before deciding on future policies or health care services in each country. The good preliminary results in Navarra, for instance, with a 85% attendance rate, a detection rate of 0.55% (age 45-64) and 63% of screen-detected tumours in stage $T_{is,1}$ seem to indicate that screening might function at regional level in these countries (Aizcorbe et al., 1993). The implementation of screening in so many countries has now also led to the development of an International Breast Cancer Screening Data Base Working Group, in which the members develop an internatio-nal applicable data information system or table-book in which standard definitions and classifications will be used. Supplemented by descriptive information on the characteristics of each programme in each country, it should enhance the

comparison of screening internationally and draw inferences about, e.g., the relative effectiveness of the components of the programme in different countries or the modification of the programmes. Again cost-effectiveness analyses could complement these comparisons. One important difference (already visible) will certainly be the relatively small number of women being referred for further assessment in the Netherlands, in the order of 1-2% or less, compared to almost all other countries with a 4-7% referral rate.

Spin-offs to other research

At the end of this thesis, one is tempted to discuss whether the analyses have had further spin-offs in breast cancer or other research and whether the conclusions drawn are leading or have led to the correct decisions. In the Netherlands, one may point at several new activities that are at least initiated due to these analyses or in which they have certainly been of help: e.g., regional analyses of breast cancer screening, the evaluation tables to be produced when implementing screening, and monitoring of the total number of mammograms being made (Health Interview Survey CBS, Continuous Morbidity Registration Sentinel Stations). The detailed analysis on the cost of the screening itself is now indispensable for finding solutions on constraining the expenses for screening (van Ineveld, 1992). The methodology on cost-effectiveness for screening has also extensively been used for screening on cervical cancer (Koopmanschap et al., 1990; van Ballegooijen et al., 1992; van Oortmarssen et al., 1992). Application to screening for prostate cancer and colorectal cancer is an obvious extension. There is no reason to assume that application of this type of research is not as useful or even more for evaluation and/or decision making on screening programmes for prostate cancer or colorectal cancer. The basic information on the disease process itself (like clues on the natural history) or on the cost for the disease without screening, which this research produces too, is worthwhile.

An important spin-off is the application to other, non-cancer, screening programmes like child health surveillance, and screening for hearing disorders in pre-school children, which are one of the largest activities in preventive health care since the eighties. In fact, new screening programmes to be set up for children in the age group 0-4 years old should include an analysis on effects, favourable and unfavourable, and costs for different policies (de Koning et al., 1992b), along the lines presented in this thesis.

Politics

Actually, both the primary and final spin-off of this research should affect the politicians. This research should not only be judged on its scientific merits,

but also on its usefulness in health policy decisions. The decision process on nationwide screening for breast cancer in the Netherlands seems to have followed a long road, starting around 1977 in which year the Health Council was assigned the task to give its opinion on this subject. The final report appeared in 1987 (Health Council, 1987), together with the first results from this cost-effectiveness analysis, which is 3-4 years after the favourable published results in the Lancet from Utrecht, Nijmegen and Kopparberg/ Östergötland. Then, within a short time a national screening programme for women aged 50-70, at 2-yearly intervals, was initiated in the regions surrounding Nijmegen and Utrecht around 1988. In 1990, when most of the analyzed material was ready, 5 more regions started to screen women, in Alkmaar, Enschede, Amsterdam, Rijswijk, (my home town) Den Haag, Goes, Rotterdam and Roermond, followed by the 2 final regions in 1991. At the end of 1992, about 50% of the target group has had at least one invitation (NETB, 1993). Decisions on the screening interval, the age group invited, and the gradual build-up period for breast cancer screening have been supported by our studies. The above discussion has shown that the predictions have appeared to be realistic on the 3 main aspects: breast cancer mortality reduction achieved in the screening trials after longer follow up, first achievements on outcomes and quality of the national programme, and autonomous scenario of earlier diagnosis anyhow compared to the retrospective clinical data from Nijmegen and Utrecht. It also confirms the integrity of the independent researchers, which had been questioned once in the political debate.

New analyses by the National Evaluation Team for Breast cancer screening supported more recently the decision to make one uniform policy on the total number of invitations per woman in the age period from 50 to 70, and to stop inviting women aged 70 and older. The question of a possible improvement in prognosis for women screened above age 70 can not be answered scientifically due to a lack of screening trials inviting these women (over a long period). Furthermore, it seems possible that above the age of 80 unfavourable effects may outweigh possible favourable effects (NETB, 1992). The new data from Sweden seem even less favourable for the age group 70+ than earlier expected (9% reduction in mortality).

I hope the politicians find time to read this thesis.

Breast cancer will remain the second most important cause of death in women for a long time, but secondary prevention by high quality mammographic screening is likely to make a strong contribution in reducing mortality, and partly preventing morbidity, of the disease. It is too early to say whether primary prevention may do the same in the future. Needless to say this should be one of the ultimate aims of medical research. The 20% higher cost per screen in the present national programme than expected is of concern and is certainly enough reason to act and intervene now. One has to realize also that mammographic screening is still a costly additional service, only possible in countries with high developed (financial) standards.

10 Conclusions

1 **All breast cancer screening trials confirm the expectation that introducing breast cancer screening by means of high quality mammography for women aged 50 to 69 years will lead to a reduction in breast cancer mortality.**

Breast cancer screening trials in several countries have proven to reduce breast cancer mortality. The precise extent of the reduction depends on various factors. The results from the Swedish randomized breast cancer screening trials are most applicable to the Dutch situation. Assuming equal mammographic quality in the Swedish breast cancer screening projects in Kopparberg/Östergötland and in Malmö, a reduction of 32% in breast cancer mortality for invited women in the age group of 50 to 69 years after 8 to 9 years for these trial situations is a good estimate for similar projects. Applying this "Swedish assumption" for the improvement in prognosis after early mammographic detection to a Dutch nationwide programme, given the Dutch model for breast cancer and screening, the 2-yearly screening programme for women aged 50-69 with a 70% attendance rate will eventually lead to a reduction in breast cancer mortality of 16% in the entire population. In the year 2015 about 700 women will not die of breast cancer who would have without screening.

2 **Screening is thus justified by the improvement in quality of life of and cost savings for women prevented from reaching advanced disease.**

As nearly all women with breast cancer in an advanced stage will ultimately die from the disease, the number of women to be treated for advanced disease will decrease proportionally to the expected decrease in breast cancer mortality due to screening. On average a woman who will not die from breast cancer will gain 15 years, and a period of almost 2 years of intensive treatment, follow up and suffering from advanced disease is prevented. During the latter period, hormonal treatment is the main modality during 14 and chemotherapy during 4 months and during this period from first recurrence until death, women remain in hospital for a mean period of one and a half months. A woman with advanced breast cancer is estimated to have a 39% loss in quality of life compared to a healthy woman. When quantifying the effect of breast cancer screening, the resulting gain in

quality of life by preventing the occurrence of advanced disease, contributes 70% of the total gain in quality of life and is the most important favourable effect other than mortality reduction. In the long term, high quality screening will result in an important benefit for the women involved and in a considerable reduction of the amount of effort and money being spent to treat women with advanced breast cancer. Total medical cost from diagnosis of advanced disease until death amounts to 21,000 US dollars.

3 **When considering all other relevant effects for the women besides breast cancer mortality reduction, breast cancer screening can still be recommended.**

In discussions on breast cancer screening, much attention has been focused on the possible morbidity generated by screening. However, a balanced judgement leads to the conclusion that the negative impact on quality of life attributes no major negative issue in the decision to undertake a large-scale breast cancer screening programme and should not be overrated. Screening will have effects on quality of life in the short run for women participating, and effects in the long run as a result of the expected shift in the number of women experiencing early and advanced phases of the disease. The increase in biopsies due to false-positive referrals is often regarded as a serious problem of mass screening. One has to realize, however, that the number of referrals from general practitioners to out-patient departments of women outside the programme is expected to diminish after the start of the screening. The most important unfavourable effect in terms of loss in quality of life for the total group of women involved is the large increase in woman-years in follow up. To assume a small loss in quality of life during years in follow-up is justified since the woman will undergo several medical follow-up procedures and will experience a negative impact on quality of life resulting from breast surgery and the knowledge that she has had breast cancer. The period that the diagnosis is brought forward compared to the situation without screening (lead time), is on average 4 years in the 2-yearly programme for women aged 50-69, for all these women irrespective of the fact whether they will benefit from the earlier detection. Those patients who will profit experience a longer disease-free interval, with the knowledge of having had breast cancer. The prevention of the occurrence of advanced disease is the most important favourable effect, followed by the decrease in adjuvant systemic treatment, due to the smaller number of women with lymph-node metastases.

4 **The temporary increase in referrals and women diagnosed as having breast cancer will have an impact on workload in the first years of screening.**

The 2-yearly screening programme for women aged 50-69 is predicted to initially result in a sharp rise in the number of newly diagnosed cases. A 20% increase in breast surgery and in breast radiotherapy due to screening in the first years is expected. The 45% increase in assessment procedures for non-palpable lesions in the build-up period gives an even higher workload for the surgeons and pathologists. The decreasing number of women with advanced disease will have the opposite effect on workload only after several years. This decrease in systemic and palliative treatments results in a prevention of drug treatment and in a decrease in hospital admissions and in-patient days, primarily noticed by the internist. The changes in palliative surgery or palliative radiotherapy will be small in contrast to the large and immediate changes in primary treatment as a result of the introduction. Actions should be taken beforehand in hospital and radiotherapy departments surrounding a starting screening unit in order to handle the increasing workload.

5 The actual number of women treated conservatively is rapidly growing irrespective of screening, and this number will increase furthermore due to screening.

Since 1985, there has been a rapid and steady increase in the proportion of breast conserving therapy in the Netherlands, resulting in 28% in 1988 and about 36% of women in 1990 (=2900). Continuation of this trend would mean that around 1992 the expected number would be reached that may be postulated on the basis of oncologic data and the treatment protocols. Data shows that breast conserving surgery is always followed by radiotherapy, whereas there has been a clear reduction in postoperative radiation after mastectomy in 1986-1988.

Due to the implementation of screening, the increase in breast conserving therapy will be influenced even more. Screening will lead to an important shift in primary treatment modalities, as 10-15% of mastectomies will be replaced by breast conserving therapy. Treatment for ductal carcinoma in situ will increase by 250 a year, but will remain a small fraction of the total number of diagnosed breast cancers. Apart from the increase in the demand for primary treatment, breast conserving therapy requires more time and effort from surgical and radiotherapy departments. At national level, we expect a steady increase in breast conserving therapy to an increase up to 34% (+1,025) in 1994 due to screening. After the first screening years, the situation stabilizes at approximately +22% per year. There will be a sharper rise at the regional level, only it lasts less longer.

6 Screening of women aged 50-69 years by mammography at 2-yearly intervals is a relatively cost-effective schedule.

In the absence of screening, the total medical cost (in 1990 prices) of breast cancer is calculated to be 340 mln Dutch guilders in 1994: 22% for assessment of women to confirm or exclude breast malignancy, 36% for primary treatment and follow up of patients, and 42% for treatment of advanced disease. Initially, the total costs of screening (in 1990 prices) have been estimated to be on average 80 Dutch guilders in 1994 per screen: the cost of screening amounts to more than 40 mln Dutch guilders in 1994. In the long run, 47% of the annual cost of screening will be offset by savings due to a decrease in the number of women with advanced disease. If costs generated by screening (false positives, more primary treatment and longer follow up) are taken into account, one third of the cost of screening will be countered by savings in the Dutch 2-yearly screening programme for women aged 50-69 years. The cost-effectiveness (CE) ratio is 3,825 US dollars per life-year gained (5% discount rate). Variation of most variables keeps the CE ratio limited to the range of 3,000 to 5,000 US dollars per life-year gained, at purchasing power rates of 1990. Cost per breast cancer death prevented or life-year gained is much lower than for most other medical interventions for which cost-effectiveness ratios are known, screening for cervical cancer included. The absolute number of breast cancer deaths prevented is higher than the total mortality from cervical cancer in the Netherlands.

If the United Kingdom policy, of offering screening to women aged 50-65 at 3-yearly intervals, would have been implemented in the Netherlands, the expected breast cancer mortality reduction would only be 11% per year. The smaller age range invited and the fact that women will have less screens with a 3-year interval (given the same time period) is responsible for this difference. In terms of cost-effectiveness, both policies are relatively cost-effective: the United Kingdom policy, applied to the Dutch situation, results in a CE-ratio of 3,235 US dollars. In fact, screening policies of even less screens in the age group 50-69 are slightly more cost-effective, but there is a trade-off with the reduction in breast cancer mortality to be achieved. A 4-year interval would result in only a 10% reduction instead of 16%.

However, further decreasing the screening interval to 1 year or less is far less cost-effective. The total number of life-years gained increases, but not in a linear fashion as the costs do. Each shortening of interval leads to much less additional effect on life-years gained than it results in additional cost to be invested, and is not to be preferred.

7 **At present, nationwide screening for women below 50 is not to be recommended. Further analysis of the most recent data from the trials in Sweden and the United Kingdom and setting up a 40-49 trial might be necessary.**

Controversy on the effectiveness of screening for women below 50 is strong, but all publications suggest that there is no short-term mortality reduction in the

present trials. When using the (non-significant) 8% breast cancer mortality reduction obtained by screening women under age 50 in Kopparberg/Östergötland in a policy of inviting women aged 40-69 at 2-yearly intervals, it still leads to a very unfavourable balance in terms of marginal cost-effectiveness and cost-utility. The additional cost per additional life-year gained is 35,000 US dollars. A new breast cancer screening trial has started for annual screening of women aged 40-41 at seven intervals. Due to a lack of important data concerning screening above age 69, it is difficult to estimate the relationship between effects and costs above this age group.

8 **National evaluations show that the expertise from pilot projects is transferable, and they support the predictions made on the effects and costs of breast cancer screening.**

A relatively strong impact on the results of cost-effectiveness and quality of life are due to the screening performances, the demand for mammograms outside screening and the possibility of a survival improvement irrespective of screening, which makes assessing the CE of breast cancer screening important research in the following years. The first results from the nationwide screening programmes in the United Kingdom and the Netherlands prove that the expertise from the experimental projects is transferable, and support the predictions made on breast cancer mortality reduction.

High quality mammographic screening of women aged 50-69, together with high quality assessment and therapeutic standards, can contribute considerably to the health and well-being of many women, and deserves full support.

Literature

Adami H-O, Malker B, Holmberg L, Persson I and Stone B. The relation between survival and age at diagnosis in breast cancer. N Engl J Med 1986;315:559-663

Aizcorbe M, Illarramendi J, Dominguez F and Cobo F. Intraductal carcinomas detected in a population screening program: a breast unit experience. Poster presented on the EUSOMA Consensus Conference on screening methodology and management of occult breast carcinomas, February 4th & 5th, Paris, 1993

Andersson I, Aspegren K, Janzon L, Landberg T, Lindholm K, Linell F, Ljungberg O, Ranstam J and Sigfusson B. Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial. Br Med J 1988;297:943-948

Andersson I. Presentation European Group for Breast Cancer Screening. Strasbourg, June 1989a

Andersson I, Aspegren K and Janzon L. Letter to the editor. Br Med J 1989b;298:49

Andersson I. Presentation European Group for Breast Cancer Screening. Rhodes, May 1992

Ashby J, Buxton M and Gravelle H. Will a breast screening programme change the workload and referral practice of general practitioners? J Epidemiol & Comm Health 1990;44:36-38

Atkinson EN, Brown BW and Montague ED. Tumor volume, nodal status, and metastasis in breast cancer in women. J Nat Cancer Inst 1986;76:171-178

Australian Institute of Health. Prevention program evaluation series no 1. Breast cancer screening in Australia: future directions. Australian Government Publishing Service. Canberra, 1990

Baines CJ, Miller AB, Kopans DB, Moskowitz M, Sanders DE, Sickles EA, To T and Wall C. Canadian National Breast Screening Study: assessment of technical quality by external review. AJR 1990;155:743-747

Baker MS, Kessler LG, Urban N and Smucker RC. Estimating the treatment costs of breast and lung cancer. Med Care, 1991;29:40-49

Ballegooijen M van, Habbema JDF, Oortmarssen GJ van, Koopmanschap MA, Lubbe JThN and Agt HME van. Preventive pap-smear: balancing costs, risks and benefits. Br J Cancer 1992;65:930-933

Bartelink H, Borger JH, Dongen JA van and Peterse JL. The impact of tumor size and histology on local control after breast-conserving therapy. Radiother Oncol 1988;11:297-303

Bennett IC, McCaffrey JF, Baker CA, Burke MF, Lee JF and Balderson GA. Changing patterns in the presentation of breast cancer over 25 years. Aust N Z J Surg 1990;60:665-671

Benraadt J, Leeuwen FE van and Visser O. Jaarverslag Kankerregistratie 1989. IKA Stedendriehoek Twente. Enschede, 1991

Bjurstam N, Cahlin E, Eriksson O, Hafström L-O, Rudenstam C-M and Södenbergh JS. Breast Cancer

Screening Project, Göteborg, Sweden. An efficient program for breast screening by mammography. Experiences from two screening rounds. 4th EORTC Breast Cancer Working Conference 30th June-3rd July 1987. Abstract Book pp. E16

Boekema AG, Dornseiffen G, Mulder HJ, Vos RAI de and Kluft-de Haas BA. De eerste ervaring met mammascreening in de regio Enschede. I. Effectiviteit. Ned Tijdschr Geneeskd 1992;136(36):1757-1760

Bombardier C, Wolfson AD, Sinclair AJ and McGreer A. Comparison of three preference measurement methodologies in the evaluation of a functional status index. In: R Deber, G Thompson (eds.). Choices in health care - decision making and evaluation of effectiveness. University of Toronto, Toronto, 1982. (cited by G Loomes. Disparities between health state measures: an explanation and some implications. CHESG paper, Centre for Health Economics, York, 1988)

Brekelmans CTM, Collette HJA, Collette C, Fracheboud J and Waard F de. Breast cancer after a negative screen: follow-up of women participating in the DOM screening programme. Eur J Cancer 1992;28A:893-895

Bruin A de, Koning HJ de and Ballegooijen M van. Cervix uitstrijkjes en mammografieën, Gezondheids-enquete 1991. Mndber Gezondheid 1993, in press

Buxton M, Ashby J and O'Hanlon M. Alternative methods of valuing health states: a comparative analysis based on an empirical study using the time trade-off approach in relation to health states one year after treatment for breast cancer. Paper presented to the Third Annual Meeting of the International Society of Technology Assessment in Health Care, Rotterdam, 1987

Carter R, Glasziou P, Oortmarssen GJ van, Koning HJ de, Stevenson C, Salkeld G and Boer R. Cost-effectiveness of mammographic screening in Australia. Austr J Public Health 1993;17:42-50

Cassileth BR, Lusk EJ, Strouse TB, Miller DS, Brown LL, Cross PA and Tenaglia AN. Psychological status in chronic illness, a comparitive analysis of six diagnostic groups. N Engl J Med 1984;311:506-511

Centraal Bureau voor de Statistiek. Mndber Gezondheid 1991:9;34-35

Chamberlain J. The United Kingdom and Edinburgh trial. Presentation NCI-international workshop on screening for breast cancer. Bethesda, February 1993

Charro, FThM de. A cost-effectiveness analysis of a programme for end stage renal disease. Thesis (in Dutch). Eburon, Delft, 1988

Chu KC, Smart CR and Tarone RE. Analysis of breast cancer mortality and stage distribution by age for the Health Insurance Plan clinical trial. J Natl Cancer Inst 1988;80:1125-1132

Coebergh JWW, Crommelin MA, Kluck HM, Beek M van, Horst F van der and Verhagen-Teulings MTh. Borstkanker in Zuidoost-Noord-Brabant en Noord-Limburg; beloop van incidentie en vervroeging van de diagnose in een niet-gescreende vrouwelijke bevolking, 1975-1986. Ned Tijdschr Geneeskd 1990;134(15):760-765

Collette HJA, Rombach JJ, Day NE and Waard F de. Evaluation of screening for breast cancer in a non-randomized study (the DOM project) by means of a case-control study. Lancet 1984;i:1224-1226

Collette C, Collette HJA, Fracheboud J, Slotboom BJ and Waard F de. Evaluation of a breast cancer screening programme - the DOM project. Eur J Cancer 1992;28A(12):1985-1988

Dam FS van and Bergman RB. Psychosocial and surgical aspects of breast reconstruction. Eur J Surg Oncol 1988;14:141-149

Day NE and Miller AB (eds.) Screening for breast cancer. Hans Huber Publishers, Toronto, 1988

Day NE. Quantitative approaches to the evaluation of screening programs. World J Surg 1989a;13:3-8

Day NE, Williams DRR and Khaw KT. Breast cancer screening programmes: the development of a monitoring and evaluation system. Br J Cancer 1989b;59:954-958

Dean C, Roberts MM, French K and Robinson S. Psychiatric morbidity after screening for breast cancer. J Epidemiol & Comm Health 1986;40:71-75

Dixon AR, Ellis IO, Elston CW and Blamey RW. A comparison of the clinical metastatic patterns of invasive lobular and ductal carcinomas of the breast. Br J Cancer 1991;63:634-635

Dixon JM and John TG. Morbidity after breast biopsy for benign disease in a screened population [Letter]. Lancet 1992;339:128

Dongen JA van, Bartelink H, Aaronson, N & 16 others. Randomised clinical trial to assess the value of breast conserving therapy in stage I and II breast cancer (trial 10801). 4th EORTC Congress London, June 30th-July 3rd, 1987. Imperial College: London, 1987

Dongen JA van, Harris JR, Peterse JL. In-situ breast cancer: the EORTC consensus meeting. Lancet 1989;i:25-27

Dutch National Council for Public Health (Nationale Raad voor de Volksgezondheid). Advice organisation mass screening for breast cancer. Report. (In Dutch). Zoetermeer, 1987

Eardley A. and Elkind A. A pilot study of attendance for breast cancer screening. Soc Sci Med 1990;30:693-699

Early Breast Cancer Trialists' Collaborative Group. Effects of adjuvant Tamoxifen and cytotoxic therapy on mortality in early breast cancer. N Engl J Med;1988:319:1681-1692

Early Breast Cancer Trialists' Collaborative Group. Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy. Lancet 1992;339:1-15

Edland RW. Presidential address: does adjuvant radiotherapy have a role in the postmastectomy management of patients with operable breast cancer - revisited. Int J Radiation Oncology Biol Phys 1988;15:519-535

Ellman R, Angeli N, Christians A, Moss S, Chamberlain J and Maguire P. Psychiatric morbidity associated with screening for breast cancer. Br J Cancer 1989;60:781-784

EORTC (European Organization for Research and Treatment of Cancer). Phase III study in the conservative management of breast carcinoma by tumorectomy and radiotherapy: assessment of the role of booster dose of radiotherapy. Protocol nr. 10882, 1989

Fallowfield LJ, Hall A, Maguire GP and Baum M. Psychological outcomes of different treatment policies in women with early breast cancer outside a clinical trial. Br Med J 1990;301:575-580

Farrow DC, Hunt WF, Samet JM. Geographic variation in the treatment of localized breast cancer. N Engl J Med 1992;326:1097-1101

Fisher B, Redmond C, Poisson R, Margolese R, Womark N, Wickerham L and 9 others. Eight-year results of a randomized clinical trial comparing total mastectomy and lumpectomy with or without irradiation in the treatment of breast cancer. N Engl J Med 1989;320:822-828

Forrest, P. Breast Cancer Screening. Report to the Health Ministers of England, Wales, Scotland and Northern Ireland by a working group chaired by Prof. Sir Patrick Forrest, London, 1986

Frisell J, Glas U, Hellström L and Somell A. Randomized mammographic screening for breast cancer in Stockholm. Design, first round results, and comparison. Breast Cancer Res Treat 1986;8:45-54

Frisell J, Eklund G, Hellström L, Glas U and Somell A. The Stockholm Breast Cancer Screening Trial - 5-year results and stage at discovery. Breast Cancer Res Treat 1989;13:79-87

Frisell J, Eklund G, Hellström L, Lidbrink E, Rutqvist L-E and Somell A. Randomized study of mammography screening - preliminary report on mortality in the Stockholm trial. Breast Cancer Res Treat 1991;18:49-56

Gram IT, Lund E and Slenker SE. Quality of life following a false positive mammogram. Br J Cancer 1990;62:1018-1022

Gravelle HSE, Simpson PR and Chamberlain J. Breast cancer screening and health service costs. J Health Economics 1982;1:185-207

Habbema JDF, Oortmarssen GJ van, Putten DJ van, Lubbe JThN and Maas PJ van der. Age-specific reduction in breast cancer mortality by screening: an analysis of the results of the Health Insurance Plan of Greater New York Study. J Natl Cancer Inst 1986;77:317-320

Habbema JDF, Lubbe JThN, Oortmarssen GJ van and Maas PJ van der. A simulation approach to cost-effectiveness and cost-benefit calculations of screening for early detection of disease. Eur J Oper Res 1987;29:159-166

Habbema JDF and Bonsel GJ. The costs and effects of liver transplantation. An analysis of the Dutch program 1977-1987. Report (In Dutch). Erasmus Universiteit, Rotterdam, 1988a

Habbema JDF, Lubbe JThN, Agt HME van, Ballegooijen M van, Koopmanschap MA and Oortmarssen GJ van. The costs and effects of screening for cervical cancer. Report (In Dutch). Erasmus Universiteit, Rotterdam, 1988b

Haes JCJM de and Knippenberg FCE van. Quality of life instruments for cancer patients: "Babel's Tower Revisited". J Clin Epidemiol 1989;42:1239-1241

Haes JCJM de, Knippenberg FCE van and Neijt JP. Measuring psychological and physical distress in cancer patients: structure and application of the Rotterdam Symptom Checklist. Br J Cancer 1990;62:1034-1038

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

Harris JR, Hellman S and Kinne DW. Limited surgery and radiotherapy for early breast cancer. N Engl J Med 1985;313:1365-1368

Harris JR, Hellman S. Put the "hockey stick" on ice. Int J Radiation Oncol Biol Phys 1988;15:497-499

Harris JR, Lippman ME, Veronesi U and Willett W. Breast Cancer. Part I, II, III. N Engl J Med 1992;327:319-328, 390-398, 473-480

Haybittle JL, Brinkley D, Houghton J, A'Hern RP and Baum M. Postoperative radiotherapy and late mortality: evidence from the Cancer Research Campaign trial for early breast cancer. Br Med J 1989;298:1611-1614

Hayward JL and Rubens RD. UICC multidisciplinary project on breast cancer. Management of early and advanced breast cancer. Meeting held in Guernsey, Channel Islands, United Kingdom, on April 22-24, 1987. Int J Cancer 1987;39:1-5

Health Council (Gezondheidsraad). The early detection of breast cancer. Report by an Expert Committee (In Dutch). 's-Gravenhage, 1987

Health Insurance Executive Board (Ziekenfondsraad). Advice to the Ministry of Welfare, Health and Cultural Affairs: Mass screening for breast cancer (In Dutch). Amstelveen, 24 March 1988.

Hendriks JHCL. Population screening for breast cancer by means of mammography in Nijmegen 1975-1980. Thesis. Katholieke Universiteit Nijmegen, 1982

Hobbs P, Kay C, Friedman EHI, Leger ASSt, Lambert C, Boggis CRM, Howard TM and Asbury DL. Response by women aged 65-79 to invitation for screening for breast cancer by mammography: a pilot study. Br Med J 1990;301:1314-1316

Holland R, Veling SHJ, Mravunac M and Hendriks JHCL. Histologic multifocality of T_{is}, T_{1-2} breast carcinomas, implications for clinical trials of breast-conserving surgery. Cancer 1985;56:979-990

Holland R, Connnolly JL, Gelman R, Mravunac M, Hendriks JHCL, Verbeek ALM, Schnitt SJ, Silver B, Boyages J and Harris JR. The presence of an extensive intraductal component following a limited excision correlates with prominent residual disease in the remainder of the breast. J Clin Oncol 1990;8:113-118

Holli K and Hakama M. Treatment of the terminal stages of breast cancer. Br Med J 1989;298:13-14

Holmberg L, Adami HO, Persson I, Lundstrom T and Tabár L. Demands on surgical inpatient services after mass mammographic screening. Br Med J 1986a;293:779-782

Holmberg L, Tabár L, Adami H-O and Bergstrom R. Survival in breast cancer diagnosed between mammographic screening examinations. Lancet 1986b;ii:27-30

Hooning MJ, Dongen JA van, Went G. Changing indications for breast conserving therapy: proportion of patients with operable breast cancer suitable for breast conservation. Neth J Surg 1991;43(4):102-104

Hopwood P, Howell A and Maguire P. Psychiatric morbidity in patients with advanced cancer of the breast: prevalence measured by two self-rating questionnaires. Br J Cancer 1991;64:349-352

Horak ER, Leek R, Klenk N, LeJeune S, Smith K, Stuart N, Greenall M, Stepniewska K and Harris AL. Angiogenesis, assessed by platelet/endothelial cell adhesion molecule antibodies, as indicator of node metastases and survival in breast cancer. Lancet 1992;340:1120-1124

Høst H and Lund E. Age as a prognostic factor in breast cancer. Cancer 1986;57:2217-2221

Ineveld BM van. Het bevolkingsonderzoek naar borstkanker: Hoe doeltreffend én doelmatig? Kanker 1992;16(5):208-212

Ineveld BM van, Oortmarssen GJ van, Koning HJ de, Boer R and Maas PJ van der. How cost-effective is breast cancer screening in different EC-countries? Eur J Cancer, in press

Integraal Kanker Centrum Oost-Nederland (IKO) and Dronkers DJ. Annual report on breast cancer screening in 1989 (in Dutch). Nijmegen, 1990

Kallioniemi O-P, Kärkkäinen A, Auvinen O, Mattila J, Koivula T and Hakama M. DNA flow cytometric analysis indicates that many breast cancers detected in the first round of mammographic screening have

a low malignant potential. Int J Cancer 1988;42:697-702

Keeler EB and Cretin S. Discounting of life-saving and other nonmonetary effects. Management Sci 1983;29:300-306

Kiebert GM, Haes JCJM de and Velde CJH van de. The impact of breast-conserving treatment and mastectomy on the quality of life of early-stage breast cancer patients; a review. J Clin Oncol 1991;9(6):1059-1070

Knox EG. Evaluation of a proposed breast cancer screening regimen. Br Med J 1988;297:650-654

Koning HJ de, Boer R, Maas PJ van der, Ineveld BM van, Collette HJA and Hendriks JHCL. Effectiviteit van bevolkingsonderzoek naar borstkanker; sterftereductie in binnen- en buitenland. NTvG 1990a;134:2240-2245

Koning HJ de, Ineveld BM van, Oortmarssen GJ van, Boer R, Collette HJA, Verbeek ALM, Hendriks JHCL, Bruyn AE de, Agt HME van, Haes JCJM de, Zwan L van der, Koopmanschap MA, Deurenberg JJM, Habbema JDF and Maas PJ van der. The costs and effects of mass screening for breast cancer. Final report (In Dutch). Erasmus Universiteit, Rotterdam, 1990b

Koning HJ de, Oortmarssen GJ van, Ineveld BM van and Maas PJ van der. Breast cancer screening: its impact on clinical medicine. Br J Cancer 1990c;61:292-297

Koning HJ de, Ineveld BM van, Oortmarssen GJ van, Haes JCJM de, Collette HJA, Hendriks JHCL and Maas PJ van der. Breast cancer screening and cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors. Int J Cancer 1991;49:531-537

Koning HJ de, Ineveld BM van, Haes JCJM de, Oortmarssen GJ van, Klijn JGM and Maas PJ van der. Advanced breast cancer and its prevention by screening. Br J Cancer 1992a;65:950-955

Koning HJ de, Juttmann RE, Panman J, Verzijl JG, Meulmeester JF, Oortmrssen GJ van, Bruyn AE de and Maas PJ van der. Cost-effectiveness analysis in child health care for 0-4 years: method and possibilities. (In Dutch). Report. Erasmus Universiteit, Rotterdam, 1992b

Koopmanschap MA, Oortmarssen GJ van, Agt HMA van, Ballegooijen M van, Habbema JDF and Lubbe JThN. Cervical cancer screening: attendance and cost-effectiveness. Int J Cancer 1990;45:410-415

Koopmanschap MA, Ineveld BM van and Miltenburg EM. Costs of home care for advanced breast and cervical cancer in relation to cost-effectiveness of screening. Soc Sci Med 1992;35:979-985

Koscielny S, Tubiana M, Lê MG, Valleron AJ, Mouriesse H, Contesso G and Sarrazin D. Breast cancer: relationship between the size of the primary tumour and the probability of metastatic dissemination. Br J Cancer 1984;49:709-715

Lancry JP and Fagnani F. Evaluation Economique du dépistage systematique des cancers du sein et du col de l'uterus. COMAC-HSR, Paris, 1989

Lazovich D, White E, Thomas DB and Moe RE. Underutilization of breast-conserving surgery and radiation therapy among women with stage I or II breast cancer. J Am Med Ass 1991;266(24):3433-3438

Levitt SH. Is there a role for post-operative adjuvant radiation in breast cancer? Beautiful hypothesis versus ugly facts: 1987 Gilbert H Fletcher Lecture. Int J Radiation Oncol Biol Phys 1988;14:787-796

Levy SM, Herberman RB, Lee JK, Lippman ME, and d'Angelo T. Breast conservation versus mastectomy: distress sequelae as a function of choice. J Clin Oncol 1989;7:367-375

Llewellyn-Thomas M, Sutherland HJ, Tibshirani R, Ciampi A, Till JE and Boyd NF. Describing health states, methodologic issues in obtaining values for health states. Med Care 1984;22:543-552

Ludbrook A. Breast cancer screening. Lancet 1987;i:1153

Maas PJ van der, Koning HJ de, Ineveld BM van, Oortmarssen GJ van, Habbema JDF, Lubbe JThN, Geerts AT, Collette HJA, Verbeek ALM, Hendriks JHCL and Rombach JJ. The cost-effectiveness of breast cancer screening. Int J Cancer 1989;43:1055-1060

Miller AB, Chamberlain J, Day NE, Hakama M and Prorok PC. Report on a workshop of the UICC project on evaluation of screening for cancer. Int J Cancer 1990;46:761-769

Miller AB, Baines CJ, To T and Wall C. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years and 2. Breast cancer detection and death rates among women aged 50 to 59 years. Can Med Assoc J 1992;147(10):1459-1476 + 1477-1488

NETB (National Evaluation Team for Breast cancer screening). National evaluation of mass screening for breast cancer. Report (In Dutch) juni 1992 (I). Erasmus Universiteit, Rotterdam, 1992

NETB (National Evaluation Team for Breast cancer screening). National evaluation of mass screening for breast cancer. Report (In Dutch). Erasmus Universiteit, Rotterdam, 1993 (in press)

Nerenz DR, Love RR, Leventhal H and Easterling DV. Psychosocial consequences of cancer chemotherapy for elderly patients. Health Serv Res 1986;20/6:961-976

Netherlands Cancer Registry. Incidence of cancer in the Netherlands 1989. Report. Utrecht, 1992

Nyström L. Combined analysis of Swedish trials. Presentation NCI-international workshop on screening for breast cancer. Bethesda, February 1993. (Nyström L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Rydén S, Andersson I, Bjurstam N, Fagerberg G, Frisell J, Tabár L and Larsson L-G. 1993, in press)

OECD, Main economic indicators, pp. 174-175, OECD, Paris, 1987

Oortmarssen GJ van, Habbema JDF, Lubbe JThN and Maas PJ van der. A model-based analysis of the HIP project for breast cancer screening. Int J Cancer 1990a;46;207-213

Oortmarssen GJ van, Habbema JDF, Maas PJ van der, Koning HJ de, Collette HJA, Verbeek ALM, Geerts AT and Lubbe JThN. A model for breast cancer screening. Cancer 1990b;66:1601-1612

Oortmarssen GJ van, Habbema JDF and Ballegooijen M van. Predicting mortality from cervical cancer after negative smear test results. Br Med J 1992;305:449-451

Osteen RT, Steele GD, Menck HR and Winchester DP. Regional differences in surgical management of breast cancer. CA - A Cancer Journal for Clinicians 1992;42:39-43

Paci E, Boer R, Zappa M, Koning HJ de, Oortmarssen GJ van, Crocetti E, Giorgi D, Rosselli del Turco M and Habbema JDF. A model based prediction of the impact on mortality reduction of a mammographic screening programme in the city of Florence - Italy. Submitted, 1992

Palli D, Rosselli del Turco M, Buiatti E, Carli S, Ciatto S and Toscani L. A case-control study of the efficacy of a non-randomized breast cancer screening program in Florence (Italy). Int J Cancer 1986;38:501-504

Palli D, Rosselli del Turco M, Buiatti E, Ciatto S, Crocetti E and Paci E. Time interval since last test in a breast cancer screening programme: a case-control study in Italy. J Epidemiol & Comm Health

1989;43:241-248

Patel JK, Nemoto T, Vezeridis M, Petrelli N, Suh O and Dao TL. Does more intense palliative treatment improve overall survival in metastatic breast cancer patients? Cancer 1986;57:567-570

Peeters PH. The extent of side-effects in screening for breast cancer. Thesis. Nijmegen, 1989

Perez JE, Machiavelli M, Leone BA, Romero A, Rabinovich MG, Vallejo CT, Bianco A, Rodriguez R, Cuevas MA and Alvarez LA. Bone-only versus visceral-only metastatic pattern in breast cancer: analysis of 150 patients. Am J Clin Oncol 1990;13:294-298

Powles TJ, Smith IE, Ford HT, Coombes RC, Mary Jones J and Gazet J-C. Failure of chemotherapy to prolong survival in a group of patients with metastatic breast cancer. Lancet 1980;i:580-582

Roberts MM. Breast screening: time for a rethink? Br Med J 1989;299:1153-1155

Roberts MM, Alexander FE, Anderson TJ, Chetty U, Donnan PT, Forrest P, Hepburn W, Huggins A, Kirkpatrick AE, Lamb J, Muir BB and Prescott RJ. Edinburgh trial of screening for breast cancer: mortality at seven years. Lancet 1990;335:241-246

Rombach JJ. Breast cancer screening; results and implications for diagnostic decision making. Thesis, Rijksuniversiteit Utrecht, 1980

Rosen PP. The pathological classification of human mammary carcinoma: past, present and future. Ann Clin Lab Sci 1979;9:144-156

Rosner D, Bedwani RN, Vana J, Baker HW and Murphy GP. Noninvasive breast carcinoma. Results of a national survey by the American College of Surgeons. Ann Surg 1980;192:139-147

Russell LB. Evaluating preventive care. Report on a workshop. The Brookings Institution, Washington DC, 1987

Rutqvist LE, Miller AB, Andersson I, Hakama M, Hakulinen T, Sigfusson BF and Tabár L. Reduced breast-cancer mortality with mammography screening - an assessment of currently available data. Int J Cancer Suppl 1990;5:76-84

Rutqvist LE. Interpretation of the combined analysis of Swedish trials. Presentation NCI-international workshop on screening for breast cancer. Bethesda, February 1993.

Schmidt JG. The epidemiology of mass breast cancer screening - a plea for a valid measure of benefit. J Clin Epidemiol 1990;43(3):215-225

Schnitt SJ, Silen W, Sadowsky NL, Connolly JL and Harris JR. Ductal carcinoma in situ (intraductal carcinoma) of the breast. N Engl J Med 1988;318:898-903

Shapiro S, Venet W, Strax P, Venet L and Roeser R. Ten- to fourteen-year effect of screening on breast cancer mortality. J Natl Cancer Inst 1982;69:349-355

Shapiro S, Venet W, Strax P and Venet L. Current results of the breast cancer screening randomized trial: the Health Insurance Plan (HIP) of Greater New York study. In: Screening for breast cancer. Day NE and Miller AB (eds.). Hans Huber Publishers, Toronto, 1988

Skrabanek, P. The debate over mass mammography in Britain. The case against. Br Med J 1988;297:971-972

Slevin ML, Stubbs L, Plant HJ, Wilson P, Gregory WM, Armes PJ and Downer SM. Attitudes to

chemotherapy: comparing views of patients with cancer with those of doctors, nurses, and general public. Br Med J 1990;300:1458-1460

Steering Committee on Future Health Scenarios. Cancer in the Netherlands, volume I, 1988

Tabár L, Gad A, Holmberg LH, Ljungquist U, Fagerberg CJG, Baldetorp L, Grontoft O, Lundstrom B, Manson JC, Eklund G and Day NE. Reduction in mortality from breast cancer after mass screening with mammography. Lancet 1985;i:829-832

Tabár L, Duffy SW and Krusemo UB. Detection method, tumour size and node metastases in breast cancers diagnosed during a trial of breast cancer screening. Eur J Cancer Clin Oncol 1987a;23:959-962

Tabár L, Fagerberg G, Day NE and 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 1987b;55:547-551

Tabár L, Fagerberg CJG and Day NE. Evaluation of the results. In: NE Day, AB Miller (eds). Screening for breast cancer. Hans Huber Publishers, Toronto, 1988

Tabár L, Fagerberg G, Duffy SW and Day NE. The Swedish two county trial of mammographic screening for breast cancer: recent results and calculation of benefit. J Epidemiol & Comm Health 1989;43:107-114

Tabár L, Fagerberg G, Day NE, Duffy SW, Kitchin RM. Breast cancer treatment and natural history: new insights from results of screening. Lancet 1992a;339:412-414

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

Tijmstra Tj. Medisch vernuft heeft ook een schaduwzijde. Nieuwsblad van het Noorden, 5 maart 1988

Toonder M. Het kukel. In: Als je begrijpt wat ik bedoel. De Bezige Bij, Amsterdam, 1967

Tomin R and Donegan WL. Screening for recurrent breast cancer - its effectiveness and prognostic value. J Clin Oncol 1987;5:62-67

Torrance GW. Social preferences for health states: an empirical evaluation of three measurement techniques. Socio-Econ Plan Sci 1976;10:129-136

Torrance GW. Measurement of health state utilities for economic appraisal. J Hlth Economics 1986;5:1-30

Trouw JM. Cost-effectiveness of breast cancer screening by means of mammography. Report (In Dutch). Rijksuniversiteit Limburg, Maastricht, 1986

Tubiana M and Koscielny S. Natural history of human breast cancer: recent data and clinical implications. Breast Cancer Res Treat 1991;18:125-140

UK Trial of Early Detection of Breast Cancer Group. First results on mortality reduction in the UK trial of early detection of breast cancer. Lancet 1988;ii:411-416

UK Working Party on Breast Cancer Screening. Screening for breast cancer. Lancet 1987;1:575-576

UN (United Nations). Demographic Yearbook 1987. New York, 1987

Uppsala-Örebro Breast Cancer Study Group. Sector resection with or without postoperative radiotherapy

for stage I breast cancer: a randomized trial. J Nat Cancer Inst 1990:82:277-282

Vandenbroucke JP and Hofman A. Vroege opsporing van ziekte. In: Grondslagen der epidemiologie. Wetenschappelijke Uitgeverij Bunge, Utrecht, 1990, pp. 61-73

Verbeek ALM, Hendriks JHCL, Holland R, Sturmans F, Mravunac M and Day NE. Reduction of breast cancer mortality through mass screening with modern mammography. First results of the Nijmegen project, 1975-1981. Lancet 1984;i:1222-1224

Veronesi U, Saccozzi R, Del Vecchio M, Banfi A, Clemente C, De Lena M and 9 others. Comparing radical mastectomy with quadrantectomy, axillary dissection, and radiotherapy in patients with small cancer of the breast. N Engl J Med 1981;305;6-11

Vessey, M. Breast cancer screening 1991: evidence and experience since the Forrest report. A report of the Department of Health Advisory Committee. NHS Breast Screening Programme. NHSBSP Publications, Sheffield, 1991

Waard F de, Rombach JJ and Collette HJA. The DOM-project for the early detection of breast cancer, Utrecht, the Netherlands. J Chron Dis 1984;37:1-44

Waard F de, Collette HJA and Rombach JJ. The DOM-project for the early detection of breast cancer 1983-1984. Part 3. Report (In Dutch). Preventicon, Utrecht, 1986

Wald N, Frost C and Cuckle H. Breast cancer screening: the current position. Br Med J 1991;302:845-846

Walter SD, Kubik A, Parkin DM, Reissigova J, Adamec M and Khlat M. The natural history of lung cancer estimated from the results of a randomized trial of screening. Cancer Causes and Control 1992;3:115-123

Warren R. The debate over mass mammography in Britain. The case for. Br Med J 1988;297:969-970

WHO (World Health Organization). Cancer incidence in five continents. Volume V. IARC Scientific Publication, Lyon, 1987

WHO (World Health Organization). World Health Statistics, Geneva, 1989

Williams A. Economics of coronary artery bypass grafting. Br Med J 1985;291:326-329

Summary

Several breast cancer screening trials have been going on before any government had decided to implement nationwide screening. In 1986, the Dutch Ministry of Welfare, Health and Cultural Affairs asked our department to start investigating the expected effect of breast cancer screening on mortality and possibly morbidity, if implemented in the Netherlands. Several data from the two Dutch breast cancer screening projects, the DOM-projects in Utrecht and the Nijmegen project, was made available. The research group consisted of members from all 3 centres. Together with research on the cost, one of the first Dutch cost-effectiveness (CE-) analyses in health care was started.

The first question in breast cancer screening concerns the prevention of advanced disease and reduction in breast cancer mortality. If screening is shown to be an effective means of reducing breast cancer mortality, it is important to establish whether the decision to implement screening could be otherwise when taking into account the impact of screening on quality of life and on health care. Cost-effectiveness and cost-utility analyses are indispensable tools in the evaluation of these questions. They result in an overview of favourable and unfavourable effects of implementing screening, of the impact on health care and of the costs and possible savings. The integration of assessing the effects for public health, for the women, and for health care together with predicted cost estimates is the main asset of such analyses. The relationship between effects and costs may be additional guides for decisions on screening policy and/or further research and follow up.

If applying screening to a population, part of the detectable lesions will become screen-detected cancers whereas the number of clinically diagnosed cancers becomes smaller. The shift from diagnosing and treating relatively large clinical cancers towards earlier screen-detected stages is reflected in a decrease in breast cancer mortality, as shown in breast cancer screening trials. The concept of screening for cancer has been incorporated in mathematical or simulation models, used for the evaluation of screening in order to try to adjust for a number of complexities involved. For breast cancer, the computer simulation package MISCAN, developed at our department, has been the underlying model for part of the research presented in this thesis.

This model and concept of disease and screening has been extended in detail in order to assess the differences on assessment and treatment if implementing screening. New research was set up in those areas where important information was missing. Quantification of the impact of breast cancer screening on the favourable on the one hand and the unfavourable effects on the other hand

was initiated. A quality of life index had to be constructed in which the gain or loss in quality of life for women was expressed for all women in or referred after the breast cancer screening programme, both short and long term effects.

In the Netherlands, a 2-yearly mammographic screening programme for women aged 50-69 is predicted to detect 26% of all diagnosed breast cancers in the population. The average size of the screen-detected cancers is small: after the build-up period of screening with a relatively high number of prevalent (large sized) cancers, 80% of all screen-detected cancers are smaller than 20 mm or are non-invasive. This would be 37% in the non-screening situation. The proportion of women with axillary lymph-node metastases is approximately 60% of the proportion in clinically diagnosed cases, given the same tumour size. The earlier detection will gradually make its impact on breast cancer mortality reduction. After 10 years, the number of women who will die from breast cancer may have fallen from 3,675 to 3,325 (in that year). A maximum annual reduction in breast cancer mortality of 16% (=700 women) in the total female population is attainable from the year 2015 onwards. On average, each year of screening will prevent 630 breast cancer deaths in the Netherlands.

Unfavourable side-effects of screening are, however, inevitable. Women who are being detected at screening of having breast cancer will know and experience their disease through a longer period, than would have been without screening. Screening may induce anxiety and result in referrals, diagnostic surgery and follow up for women who appear not to have breast cancer (false positives). The most important unfavourable effect in terms of loss in quality of life is, the large increase in women-years in follow up. Per every million screens, approximately 4,500 women will be diagnosed as having breast cancer on average 4 years earlier (lead time) and more than 1,000 patients will experience a longer disease-free interval (of 15 years), with the knowledge of having had breast cancer.

On the other hand, there are the favourable side-effects of screening, some reflected by primary treatment changes. However, the main expected favourable effect, on morbidity, is the prevention of the occurrence of advanced disease, due to screening. When computing quality-adjusted life-years, taking the expected changes in morbidity of the favourable and unfavourable effects into account, these have only limited impact and appear to cancel each other out. The adjustment is too small to attribute a major negative role to quality of life in the decision to undertake a large-scale breast cancer screening programme, if done with high quality and an appropriate invitation system. A less positive view on breast cancer screening, when taking side-effects into account, is not justified by facts.

If the United Kingdom policy, of offering screening to women aged 50-64 at 3-yearly intervals, would be implemented in the Netherlands, the expected breast cancer mortality reduction would be 11% per year. The smaller age range invited and the fact that women will have less screens with a 3-year interval (given the same time period) is responsible for this difference. In terms of cost-

effectiveness, both policies are relatively cost-effective: the Dutch policy results in a CE-ratio of 3,825 US dollars per life-year gained, and the United Kingdom policy applied to the Dutch situation of 3,235 US dollars (5% discount rate both effects and costs). In fact, screening policies of even less screens in the age group 50-69 are slightly more cost-effective, but there is a trade-off with the reduction in breast cancer mortality to be achieved. However, decreasing the screening interval to 1 year leads to much less additional effects on breast cancer mortality than it results in additional cost to be invested, and is not to be preferred.

At present, nationwide screening for women below 50 is not to be recommended. Controversy on the effectiveness of screening in this group is strong. When using the non-significant 8% breast cancer mortality reduction obtained by screening women under age 50 in Kopparberg/Östergötland in the policy of inviting women aged 40-69 at 2-yearly intervals, it still leads to a very unfavourable balance in terms of marginal cost-effectiveness and cost-utility. Further research is needed. Screening women over 70 is an option, but quality of life plays a stronger role. In this age-group the incidence and mortality rates for breast cancer and the sensitivity of mammography are relatively high. And although the attendance rate for screening and women's life-expectancy are decreasing, it appears that both the number of prevented breast cancer deaths and the costs increase by approximately 15% when including 2 additional screens up to the age of 75. However, increasing the upper age limit even further will lead to a situation where unfavourable side-effects are by far outweighing favourable side effects. Relatively more women will be diagnosed as having breast cancer, who will die from other causes than breast cancer shortly after the detection (without gaining life-years), and the unfavourable effect of lead-time years increases with age.

Whether the decision to implement breast cancer screening should be followed by other countries in Europe is not to be answered straight forwardly. Country-specific data on incidence, mortality, demography and screening costs and price levels in health care do have to be taken into account. The marked North-South gradient in incidence and mortality is responsible for a relatively higher effect of screening in e.g. the United Kingdom and the Netherlands compared to Spain and France. The often higher cost of the more decentralized health care (and screening) systems in the latter countries is another reason why the cost per life-year gained are higher in Spain than in the Netherlands. One uniform policy for all countries of the European Community is therefore maybe less obvious than expected.

The CE-approach is one of the important tools in evaluating and/or modifying screening programmes. Consensus may be reached on aspects that affects so many different factors, like age groups, screening intervals, that comparison of the expected effects and costs at national level are needed. Priorities are to be made compared to other health care programmes, such as cervical cancer screening. The rather complicated approach, we use at our department, is necessary to get the integration of as much as aspects as possible

and for answering as much as questions as possible. Integrating the different point estimates on mortality-reduction in different trials, integrating new insights in biology of the disease like for DCIS, calculating the impact on health care and monitoring the effects and costs in practice with the expectations made beforehand are probably hardly possible without this approach. Open questions have remained up till now. The influence of the trials concerning adjuvant systemic treatment on daily practice and the consequences for observed mortality reduction attributable to screening in the near future is one of them. Assessing the effects and costs of breast cancer screening will stay important research in the following years.

Samenvatting

Er zijn verschillende proefprojecten voorafgegaan aan de beslissingen van regeringen om bevolkingsonderzoek naar borstkanker in te voeren. In 1986 verzocht het Ministerie van WVC ons instituut onderzoek te doen naar het effect van een mogelijke invoering van landelijk bevolkingsonderzoek naar borstkanker in Nederland. Gegevens uit de Nederlandse proefprojecten in Utrecht en Nijmegen werden beschikbaar gesteld en er werd een onderzoeksgroep samen gesteld met leden uit deze 3 centra. Hiermee werd de start gemaakt met een van de eerste kosten-effectiviteitsanalyses in de gezondheidszorg in Nederland.

Allereerst is het van belang te kwantificeren in hoeverre screening leidt tot het vervroegd vast stellen van de ziekte borstkanker bij een aanzienlijke groep vrouwen zodat door vroegtijdige behandeling wordt voorkomen dat de ziekte in een vergevorderd stadium geraakt, en de sterfte aan borstkanker zal kunnen dalen. Als dit duidelijk het geval is, resteert de vraag of invoering van bevolkingsonderzoek ook gewenst is indien gekeken wordt naar de consequenties voor de vrouw en voor de gezondheidszorg. Kosten-effectiviteitsanalyses en kosten-utiliteitsanalyses zijn hier uitstekend geschikt voor: er ontstaat een overzicht van gunstige en minder gunstige effecten van de invoering van bevolkingsonderzoek, van de invloed van het programma op kwaliteit van leven en de gezondheidszorg, en van de kosten en mogelijke besparingen. De integratie van al deze facetten in dergelijke analyses maken ze zo interessant en kunnen leiden tot beslissingen over het te volgen beleid.

Bij het screenen van een bevolking op kanker zal een gedeelte van de detecteerbare afwijkingen worden ontdekt (als via het bevolkingsonderzoek vastgestelde borstkankers), terwijl het aantal tumoren dat op andere wijze aan het licht zou komen vermindert. De verschuiving van het ontdekken en behandelen van relatief grote kwaadaardige tumoren (in de kliniek) naar het ontdekken van meer kleinere tumoren via het bevolkingsonderzoek gaat gepaard met een daling in de sterfte aan borstkanker, zoals zichtbaar is in verschillende proefprojecten op het gebied van borstkankerscreening. Dit concept van screening is in verschillende mathematische of simulatiemodellen neergelegd om zodoende met een groot aantal complexe factoren rekening te kunnen houden bij de evaluatie en/of beslissing omtrent screening. Het computersimulatie-programma MISCAN, ontwikkeld op ons instituut, is gebruikt als onderliggend model voor een gedeelte van het onderzoek beschreven in dit proefschrift.

Dit model en het concept van screening is daarnaast uitvoerig uitgebreid om o.a. rekening te kunnen houden met de verschuivingen in diagnostiek en therapie als gevolg van screening. Hiervoor is nieuw onderzoek opgezet. Kwantificeren van de gunstige en ongunstige effecten werd als belangrijk doel

gezien, en hiervoor werd een kwaliteit van leven-index geconstrueerd. De effecten voor de vrouw, zowel op korte als op langere termijn, van een bevolkingsonderzoek werden hiermee op een rij gezet.

Naar verwachting zal in een 2-jaarlijks screeningsprogramma voor vrouwen van 50 tot 70 met behulp van mammografie (rontgenfoto's van een of beide borsten) 26% van alle gevallen van borstkanker in de populatie worden ontdekt (dit is het huidige goedgekeurde programma). Gemiddeld zijn deze tumoren klein: na de eerste jaren, waarin nog relatief veel grote tumoren aan het licht komen, zal 80% van de borstkankers ontdekt bij het bevolkingsonderzoek kleiner dan 20 mm in doorsnede zijn, of niet-invasief zijn. In de situatie zonder screening behoort slechts 37% van de borstkankers tot deze relatief gunstige categorie. Ook het aantal vrouwen met uitzaaiingen naar de oksel is met 40% gedaald. Geleidelijk aan zal deze verschuiving naar vroege detectie van tumoren er toe leiden dat minder vrouwen aan borstkanker zullen overlijden. Na 10 jaar is dit aantal gedaald van 3675 naar 3325 (in dat jaar). Uiteindelijk zal er vanaf het jaar 2015 een reductie in de sterfte aan borstkanker kunnen worden bereikt van 16% (=700 vrouwen) in de vrouwelijke populatie. Elk jaar van screening voorkomt gemiddeld 630 sterfgevallen aan borstkanker in Nederland.

Onvermijdelijk zijn er ook ongunstige neven-effecten. Gemiddeld wordt een via bovengenoemd bevolkingsonderzoek ontdekte borstkanker vier jaren eerder opgespoord. Per elke 1 miljoen screeningsonderzoeken zullen er ongeveer 4500 vrouwen zijn waarbij de diagnose borstkanker zo veel vroeger wordt vastgesteld. Naast het psychologisch nadelig effect dat de vrouw een aantal jaren langer weet dat ze borstkanker heeft, betekent deze extra periode ook extra nacontroles en bezoeken aan de polikliniek. Andere ongunstige neven-effecten zijn het aantal (fout-positieve) verwijzingen van vrouwen bij wie uiteindelijk geen borstkanker te constateren blijkt, en natuurlijk kan het screeningsonderzoek zelf voor sommige vrouwen beangstigend en/of gevoelig zijn. De totale toename in het aantal jaren dat vrouwen langer met (de notie van) borstkanker leven is het belangrijkste ongunstige neven-effect van het bevolkingsonderzoek.

Aan de andere kant zijn er duidelijk gunstige neven-effecten te verwachten, bijvoorbeeld in de primaire behandeling van borstkanker. Het belangrijkste is echter de preventie in morbiditeit ten gevolge van het minder frequent voorkomen van vergevorderde (en uitgezaaide) borstkankers. Uit het oogpunt van kwaliteit van leven blijken de ongunstige en gunstige neveneffecten elkaar vrijwel in evenwicht te houden. Bij een goed georganiseerd screeningsprogramma van hoge kwaliteit en met een uitnodigingssysteem (zoals in Nederland) blijkt kwaliteit van leven voor de groep vrouwen als geheel geen negatieve factor te zijn. Een pessimistische blik en/of negatieve beslissing omtrent invoering is dus niet gerechtvaardigd, ook als naar neven-effecten wordt gekeken.

Indien voor een programma zou zijn gekozen voor vrouwen van 50 tot 65 met een 3-jaarlijkse uitnodiging (zoals in het Verenigd Koninkrijk het geval is), zou de reductie in de sterfte aan borstkanker 11% zijn (in plaats van de verwachte 16% in het huidige programma). De minder brede leeftijdsgroep en

het feit dat vrouwen minder screeningsonderzoeken zouden krijgen met een 3-jaars interval zijn verantwoordelijk voor dit verschil. Beide voorstellen zouden echter kosten-effectief zijn: het Nederlandse schema leidt tot een kosten-effectiviteitsratio van 7650 gulden per gewonnen levensjaar, en het andere schema tot 6470 gulden (5% discontering van effecten en kosten). Het is zelfs zo dat programma's met langere intervallen tussen twee uitnodigingen (of onderzoeken) enigszins kosten-effectiever zijn, maar daarentegen zal dus de te bereiken reductie in de sterfte aan borstkanker minder zijn. Screeningsprogramma's met een korter interval dan 2 jaren is voor deze leeftijdsgroep niet aan te bevelen: de extra winst in het aantal voorkomen sterfgevallen is beduidend minder dan de extra kosten die er voor gemaakt moeten worden.

Op dit moment is een landelijk bevolkingsonderzoek voor vrouwen jonger dan 50 jaar niet aan te bevelen. Er is veel discussie in hoeverre screening in deze groep ook tot een reductie in de mortaliteit kan leiden. Uitgaande van een proefbevolkingsonderzoek in Zweden (waarbij een kleine maar verre van significante reductie werd gevonden), blijkt de verhouding tussen mogelijk te behalen effectiviteit (en kwaliteit van leven) en kosten extreem ongunstig te zijn. Nader onderzoek op dit gebied is echter gewenst. Screening van vrouwen van 70 jaar en ouder is wel een optie. Het aantal nieuwe gevallen van (incidentie) en de sterfte aan borstkanker is in deze leeftijdsgroep relatief hoog. Tevens is de screeningstest relatief goed in staat maligne afwijkingen te ontdekken bij deze vrouwen. Hoewel de opkomst voor het bevolkingsonderzoek relatief lager ligt, evenals de resterende levensverwachting van de vrouw, blijkt toch dat twee additionele screeningen tussen het 70ste en 75ste jaar zowel in kosten als effecten met een gelijke 15% toenemen. Op nog hogere leeftijd krijgen ongunstige neven-effecten echter de overhand: bij relatief veel vrouwen zal borstkanker worden ontdekt, terwijl de vrouw een korte periode erna zal overlijden aan andere doodsoorzaken zonder ten gevolge van de screening levensjaren te hebben gewonnen.

Het is niet zonneklaar dat deze gunstige resultaten en de invoering van bevolkingsonderzoek in Nederland en het Verenigd Koninkrijk automatisch betekent dat het voorbeeld door de verschillende andere landen in Europa gevolgd moet worden. Voor elk land dient de incidentie, mortaliteit, demografie en kosten afzonderlijk te worden bestudeerd. De duidelijke Noord-Zuid verschillen in deze getallen maken een screeningsprogramma in Nederland en het Verenigd Koninkrijk relatief effectiever dan bijvoorbeeld in Spanje of Frankrijk. Daarnaast zijn de kosten in de laatstgenoemde landen met vaak meer gedecentraliseerde screeningssystemen vaak relatief hoger zodat de kosten-effectiviteit alles bij elkaar minder gunstig kan worden. Een uniform beleid voor alle Europese landen lijkt derhalve minder voor de hand te liggen dan vantevoren verwacht mocht worden.

Tenslotte is de aanpak van het schatten van de effecten en kosten van een voorziening een belangrijk gegeven in de evaluatie en eventuele modificatie van screeningsprogramma's. De keuze voor de uit te nodigen leeftijdsgroepen en het

interval kunnen op deze manier beter worden bepaald, en de resultaten daadwerkelijk in de praktijk van het landelijk programma behaald, kunnen worden vergeleken met de eerder gedane verwachtingen. Tevens kunnen prioriteiten worden gemaakt ten opzichte van andere voorzieningen zoals screening op baarmoederhalskanker. De gedetailleerde (en daardoor soms ingewikkelde) aanpak die ons instituut met de evaluatie van screening voorhanden heeft blijkt wel nodig te zijn om met zoveel verschillende factoren rekening te kunnen houden en op zoveel vragen redelijk antwoord te kunnen geven. Nagaan of de sterftereducties in de verschillende proefprojecten, die allemaal uiteraard verschillend van opzet en kwaliteit zijn, tot een zelfde conclusie leiden voor de specifieke vraag wat de te verwachten reductie in Nederland is, nagaan of nieuwe inzichten over het beloop van niet-invasieve vormen van borstkanker tot andere resultaten leiden, een kwantitatief overzicht krijgen van de gevolgen voor de gezondheidszorg en kwaliteit van leven, en het periodiek kunnen vergelijken van de verwachtingswaarden met de praktijk lijken vrij ondoenlijk zonder deze aanpak. Natuurlijk blijven er op dit moment nog vragen over. De gevolgen van de gunstige gegevens over adjuvante chemo- en/of hormonale therapie in de praktijk, ook voor de vrouwen met borstkanker ontdekt via screening, en de verwachte interferentie met de te behalen sterfte reductie als gevolg van screening is hier slechts een van. Analyses van de effecten en kosten van bevolkingsonderzoek naar borstkanker zullen nog belangrijk blijven in de komende jaren.

Acknowledgements

Without breast cancer screening trials, without women participating and without people organizing and evaluating the programmes, there would have been no thesis. At least not this one.

Furthermore, without the initial request of the Ministry of Welfare, Health and Cultural Affairs and the financial support of the Health Insurance Executive Board to start investigating the expected effects and costs of breast cancer screening in the Netherlands, I am not sure the topic of the thesis would have been the same. Very difficult would it have been without the already existing base-line experience at the Dept. of Public Health on screening evaluation and without the knowledge and experience from the screening groups in Nijmegen and Utrecht.

Therefore, everyone knows that the "acknowledgements"-section of a dissertation is as difficult to write as it is to write several of the other chapters: having to be concise on the one hand without being inadequate on the other hand. Therefore, I would like to thank everyone with whom I have had contact in the past six years, whether it was on a scientific basis, on a informal basis, on a friendly or perhaps sometimes more hostile basis, or at home. The fact that this thesis has been written means you have all kept me on the right track.

There are some "special" acknowledgements I do have to make. Regarding my daily work, I am much obliged to Martin van Ineveld with whom I worked close together in the first 3 years. Although working on separate aspects, we learnt a lot from each other, and I don't mean about comic books, and the friendship made it work. On a more irregular basis, it was Gerrit van Oortmarssen whom I could always consult about any difficult methodological questions or if I wanted to have a manuscript returned full of red remarks and question marks. In the last 3 years, of course more and more new persons got involved, of which I have to thank Rob Boer.

Three women need to be mentioned here: Hanneke de Haes, not only for the work we did together but who always inspired me on thinking about my future. Arry de Bruyn should be thanked in capitalized letters, because in all the years I have never met such a fine, considerate and dedicated colleague to work with. I just know that she would have lowercapitalized it again...you see... Finally, I like to thank Bianca (and Spip and Enrico) for her patience, her dish washing and taking care of Enrico more than she should have had to.

I thank Paul van der Maas for giving me all opportunities, and Joop van Dongen for his support all through the years of past research. I thank all colleagues at the Department, all colleagues from the screening section, all colleagues from the National Evaluation Team, all international colleagues, all

clinicians in the breast cancer field and respondents on our questions or questionnaires who helped me a lot, all co-authors, all secretaries, and since I know I have forgotten some of you, I thank you all.

About the author

All his life, since 1961 up till now, he has lived in Den Haag. He only moved twice, once he was a half year old and then again when he started living together with Bianca. The first ten years of his life were predominated by a warm family including 3 older sisters. In the second decade he went to Thomas More College, where the relatively good figures for French an Science did not change his mind to start studying medicine at Rijks Universiteit Leiden in 1979. A busy period follows in which friends, volleyball and friends from volleyball play an important role.

In 1986, with a good an satisfying degree in his hands, he is asked to assist in a physician practice and starts teaching teenagers at school basics of anatomy/pathology. The multitude of medical doctors in that period makes it difficult to achieve the initially thought follow up in medicine, in retrospect without regret. In 1987, he gets the opportunity to introduce its rationality, sincereness and also his good theoretical approach in public health research at the Department of Public Health at the Erasmus Universiteit Rotterdam. It is the start of the project on the costs and effects of breast cancer screening, of which this thesis is the happy side-effect.

Since 1990, the second busy period has followed, as the coordinator in more and more projects in the field of "early detection of disease". Apart from screening for cancer, new research is set up in the field of child health surveillance, where the experienced approach on early detection and cost-effectiveness analysis applies well too. Again, the direct applicability of the research in relation to policy decision is an important asset to him. Both nationally, as a member of the National Evaluation Team for Breast cancer screening, and internationally, breast cancer (screening) stays the important topic. Teaching stays another, but now to both older and more educated persons than in 1986, as well as at home to his son Enrico, born in 1991.

All these years, his passion for Dinky Toys and comic (strips) has not changed much, hopefully become even more tasteful and sophisticated. He is only getting a little bit old for volleyball.

Appendices

Appendix 5.1

(1) The disease and treatment states considered relevant are derived from the subdivision used in the cost-effectiveness analysis (de Koning et al., 1991). If the illness process was described less specifically than necessary for the framework used here, data have been inserted in all relevant categories. For example, if papers were related to the palliative treatment phase in general, they are reflected in the different palliative treatment columns.

(2) Following the WHO definition of health, the described disease states were subdivided in a physical (covering symptoms and the activity level), a psychological and a social domain. Sexual dysfunction is supposed to be related to the physical, feeling less sexually attractive to the psychological and deteriorated sexual relations to the social domain. For quality adjusting survival, a global value should be inserted in the model. Preferably such value would reflect the willingness to trade off length of life against quality. If so, they are called utilities.

(3) Descriptive data are more easily translated into health state descriptions than the group means reported in most comparative studies. In descriptive papers prevalences are given in percentages.

(4) The distinction between an objective and a subjective approach to quality of life has often been stressed often. Objective quality of life indicators are those reflecting the evaluation of others than the person concerned or those measured by standard psychiatric measurement tools. Subjective indicators are those reflecting the evaluation or perception of the person concerned. The distinction objective/subjective is applicable to empirical studies only.

Appendix 5.2

Computation of Quality-Adjusted Life-Years gained

The calculations can be represented as follows. The expected number of lifeyears S^- without a screening programme is obtained by taking the sum of the lifeyears S_j^- in each of the 15 phases j, and in phase 0 before the diagnosis of breast cancer in which the quality of life is assumed to be 1.0:

$$S^- = \sum_{j=0}^{15} S_j^-$$

In case of a screening programme the expected number of lifeyears S^+ is:

$$S^+ = \sum_{j=0}^{15} S_j^+$$

If v_i is the quality of life in phase i, then the quality-adjusted lifeyears without (Q^-) and with (Q^+) a screening programme are:

$$Q^- = \sum_{j=0}^{15} v_j \, S_j^-$$

$$Q^+ = \sum_{j=0}^{15} v_j \, S_j^+$$

The gain in lifeyears is $G_L = S^+ - S^-$, and the gain in quality-adjusted lifeyears is $G_Q = Q^+ - Q^-$. Calculation of the gain in lifeyears is relatively straightforward. Calculation of the gain in quality-adjusted lifeyears G_Q is most easily done by correcting G_L for lost quality of life. In calculating the correction factor $C = G_Q - G_L$, the troublesome phase 0 disappears from the formula since $v_0 = 1.0$, and consequently:

$$C = \sum_{j=1}^{15} (1-v_j) \; (S_j^- - S_j^+)$$

Appendix 6.1

Number of hospital admissions (females) in 1983-1988 for invasive breast cancer as a main diagnosis (ICD=174) in which limited surgical procedures probably part of breast conserving therapy, and corresponding mean number of nursing days. The Netherlands

surgical procedures	1983	1984	1985	1986	1987	1988
other partial mastectomy and axillary dissection[1)]			271	650	861	1136
other partial mastectomy alone[2)]			270	341	394	497
total	477	507	541	991	1255	1633
mean number of nursing days	(14.4)	(13.2)	(12.8)	(12.5)	(11.8)	(11.7)
data coverage	97%	99%	82%	83%	83%	83%
extrapolation (total)	490[4)]	515[4)]	660	1195	1510	1965
excision-biopsy and axillary dissection[3)]			297	226	229	286
mean number of nursing days			(12.5)	(12.9)	(12.1)	(11.5)

Classification of surgery, 1977:

[1)] 391.9 + 372
[2)] 391.9
[3)] 391.1 + 372
[4)] subdivision not known

Appendices 6.2 and 7.1

Appendix 6.2:

a) Chance of breast conserving therapy for operable women with invasive breast cancer (not stage IIIb, not stage IV, not primary Tamoxifen); 1990-level (no trend)

	Size of tumour		
	< 10 mm	10-19 mm	≥ 20 mm
clinically diagnosed	0.7	0.6	0.3
screen-detected	0.8	0.7	0.5

b) Chance of postoperative radiotherapy for operable women with invasive carcinoma treated by mastectomy; 1990-level (no trend)

	Size of tumour		
	< 10 mm	10-19 mm	≥ 20 mm
clinically diagnosed	0.30	0.33	0.41
screen-detected	0.25	0.29	0.35

Appendix 7.1:

Duration of preclinical stages and lead time (from van Oortmarssen et al., 1990b)

The average duration of the preclinical stage is proportional to the ratio between the detection rate in the initial screening round and the clinical incidence rate. This ratio clearly increases with age: in Sweden, from 1.9 in the age group from 40 to 49 years to 3.5 in the age group from 60 to 69 years, and in Nijmegen, from 2.1 to 3.1 for the same age groups.

When searching for appropriate model assumptions about the relationship between the average duration and age, the MISCAN-model was tested against the age-specific detection rates and the age-specific incidence of interval cancers as observed in Nijmegen.

The model assumptions about preclinical disease and mammography were checked against the following results of the Dutch screening projects in Nijmegen and Utrecht. (1) detection rates by age and screening round; (2) incidence of interval cancers by age, screening round, and time since last screening; and (3) stage distribution of screen-detected cancers, interval cancers, and unscreened cancers.

In this chapter and in the first preliminary analyses, for Nijmegen, the results pertain to women in the birth cohorts (1910-1939) invited for the first round. This wide age range allowed for an age-specific comparison of model outcomes and screening results. For Utrecht, only results for birth cohorts from 1911 to 1925 from the city of Utrecht were used. A reasonable fit was obtained using the following assumptions and parameter values:

1 the average total length of the preclinical period ranging from 1.6 to 4.7 years (age 40 to age 70)
2 ratios 4:6:7 for the average durations of the disease in the preclinical stages < 10 mm: 10 to 19 mm: ≥ 20 mm;
3 the duration of the preclinical stages follows an exponential distribution. This distribution has been proven to give an adequate fit in the HIP model, as in other model-based analyses;
4 for the stages dCIS and preclinical < 10 mm, the sensitivity of mammography is 70%. For tumours ≥ 10 mm in diameter, the value is 95%.