Evaluation of service screening with mammography in Sweden with special regard to its impact on breast cancer mortality

by

Håkan Jonsson

Umeå 2002
Evaluation of service screening with mammography in Sweden with special regard to its impact on breast cancer mortality

AKADEMISK AVHANDLING

Som med vederbörligt tillstånd av Rektorsämbetet vid Umeå Universitet för avläggande av filosofie doktorsexamen offentligen kommer att försvaras i sal D, 9 tr, Tandläkarhögskolan, Umeå Universitet fredag den 17 maj 2002 kl 9.00

Fakultetsopponent:
Professor Steinar Tretli
Krefteregistret, Oslo

av

Håkan Jonsson

Department of Radiation Sciences, Oncology, Umeå University, Sweden

Umeå 2002
Evaluation of service screening with mammography in Sweden with special regard to its impact on breast cancer mortality

Håkan Jonsson, Department of Radiation Sciences, Oncology, Umeå University, SE-901 85 Umeå, Sweden

Female breast cancer is a common disease that is responsible for 10-20% of the total number of deaths in Sweden in the age group 35-59 years. Despite considerable effort to improve treatment, breast cancer mortality in many countries including Sweden has been constant over time. As there are no known risk factors suitable for effective primary prevention, efforts to reduce the number of breast cancer deaths have been directed towards mammography screening for early detection of the disease. National screening programmes have been launched in several European countries since the mid 1980s. The main objectives of this thesis were to evaluate the effects of the Swedish national service screening programme with mammography on the mortality from breast cancer and to develop models suitable for such evaluation.

In 1986 the National Swedish Board of Health and Welfare recommended service screening with mammography. The first counties launched screening programmes the same year and in 1997 the last county started screening. Service screening in Sweden is organised at the county level or sometimes at the level of smaller administrative units. Women aged 50-69 years were invited, although in many counties those aged 40-49 and 70-74 years were also invited. This led to differences among the counties in time for start and age limits, which made some geographical comparisons possible.

Data from the cancer register together with information on date and cause of death from the cause of death register were used. Information about screening characteristics within the counties was obtained by questionnaires sent to the screening centres.

Breast cancer mortality was only considered for cancer cases diagnosed after start of screening and for a certain age interval. Data from an earlier time period were used to adjust for possible baseline differences in breast cancer mortality. Two outcome measures were used, underlying cause of death as reported in the cause of death register and excess mortality. As a complement to crude relative risks based on the cumulative number of breast cancer deaths Poisson regression was also used.

Due to lack of individual data on invitation to screening a number of breast cancer cases were included that were diagnosed after the start of screening but before the individual invitation. This bias might have caused a dilution of the observed mortality reduction. Another bias was due to lead time because of the age limits. Adjustments were made for both these biases.

In a pilot study in northern Sweden two counties that started service screening in 1990 were compared with two counties without screening. A significant (P<0.05) breast cancer mortality reduction of 0.28% was found for women 40-74 years.

Three of the studies focused on different age groups using data from all of Sweden. In the age group 50-69 years breast cancer mortality was 20% lower in the counties that started screening early compared to those that started late. In comparing a group invited to screening with a group not invited to screening, a 12% reduction in breast cancer mortality was found for women aged 40-49 years. In women 70-74 years, an estimated 24% reduction of the excess mortality was found. The total numbers of person-years were 5.6, 3.3 and 1.8 million for the age groups 40-49, 50-69 and 70-74 years, respectively. The average follow-up time was 8-10 years.

A pilot service screening project started in Gävleborg county between 1974 and 1979. A long-term evaluation of the first 10 years of screening was performed with a follow-up of 22 years. The relative risk was estimated at 0.79 when compared to the neighbouring counties.

In conclusion, a reduction of breast cancer mortality related to service screening with mammography in Sweden was observed in all invited age groups.

The ratio between prevalence rate and expected incidence rate was studied for three randomised trials and service screening in six Swedish counties. An increasing ratio was found according to age, which is of interest regarding the much-debated question of the efficacy of mammography screening in different age groups.

Keywords: Breast cancer, mortality, mammography screening, evaluation, prevalence, incidence
"It's never too late to give up"

Ronny Eriksson

To Erika, Frida and David
### CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>6</td>
</tr>
<tr>
<td>ORIGINAL PAPERS</td>
<td>8</td>
</tr>
<tr>
<td>ABBREVIATIONS</td>
<td>9</td>
</tr>
<tr>
<td>BREAST CANCER EPIDEMIOLOGY</td>
<td>11</td>
</tr>
<tr>
<td>Incidence</td>
<td>11</td>
</tr>
<tr>
<td>Mortality</td>
<td>13</td>
</tr>
<tr>
<td>Survival</td>
<td>16</td>
</tr>
<tr>
<td>Risk factors (aetiology)</td>
<td>16</td>
</tr>
<tr>
<td>BREAST CANCER DIAGNOSIS AND TREATMENT</td>
<td>19</td>
</tr>
<tr>
<td>Symptoms and Diagnosis</td>
<td>19</td>
</tr>
<tr>
<td>Treatment</td>
<td>19</td>
</tr>
<tr>
<td>BREAST CANCER PREVENTION</td>
<td>21</td>
</tr>
<tr>
<td>Primary prevention</td>
<td>21</td>
</tr>
<tr>
<td>Secondary prevention</td>
<td>21</td>
</tr>
<tr>
<td>BREAST CANCER SCREENING</td>
<td>22</td>
</tr>
<tr>
<td>Theoretical basis of screening</td>
<td>22</td>
</tr>
<tr>
<td>Mammography</td>
<td>24</td>
</tr>
<tr>
<td>Alternative screening methods</td>
<td>27</td>
</tr>
<tr>
<td>Other diagnostic tools</td>
<td>28</td>
</tr>
<tr>
<td>EVALUATION OF MAMMOGRAPHY SCREENING</td>
<td>30</td>
</tr>
<tr>
<td>Considerations</td>
<td>30</td>
</tr>
<tr>
<td>Randomised trials</td>
<td>31</td>
</tr>
<tr>
<td>Non-randomised trial</td>
<td>35</td>
</tr>
<tr>
<td>Case-control design</td>
<td>35</td>
</tr>
<tr>
<td>Observational studies</td>
<td>36</td>
</tr>
<tr>
<td>Service screening</td>
<td>40</td>
</tr>
<tr>
<td>AIMS</td>
<td>42</td>
</tr>
<tr>
<td>MATERIAL AND METHODS (PAPER I-V)</td>
<td>43</td>
</tr>
<tr>
<td>Service screening in Sweden</td>
<td>43</td>
</tr>
<tr>
<td>Data collection</td>
<td>44</td>
</tr>
<tr>
<td>Breast cancer definition</td>
<td>45</td>
</tr>
<tr>
<td>Residence codes</td>
<td>45</td>
</tr>
<tr>
<td>Refined mortality</td>
<td>45</td>
</tr>
<tr>
<td>Cohort definitions</td>
<td>45</td>
</tr>
<tr>
<td>Mean individual follow-up</td>
<td>46</td>
</tr>
<tr>
<td>More than one breast cancer</td>
<td>47</td>
</tr>
<tr>
<td>Study and control groups</td>
<td>47</td>
</tr>
<tr>
<td>Mortality measures</td>
<td>48</td>
</tr>
<tr>
<td>Cumulative mortality and relative risk</td>
<td>49</td>
</tr>
<tr>
<td>Adjustment for reference period</td>
<td>50</td>
</tr>
<tr>
<td>Estimation of mortality (UCD and EM)</td>
<td>51</td>
</tr>
<tr>
<td>Poisson models</td>
<td>52</td>
</tr>
<tr>
<td>Inclusion bias</td>
<td>53</td>
</tr>
<tr>
<td>Lead time bias</td>
<td>54</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Comparison of variances</td>
<td>56</td>
</tr>
<tr>
<td>MATERIAL AND METHODS (PAPER VI)</td>
<td>58</td>
</tr>
<tr>
<td>RESULTS</td>
<td>60</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>65</td>
</tr>
<tr>
<td>Effect of intervention</td>
<td>65</td>
</tr>
<tr>
<td>Non-constant relative risk</td>
<td>65</td>
</tr>
<tr>
<td>Misclassification in exposure of invitation to screening</td>
<td>67</td>
</tr>
<tr>
<td>mammography outside the service screening</td>
<td>67</td>
</tr>
<tr>
<td>Misclassification in cause of death</td>
<td>67</td>
</tr>
<tr>
<td>Total mortality</td>
<td>69</td>
</tr>
<tr>
<td>Refined breast cancer mortality compared to total breast cancer mortality</td>
<td>69</td>
</tr>
<tr>
<td>Cumulative mortality</td>
<td>70</td>
</tr>
<tr>
<td>Adjustment for reference period</td>
<td>70</td>
</tr>
<tr>
<td>Power</td>
<td>71</td>
</tr>
<tr>
<td>Inclusion bias</td>
<td>74</td>
</tr>
<tr>
<td>Lead time</td>
<td>74</td>
</tr>
<tr>
<td>Comments to paper I-V</td>
<td>76</td>
</tr>
<tr>
<td>Comments to paper VI</td>
<td>76</td>
</tr>
<tr>
<td>CONCLUSIONS</td>
<td>77</td>
</tr>
<tr>
<td>POPULÄRVETENSKAPLIG SAMMANFATTNING PÅ SVENSKA (SUMMARY IN SWEDISH)</td>
<td>78</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>80</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>82</td>
</tr>
</tbody>
</table>
ABSTRACT

Female breast cancer is a common disease that is responsible for 10-20% of the total number of deaths in Sweden in the age group 35-59 years. Despite considerable effort to improve treatment, breast cancer mortality in many countries including Sweden has been constant over time. As there are no known risk factors suitable for effective primary prevention, efforts to reduce the number of breast cancer deaths have been directed towards mammography screening for early detection of the disease. National screening programmes have been launched in several European countries since the mid 1980s. The main objectives of this thesis were to evaluate the effects of the Swedish national service screening programme with mammography on the mortality from breast cancer and to develop models suitable for such evaluation.

In 1986 the National Swedish Board of Health and Welfare recommended service screening with mammography. The first counties launched screening programmes the same year and in 1997 the last county started screening. Service screening in Sweden is organised at the county level or sometimes at the level of smaller administrative units. Women aged 50-69 years were invited, although in many counties those aged 40-49 and 70-74 years were also invited. This led to differences among the counties in time for start and age limits, which made some geographical comparisons possible.

Data from the cancer register together with information on date and cause of death from the cause of death register were used. Information about screening characteristics within the counties was obtained by questionnaires sent to the screening centres.

Breast cancer mortality was only considered for cancer cases diagnosed after start of screening and for a certain age interval. Data from an earlier time period were used to adjust for possible baseline differences in breast cancer mortality. Two outcome measures were used, underlying cause of death as reported in the cause of death register and excess mortality. As a complement to crude relative risks based on the cumulative number of breast cancer deaths Poisson regression was also used.

Due to lack of individual data on invitation to screening a number of breast cancer cases were included that were diagnosed after the start of screening but before the individual invitation. This bias might have caused a dilution of the observed mortality reduction. Another bias was due to lead time because of the age limits. Adjustments were made for both these biases.

In a pilot study in northern Sweden two counties that started service screening in 1990 were compared with two counties without screening. A significant (P<0.05) breast cancer mortality reduction of 0.28% was found for women 40-74 years.

Three of the studies focused on different age groups using data from all of Sweden In the age group 50-69 years breast cancer mortality was 20% lower in the counties that started screening early compared to those that started late. In comparing a group
invited to screening with a group not invited to screening, a 12% reduction in breast cancer mortality was found for women aged 40-49 years. In women 70-74 years, an estimated 24% reduction of the excess mortality was found. The total numbers of person-years were 5.6, 3.3 and 1.8 million for the age groups 40-49, 50-69 and 70-74 years, respectively. The average follow-up time was 8-10 years.

A pilot service screening project started in Gävleborg county between 1974 and 1979. A long-term evaluation of the first 10 years of screening was performed with a follow-up of 22 years. The relative risk was estimated at 0.79 when compared to the neighbouring counties.

In conclusion, a reduction of breast cancer mortality related to service screening with mammography in Sweden was observed in all invited age groups.

The ratio between prevalence rate and expected incidence rate was studied for three randomised trials and service screening in six Swedish counties. An increasing ratio was found according to age, which is of interest regarding the much-debated question of the efficacy of mammography screening in different age groups.

Keywords: Breast cancer, mortality, mammography screening, evaluation, prevalence, incidence
ORIGINAL PAPERS

The thesis is based on the following papers:


Reprints were made with permission from the publishers
ABBREVIATIONS

BCDDP  Breast Cancer Detection and Demonstration Project
BSE   Breast self-examination
CBE   Clinical breast examination/ physical examination
CI    Confidence interval
CIS   Carcinoma in situ
CM    Cumulative mortality
CNBSS Canadian National Breast Screening Study
DCIS  Ductal carcinoma in situ
DPCP  Detectable preclinical phase
DOM   The Diagnosisch Onderzoek Mammarcancinoom Project
EM    Excess mortality
HIP   Health Insurance Plan of Greater New York
HRT   Hormone replacement therapy
LCIS  Lobular carcinoma in situ
MMST  Malmö Mammographic Screening Trial
MRI   Magnetic resonance imagining
NBHW  The National Board of Health and Welfare in Sweden
OR    Odds ratio
RCT   Randomised controlled trial
RR    Relative risk
TEDBC The UK trial of Early Detection of Breast Cancer
UCD   Underlying cause of death
UICC  International Union against Cancer
In Sweden, in 1999, about 29% of all cancer occurring in females is breast cancer (NBHW 2001) and for women 35 to 59 years old breast cancer is responsible for 10-20% of the total number of deaths (NBHW 2001). Despite considerable efforts to improve treatment, the mortality from breast cancer has been constant in many countries including Sweden. Because there are no known risk factors suitable for effective primary prevention, the attempts to reduce the number of breast cancer deaths rely on screening for early detection of the disease. An effective screening method is mammography, which has been studied and evaluated in several randomised controlled trials. Beside local routine screening programmes that have been performed for a long time, national screening programmes have been launched in several European countries since the mid 1980s. This thesis evaluates the effects of the Swedish national service screening programme with mammography on mortality from breast cancer.

BREAST CANCER EPIDEMIOLOGY

INCIDENCE

Breast cancer is the most common incident female cancer in many countries. In Europe, in 1995, breast cancer constituted 26% of the total number of new cancer cases (Bray et al 2002). The crude incidence in Sweden in 1999 was 141 per 100,000 (NBHW 2001) and has increased steadily during the last decades. As with many other cancer forms, breast cancer incidence increases with age. The increase over time is partly due to the increased proportion of older women in the population. However, even the age adjusted incidence has increased almost linear since the start of cancer registration in the late 1950s. Figure 1 shows the time trends in Sweden for age specific incidence in 10-year intervals.

The incidence also varies between countries. Data from 1988 to 1992 shows that Scandinavia had three times higher incidence than Japan (age standardised incidence, world population) (IARC 1997). In addition, within a country there can be some variation e.g. in Sweden the incidence was higher in big cities and urban areas (SOC 1995). Even within the rural areas a variation can be seen in Sweden as well as in Finland from 1973 through 1982; there was a lower incidence in northern areas (Pukkala et al 1987).
Cancers in the breast roughly can be divided histologically into three groups: invasive carcinoma, carcinoma in situ (CIS), and other malignant tumours (Table 1). These three groups can be histo-pathologically classified into a number of different subgroups. The reported incidence of CIS has increased during the last two decades probably mainly due to the increased use of clinical mammography and the successive introduction of mammography screening.

Table 1. Distribution of invasive breast cancer, carcinoma in situ, and other malignant tumours in the breast in women 40-74 years in Sweden 1998.

<table>
<thead>
<tr>
<th>Breast cancer</th>
<th>No of cases</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invasive carcinoma</td>
<td>4438</td>
<td>88.2</td>
</tr>
<tr>
<td>Carcinoma in situ</td>
<td>543</td>
<td>10.8</td>
</tr>
<tr>
<td>Other malignant tumours</td>
<td>49</td>
<td>1.0</td>
</tr>
<tr>
<td>Total</td>
<td>5030</td>
<td>100.0</td>
</tr>
</tbody>
</table>

For some cancer sites the autopsy frequency can have an impact on the observed incidence. This is not the case for breast cancer because the number of breast cancers incidentally found by autopsy is very low, 0.2% in 1999 (NBHW 2001). In average 2% of all female cancers were incidentally found by autopsy.
MORTALITY

In 1998, the breast cancer mortality in Sweden was 34.6 per 100,000 (NBHW 2001). In contrast to the incidence no obvious increase has been seen and the mortality has been rather constant over four decades (Fig 2). In the last two decades, however, there has been a small decrease in women 60-79 years. As for the incidence, there are significant differences in mortality between countries. In addition, the time trends can be different. In the UK for example, breast cancer mortality increased between 1960 and the mid-1980s, followed by a steep decrease for a decade to a level below that of 1960 (Peto et al 2000). The level before this decrease was clearly higher than in Sweden.

Breast cancer is a common cause of death in women aged 35-59 years (10-20% of the total number of deaths and 20-35% of all cancer death, Table 2, Figure 3). For women over 75 years old breast cancer is responsible for less than 3% of the total number of deaths and 16% of all cancer death.

Figure 2. Breast cancer mortality, in 10-year age groups, for women 40-49, 50-59, 60-69 and 70-79 years old at death, between 1960 and 1996 in Sweden.
Figure 3. Number of breast cancer deaths among women in proportion to all cancer deaths (triangle) and to all causes of deaths (square) in different age groups.
Table 2. Number of deaths and mortality/100,000 from breast cancer, all cancer and all causes by age group among women in Sweden 1998.

<table>
<thead>
<tr>
<th>Age at death</th>
<th>Breast cancer</th>
<th>All cancer</th>
<th>All causes</th>
<th>Breast cancer deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of deaths</td>
<td>Mortality /100,000</td>
<td>No. of deaths</td>
<td>Mortality /100,000</td>
<td>No. of deaths</td>
</tr>
<tr>
<td>0-24</td>
<td>0</td>
<td>0</td>
<td>47</td>
<td>4</td>
</tr>
<tr>
<td>25-29</td>
<td>3</td>
<td>1</td>
<td>19</td>
<td>7</td>
</tr>
<tr>
<td>30-34</td>
<td>9</td>
<td>3</td>
<td>35</td>
<td>11</td>
</tr>
<tr>
<td>35-39</td>
<td>28</td>
<td>10</td>
<td>89</td>
<td>31</td>
</tr>
<tr>
<td>40-44</td>
<td>57</td>
<td>20</td>
<td>158</td>
<td>55</td>
</tr>
<tr>
<td>45-49</td>
<td>88</td>
<td>29</td>
<td>249</td>
<td>83</td>
</tr>
<tr>
<td>50-54</td>
<td>131</td>
<td>40</td>
<td>472</td>
<td>146</td>
</tr>
<tr>
<td>55-59</td>
<td>138</td>
<td>55</td>
<td>636</td>
<td>253</td>
</tr>
<tr>
<td>60-64</td>
<td>132</td>
<td>63</td>
<td>731</td>
<td>348</td>
</tr>
<tr>
<td>65-69</td>
<td>163</td>
<td>79</td>
<td>1037</td>
<td>506</td>
</tr>
<tr>
<td>70-74</td>
<td>166</td>
<td>81</td>
<td>1303</td>
<td>634</td>
</tr>
<tr>
<td>75-79</td>
<td>177</td>
<td>89</td>
<td>1739</td>
<td>878</td>
</tr>
<tr>
<td>80-84</td>
<td>182</td>
<td>126</td>
<td>1545</td>
<td>1068</td>
</tr>
<tr>
<td>≥85</td>
<td>275</td>
<td>203</td>
<td>1827</td>
<td>1348</td>
</tr>
<tr>
<td>Total</td>
<td>1549</td>
<td>35</td>
<td>9887</td>
<td>221</td>
</tr>
</tbody>
</table>
Survival

The different trends in incidence and mortality correspond to increased breast cancer survival. This may be explained by earlier diagnosis and improved treatment. The early diagnosis can be ascribed to increased awareness among women and early detection methods such as clinical and screening mammography. Probably improved diagnostics have also led to an increase of non-lethal breast cancer. Breast cancer survival can be measured either by finding the cause of death for each individual or comparing observed and expected survival (relative survival) (Hakulinen 1982). Table 3 shows the estimated relative survival in Sweden. Approximately 20% of the patients diagnosed in the late 1980s died from the disease within five years. A steadily increase over time is evident. In addition, an excess mortality from breast cancer after more than 15 years of follow-up still was evident (Stenbeck 1995).

Risk factors (aetiology)

One easily observable factor is of geographic or demographical nature since the incidence between different parts of the world varies substantially. In addition, within countries differences can be seen between races or social classes. In 1988-92, the age standardised incidence per 100,000 was low in Asia and Africa. For example, in Japan the incidence per 100,000 was 20-30 and 30-60 in Eastern Europe as well as in Central and South America. Most countries in Western Europe had a high incidence, 60-80, Sweden 73. White non-Hispanic women in the USA had the highest incidence of breast cancer, 86-103 (IARC 1997).

Migration data studies show that Asian immigrants in second and later generation in the USA gradually have adopted the same breast cancer risk as in the USA (Ziegler et al 1993); this is a phenomenon that has been especially well studied among Japanese immigrants in Hawaii. Hence, hereditary factors do not explain the large differences between countries. Although most of the variation is suspected to be explained by lifestyle, so far no specific causal factor have been found.
The knowledge about hereditary factors has increased greatly during the last decades. Today about 10% of breast cancer cases are considered to be hereditary. Specific mutated genes (BRCA1, BRCA2, p53) can result in a hereditary disposition for breast cancer. Women with these mutations have a very high risk for breast cancer especially in early ages. For example, a woman with a mutated BRCA1 gene has greater than a 50% risk of developing breast cancer before the age of 50. These genes, however, are responsible for less than 5% of breast cancer cases overall; the percentage is higher in younger women (Adami et al 1998).

One well-known risk factor is ionising radiation such as diagnostic radiation exposure and radiotherapy. Most of the knowledge comes from epidemiological studies on atomic bomb survivors and cohorts of women exposed to ionising radiation for medical reasons (UNSCR 1994). The lowest dose at which radiation induced breast cancer has been shown is about 0.5 Gy. A dose at 1-1.5 Gy was estimated to double the risk for breast cancer (Larsson 1996). The main exposure due to radiation in the population comes from diagnostic examinations. However, the number of breast cancer cases caused by diagnostic radiation is small, about 1% (Evans et al 1986). The radiation dose at a mammography examination is very low and the estimated risk has to be based on extrapolations from larger doses. However, at mass screening with mammography, it seems that the possible risk for inducing breast cancer is low compared to the expected gain (Mattsson et al 2000). Radiation in connection with breast conserving therapy seems to give a marginal if any increased risk for contralateral breast cancer (Storm et al 1992, Boice et al 1992).

Some well-known risk factors are related to female sex hormones especially oestrogen. Endogenous endocrine factors such as early menarche and late menopause increase the risk while high parity and low age at first full-term pregnancy seems to reduce the risk (Kelsey et al 1993). Even lactation seems to be a protecting factor for premenopausal breast cancer (Newcomb et al 1994). In addition, exogenous hormones are suspected to increase breast cancer risk. Long term use of oral contraceptives have been found to be a risk factor although the results varies (Meirik et al 1986, Olsson et al 1989). Hormone replacement therapy (HRT) for women with climacteric problems also seems to increase the risk (Persson et al 1990).

Invasive breast cancer increases the risk for contralateral breast cancer (Adami et al 1985). Also, non-invasive breast cancer such as DCIS and LCIS (Larsson 1996) as well as some benign breast diseases seems to increase the risk (Vogel 1998).

Metabolic factors such as obesity are risk factors for postmenopausal breast cancer. In addition, body length has been related to breast cancer risk (de Waard 1986, Tömberg et al 1988). Dietary factors are strongly suspected to have an impact on the breast cancer incidence. Among possible dietary factors fat have been suggested to be important; fat intake correlates to breast cancer incidence and mortality in certain countries. Fat intake is also correlated to the increase in breast cancer incidence in high incidence countries. However, there are methodological problems when studying dietary factors because the variation of exposure within a population usually is small and it is difficult to measure the intake over a long time especially early in life. Studies
on fat intake have been carried out but have not led to any evidence of dietary fat as a risk factor (Hunter et al 1996, Willett et al 1987). There might also be protective factors in food. One such suspected factor is soy protein that seems to correlate to breast cancer incidence. In a study it was shown that the biological response to soy protein were similar as for tamoxifen (Cassidy et al 1994).

Some other risk factors to mention are alcohol (Schatzkin et al 1987, Willett et al 1987) and smoking. The increased risk due to smoking was mainly seen among ex-smokers (Manjer 2001).

As a summary there is still a lack of evidence for etiological factors that can explain most of the breast cancer incidence.
BREAST CANCER DIAGNOSIS AND TREATMENT

SYMPTOMS AND DIAGNOSIS

The main symptoms is usually a lump in the breast that is often noticed by the patient. The diagnosis can also be an incidental finding of pre-clinical asymptomatic breast cancer at a health control. The increased use of mammography, especially in screening, has increased the number of newly detected breast cancers without clinical signs and symptoms. The diagnosis is verified by clinical examination, mammography and fine needle aspiration (FNA) biopsy with cytologic examination – the “triple diagnostic” approach. If the diagnose still is uncertain biopsy can be necessary.

TREATMENT

In Sweden, in the late 1970s, health care programs (guidelines for treatment, diagnosis and follow-up) were introduced in each of the six health care regions in order to provide quality and similar health care for all residents of Sweden. The use of health care programs has continued successfully. Thus, the treatment of breast cancer is generally consistent. That is, no substantial differences in treatment between different parts of the country can be expected.

Until the 1970s most treatment focused on surgical removal of the primary tumour and regional lymph nodes. The development has led to an increased use of adjuvant therapy, which means treatment trying to prevent recurrence of the disease after the primary treatment. The therapy has also become more differentiated depending on the risk ascertained using by prognostic factors. Earlier most of the surgery was radical mastectomy, which means removal of breast and axillary lymph nodes. Today, due to earlier detection and hence smaller tumours, breast-conserving surgery is used extensively especially in younger women and for less advanced tumours.

The treatment can be grouped into treatment of primary tumour and regional lymph nodes, adjuvant therapy and treatment for metastatic disease. Treatment of a primary tumour is usually surgical removal of the tumour together with axillary dissection sometimes supplemented by radiation treatment against the breast, chest wall and regional lymph node regions. In especial risk groups adjuvant chemotherapy and/or endocrine treatment are used. In advanced breast cancer, at time for diagnosis or later, several treatment alternatives are considered usually with a palliative purpose. There are many clinical trials on breast cancer therapy ongoing. The worldwide co-operation Early Breast Cancer Trialists Collaborative Group (EBCTCG) in Oxford performs overview analyses on adjuvant treatment on early breast cancer. In 2000 EBCTCG had identified more than 180 trial groups and collected data from 290 trials that recruited 190,000 women (EBCTCG 2000).

Breast cancer is not a homogeneous disease. The disease ranges from indolent tumours to very aggressive cancers that, despite the best treatment, will kill the patient within a
short time. Death from breast cancer is usually caused by distant metastases in bone, lung or liver.

There are a number of prognostic factors that can be used to classify breast cancer into risk groups with lower or higher risk. The prognostic factors most used in clinical practice are tumour size, presence of axillary metastases, hormone receptors and tumour growth rate indicators (as S-phase). Generally axillary metastases and/or large tumour size are high risk prognostic factors and together with distant metastases they are a basis for stage classification according to the TNM system (UICC 1997). Also negative hormone receptors indicate higher risk and in such cases hormonal treatment is ineffective. Presence of distant metastases indicates a very high risk and thus the disease can be regarded as incurable. Other factors, although not strongly prognostic, such as menopausal status, can also be important for the choice of treatment.
BREAST CANCER PREVENTION

PRIMARY PREVENTION

Primary prevention is usually defined as reducing the risk of a disease by removing one or more known risk factors. However, most of the risk factors for breast cancer are either biologically related to fertility hormones and/or related to lifestyle. Not many of the known risk factors for breast cancer are suitable for prevention in the natural sense of avoiding or limiting the exposure of the risk factor.

Another kind of prevention is to expose the population to a protecting factor. This should be carefully considered because the exposure can be a risk factor for other diseases. One such idea of prevention is use of tamoxifen to reduce the risk for breast cancer. In several trials this has been studied. In the Breast Cancer Prevention Trial (P-1), initiated in 1992, 13,000 volunteer women with increased risk for breast cancer were randomly given a placebo or tamoxifen for five years. A considerable reduction (49%) of the risk of invasive breast cancer was seen after 69 months of follow-up (Fisher et al 1998). However, there are also negative consequences of tamoxifen intervention such as an increased risk for endometrial cancer (Vogel 1998).

A very different kind of prevention is based on genetic risk. Women with mutated genes and who are at very high risk for breast cancer have one drastic measure they can take to prevent the disease: the surgical removal of the breasts.

SECONDARY PREVENTION

The possibilities of affecting the incidence of breast cancer by primary prevention are at present very limited. Another alternative is secondary prevention i.e. early detection and treatment of the disease in order to prevent generalised disease and breast cancer death.

To detect breast cancer early different methods are used: breast self-examination (BSE), clinical breast examination (CBE), or mammography. Other alternatives are magnetic resonance imaging (MRI), light scan, ultrasound and thermography. After this first step to select suspicious cases a clinical and microscopic procedure has to follow in order to verify the malignant diagnoses.
BREAST CANCER SCREENING

THEORETICAL BASIS OF SCREENING

Screening is performed in a population by a test that is capable of detecting a cancer before it becomes clinically evident or symptomatic. At a certain time point, asymptomatic breast cancer will be detectable by the test. The cancer is said to enter the detectable pre-clinical phase (DPCP) (Fig 4). The duration of DPCP is called sojourn time while the time gained by the test is called lead time. If the test is repeated, the time between two tests is called the screening interval.

If a test is performed during the DPCP there is a possibility that the cancer will be detected. The probability that a test will detect a cancer is called sensitivity. Estimation of sensitivity (especially in the first screening round) is not straightforward because the total number of cases exceeds the expected number without screening. This is due to lead time and possible overdiagnosis (cases that never become clinical). Furthermore, clinical cancers diagnosed between two tests (interval cancers) might not all have been missed by the test. Interval cancers with short sojourn times had probably not entered the DPCP at the time for the test and were therefore not detectable by the test. Sometimes the test will be positive although there was no cancer. The specificity of the test is the probability of a negative test for women without breast cancer.

Figure 4. Natural history of breast cancer.
The goal of screening is to find cancer at an early stage and by treatment, give the patient a better chance of cure. Theoretically, there is a time point, individual for each cancer case, before which the patient can be cured. In Figure 5, some examples are given of the relationship between the cure point and the screening test (Armitage and Colton 1998). In (a), death from breast cancer, if the test is positive, can be avoided by screening because the cancer is detected before the cure point. In (b), screening cannot help the patient since the test will be too late. The cancer in case (c) will be missed by the screening because the DPCP is short. Together with (a), this is an example of length bias. Among cancers detected by the test, those with a long sojourn time will be over-represented. In (d), screening causes the woman a longer time with the knowledge of breast cancer because the cancer was possible to cure after clinical diagnosis. In (e), the case would not have had the diagnosis without screening.

To achieve effective screening, the test needs to be sensitive and have high specificity. In addition, a high acceptability in the population and thus a high attendance rate is necessary. There must also exist an DPCP of the disease, in which curable treatment is available. The harmful effects caused by the test should also be low compared to the gain of the screening. To achieve measurable effects on a population level the disease should be reasonably common.

Figure 5. Examples of screening consequences in different individuals (a detectable cancer is assumed to be detected at the test).
MAMMOGRAPHY

The typical feature of mammography is the use of low energy X-rays which makes it possible to visualise the breast parenchyma in some detail. The history of mammography goes back to 1913 when Albert Salomon, a surgeon in Berlin, used radiography on a large number of amputated breasts. He was the first to demonstrate invasive characteristics of breast cancer including calcifications and occult cancers. During the 1940s through the 1960s Egan in Philadelphia and Gershon-Cohen in Atlanta studied the correlation between the roentgenologic picture and the pathology. In 1951, Leborgne in Montevideo, Uruguay, described roentgenologic findings especially the significance of the calcifications. He also developed the technique, for example of local compression. In 1954, Riegler attempted roentgenologic screening on asymptomatic patients for breast cancer. Further development of the mammography technique was made by Gros in Strasbourg (Andersson 1996, Harris and Vogel 1997). The first randomised trial on breast cancer screening with mammography and clinical examination was launched in 1963. It was the Health Insurance Plan (HIP) in New York (Shapiro et al 1966). In the early 1970s, it became technically possible to reduce the radiation dose to a tenth.

During the 1970s several randomised trials and pilot screening projects started. Between 1986 and 1989 nation-wide service screening started in Sweden, Finland, the UK, and the Netherlands.

The mammogram projections used during the 1960s and 1970s were the cranio caudal and the medio lateral view. However, in the mid-1970s the Swedish radiologist Bengt Lundgren introduced the so-called medio lateral oblique view (Lundgren and Jakobsson 1976). This view is the most informative because the entire glandular tissue of the breast and a part of the axilla is visible. With this method it is possible to use one-view mammograms that could increase the patient flow and hence reduce the costs while the radiation dose was reduced.

Today there are two methods for mammography: traditional screen-film and digital. The second method has been developed in recent years. Some of the advantages are simplified image handling and possibilities for computerised image analysis. However, screen-film is still the dominating method.

Mammography can detect small lesions and impalpable tumours, down to a few mm. Also the possibility to detect CIS increases with mammography. The proportion of reported CIS for women aged 40-74 years in Sweden has increased from 1% in 1975 to 11% in 1998. Perhaps the detection and removal of CIS can be important for reducing mortality but there is still uncertainty on how often CIS converts to invasive cancer.

The possibility to find a breast cancer is dependent on the density of the breasts. Breast density is individual, but it also varies with age – dense breasts are more common among young women. In addition, hormone replacement therapy (HRT) can make the
breasts denser. Today the choice of one or two views can be dependent on the breast density.

When mammography screening starts in a population the reported breast cancer incidence increases especially during the first screening round. This is illustrated in the Swedish county of Västernorrland where screening started in January 1990 (Figure 6). The reason for the increase is earlier detection but there might also be cancers detected that never should had become clinical (overdiagnosis). Since tumours are detected earlier with mammography the average stage will be lower in a screened population. In addition, grade is on average lower in screening detected cancer.

Mammography has a high sensitivity and specificity (Mushlin et al 1998). Sensitivity has been estimated as the proportion screen-detected cancers out of the total cancers among the acceptors (screen detected plus interval cancer within 12 months), but since this estimate is biased, other estimates have been suggested (Day 1985). Sensitivity has been estimated according to this method in the UK Trial of Early Detection of Breast Cancer (TEDBC) to 72% and 83% for the age groups 45-54 and 55-64 years, respectively (Moss et al 1993). In the Swedish Two-County study the sensitivity was estimated to 86%, 92%, and 94% for women 40-49, 50-59, and 60-69 years, respectively (Tabár et al 1995).

Specificity can be approximated to be one minus the recall rate since the number of true positives and false negatives are relatively low. In the Swedish service screening, the average recall rate was 3.5% in 1988 and it had in 1995 decreased to 2.2% (Olsson et al 2000). This means a specificity around 97%. In 1995-96 the attendance rate was on average 81% and was in range 66-91% for the different programmes.

Age specific mean sojourn time has been estimated using data from the Swedish Two-County study using Markov-chain models. The estimates were 2.4, 3.7, and 4.2 years for the age groups 40-49, 50-59, and 60-69 years, respectively (Duffy et al 1997).

Factors that can influence the sensitivity of a screening programme are recall rate, screening intervals, the skill and experience of the radiologist, one or two view mammography, single or double reading, and the technical equipment in order to select the suspicious cases. In the further work-up of positive mammography findings, the diagnostic methods, with the exception of mammography, is palpation often combined with ultrasound and aspiration cytology, the triple diagnostics. The time between two examinations (screening interval) can vary. In some trial programmes, a one year interval was used, such as in the HIP trial (Shapiro et al 1966) while in other programmes – e.g. the national screening programme in England and Wales – three years have been used (Blanks et al 2000). In the Swedish service screening, 18-24 months have been used. In some programmes different intervals were used for young women and old women (Olsson et al 2000).
Figure 6. Age-adjusted breast cancer incidence between 1960 and 1998 for women aged 40-74 years in the county of Västernorrland.

There are some disadvantages such as radiation exposure, possible discomfort and pain, and biopsies on some of the false positives which cannot be ruled out with the triple diagnostic. Anxiety as the result of false positives (Olsson et al 1999) and the fact that a larger number of women will be diagnosed with non-lethal breast cancer, are negative aspects. The radiation dose in the breast at a mammography examination is about 1-2 mGy per view. Thus, any possible rise of breast cancer risk due to the total life-time dose from mammography screening has to be a theoretical extrapolation because the lowest dose where an increased breast cancer risk is shown is about 0.5 Gy. It can also be compared to the background radiation which gives a dose at 1 mGy per year to the whole body.

There is a distinction between routine screening where a whole population is the targeted and selective screening where special high-risk groups are invited. For example, women with a family history of breast cancer or are known to carry a mutated gene (BRCA1, BRCA2, p53) is a high-risk group. Projects where such women with a high genetic risk for breast cancer are invited to selective screening have recently started in Sweden and in other countries (Henrik Grönberg personal communication).
ALTERNATIVE SCREENING METHODS

Breast self-examination (BSE)

One of the advantages of breast self-examination (BSE) is the low cost. It can be questioned if BSE is screening or not because the intervention is not an invitation to a test but an invitation to be taught how to do the test (self-examination). So far, no studies have been able to show any effect of BSE on breast cancer mortality. In a study in the UK (TEDBC), women in two areas were invited to be taught BSE (Ellman et al 1993). The study had low compliance with attendance rates between 31% and 53%. The relative risk compared to the comparison centres included in the trial was 1.07 after 10 years follow-up. A 16-year follow-up did not show any effects either (RR=0.99) (TEDBC 1999).

In Finland, women in women’s clubs, included in the Mama breast self-examination program, were studied. The mortality of those who returned calendars of recorded practice of BSE over a two-year period was compared to the expected mortality based on Finnish national data. The result showed a significant breast cancer mortality reduction after 14 years, but since the total mortality was reduced to the same level a probable explanation is selection bias in the recruitment and in those 52% who returned calendars (Gastrin et al 1994).

In another study in Finland, breast cancer mortality among breast cancer patients applying for breast prostheses was studied. No survival advantage was associated with detection of breast cancer by means of BSE (Auvinen et al 1996). Of the breast cancer patients, 74% practised BSE but only 8% of these patients had found the breast cancer by means of BSE.

There are two ongoing randomised trials in Russia (Leningrad) and China (Shanghai). In the trial in Leningrad, 120,000 women 40-64 years of age were randomised to BSE instruction or to control (Semiglazov et al 1992). After 8 years of follow-up, 81 and 74 breast cancer deaths in the BSE group and the control group, respectively, were noted. The self-reported BSE compliance decreased from 82% in the first years to 56% in year 5.

In the trial in Shanghai, 267,000 women, born between 1928 and 1958 and employed in the textile industry, were cluster randomised by factory to be taught BSE or to be a control group. A follow-up from start in 1989-1990 through 1994 showed equal numbers of breast cancer detected in both arms (Thomas et al 1997). In this study, the attendance was high. Compliance in the instruction sessions was 84% on average. In a new follow-up no significant effect was found (David Thomas, personal communication).

Clinical breast examination (CBE)

Although several studies have been performed where both CBE and mammography were used, there are few results on CBE only and no randomised controlled trials
In a Canadian RCT, the Canadian National Breast Screening Study (CNBSS), one of the trial arms offered annual screening with CBE. The sensitivity was reported between 57% and 83% and specificity in range 88-96% (Baines et al 1989).

In some studies, the study group was invited to a combination of mammography and CBE, but the number of cancers could be compared by mode of detection. In the Health Insurance Plan (HIP) trial performed in the 1960s, mammography and CBE were combined in annual screening examinations. After 5 years of follow-up, 67% of the tumours were detected by CBE while 55% were detected by mammography (cases detected by both instruments included) (Shapiro 1977).

In the Breast Cancer Demonstration Detection Project (BCDDP) 280,000 women were screened for breast cancer with mammography and CBE annually for five years. Out of the 3,557 cancers detected in the project, 56% were detected by CBE and 89% by mammography (cases detected by both modalities included) (Baker 1982).

In the mid 1960s a population-based study of screening with CBE was performed in Sweden (Langeland 1970). In women aged 40-70 years breast cancer was detected in 2.1 per 1000 women at the first screening. The rate of benign lesions was very high (about 10 times as high) and a large number of diagnostic biopsies had to be performed. The conclusion was that screening with CBE was inefficient and would create practical problems.

In a non-randomised trial in the UK (TEDBC), women in two districts were invited to mammography and CBE. The sensitivity of CBE was 64% (Moss et al 1993) and the specificity 95% (TEDBC 1981). However, the findings of mammography and CBE were dependent because mammography examinations were biennial while the CBE were annual. Specificity has been found to be approximately equal to that for mammography while sensitivity was lower.

In a study in Japan, women 30-69 years in municipalities with high coverage (20-40%) of screening with CBE were compared to control municipalities. The reduction in the age-adjusted death rate from cancer of the breast in the high coverage-rate municipalities was significantly greater than women in the control groups (Kuroishi et al 2000).

A RCT was initiated in the Philippines 1995 where the CBE was conducted by trained paramedical workers (Parkin et al 1999). However, the study had to be closed down due to low compliance.

**OTHER DIAGNOSTIC TOOLS**

Ultrasound is an important adjunct to mammography and clinical examination in the further assessment of both palpable and impalpable breast abnormalities. However, the use of ultrasound in population screening of asymptomatic women is associated with unacceptable high rates of both false positive and false negative outcomes (Teh and Wilson 1998). Thermography seems to be unsuitable for screening purposes because
both sensitivity and specificity are low (Williams et al 1990). Magnetic Resonance Imaging (MRI) has high sensitivity but lower specificity (Bone et al 1996). The method, however, is time consuming and expensive. In a multi-centre study in the UK, the sensitivity and specificity of MRI will be compared to 2-view mammography. About 500 women below the age of 50 at high genetic risk of breast cancer will be recruited (Brown et al 2000). The method of lightscanning has been compared to mammography in asymptomatic women. Lightscanning was poor to pick up small cancers and had a low specificity (Alveryd et al 1990).
CONSIDERATIONS

The aim of secondary prevention using mammography screening is to reduce breast cancer mortality. Thus, the final measure of the effect should be breast cancer mortality. However, to observe an effect on mortality a certain time must pass until enough breast cancer cases in the control group are both diagnosed and have died from the disease. Waiting for results on breast cancer mortality screening can be evaluated by monitoring of measures such as compliance rate, detection rate, sensitivity, specificity, number of interval cancers, stage distribution, and rate of advanced cancers (Day et al 1989). The purpose of monitoring the screening using these measures is to check the necessary conditions for reducing breast cancer mortality. All these measures, except specificity, are surrogate measures for breast cancer mortality. High specificity on the other hand is a prerequisite for causing as little harm as possible. Secondarily it will also reduce the costs.

Choosing the appropriate population to be evaluated is important. Comparisons of mortality in the diseased cases only (case fatality) will be biased because the study group, due to length bias, will include an excess of slow-growing tumours and, due to lead time, tumours diagnosed earlier than in the control group.

There are certain questions to consider when evaluating mammography screening. One is to find a control group. The best design for choosing a control group comparable to the group invited to screening is the randomised controlled trial (RCT). Randomisation can be done individually but due to practical reasons also cluster randomisation has been used. For non-randomised studies, geographical (TEDBC 1999) and/or historical control groups (Blanks et al 2000) can be created. Another alternative was used in Finland where some year of birth cohorts were offered screening while the remaining birth years had a 2-4 year delayed screening start (Hakama et al 1997). Several case-control studies have been carried out but these are associated with problems such as self-selection bias. However, the problems seem to be related to the definition of exposure and not to the case-control design itself.

One important question is the definition of exposure. Exposure can be defined as the real effect of screening examination with mammography or as the effect of invitation to screening with mammography. The real efficacy of mammography examination is not measured among those invited as long as the attendance rate does not reach 100% which it never does in practice. From public health perspective, the interesting effect is that of invitation. Both these measures are of interest but if the effect of a mammography examination should be measured among attendees only, it can be biased. Such self-selection bias can arise if the breast cancer mortality among the attendees and the non-attendees differs.
Screening can only have an effect on breast cancer cases diagnosed after the start of screening. To measure the effectiveness of screening, mortality should be based on these cases only (refined mortality). When using the standard mortality rates cases diagnosed before screening will dilute the observed effectiveness. When evaluating an age specific screening effect the only interesting cases for evaluation are those diagnosed in the studied age group. However, the latter condition is not without problems due to lead time.

The control group can utilise clinical mammography or opportunistic screening, if available, which will dilute the observed effect. This will be more probable to happen if the control group was informed about the ongoing trial or if the trial recruited volunteers only (e.g. CNBSS). Because mammography screening is meant as a service to healthy women, this is a situation quite different from clinical therapy trials with informed consent.

**RANDOMISED TRIALS**

Several randomised controlled trials have been carried out. The first trial on mammography screening was the Health Insurance Plan (HIP) of Greater New York.

**Health Insurance Plan of Greater New York (HIP)**

Women 40-64 years in 23 out of the 31 medical groups in the Health Insurance Plan (HIP) of greater New York were individually randomised (matched by medical group) to screening with mammography and CBE annually for four years (study group) or to a control group who regularly receive medical care (Shapiro et al 1966). The size of the study group and the control group was 31000 women each. The attendance rate was 67% to 80% at the four invitations. After 9 years of follow-up, the breast cancer mortality was 30% lower in the invited group (Shapiro 1977). Long term results have been reported after 18 years of follow-up and showed a 23% reduced breast cancer mortality in women invited to screening (Shapiro 1997).

**Malmö Mammographic Screening Trial (MMST)**

Women born from 1908 to 1932 in Malmö were individually randomised to a study group that was invited to mammography screening and to a control group (MMST I) (Andersson et al 1988). The number of women in each group was 21,000. The screening programme started in 1976 and the screening intervals were 18 to 24 months. The control group was invited from 1992. First results on breast cancer mortality were reported in 1988. After 8.8 years of follow-up the relative risk (RR) for breast cancer death in the invited group compared to the control group was estimated at 0.96. The attendance rate was 70-74%. In the Swedish overview the trial was followed through 1989 (Nyström et al 1993) and the RR was estimated at 0.81. A second cohort (MMST II) comprising 18,000 women born 1933 to 1945 was randomised (this was started 1978) or when they turned the age of 45. This study has not been evaluated separately but together with women below 50 in MMST I
(Andersson and Janzon 1997). A RR at 0.64 was found. The average follow-up was 15.5 and 10 years in MMST I and MMST II, respectively.

The Two-County trial

In 1977, on the initiative of the National Board of Health in Sweden a randomised study in the counties Kopparberg and Östergötland was started. The counties were divided into 19 blocks. Each block was then divided into pairs (Östergötland) or triplets (Kopparberg). One of these was randomly allocated to a control group and the other(s) to a study group. The age of the women recruited was 39 years and over. The study group was offered screening with mammography every two years (women aged 40-49 years) or every 3 years (women aged 50-74 years). The compliance rate was 89%. In 1985, the first results showed a relative risk of breast cancer death at 0.69 (95% CI 0.51-0.92) (Tabár et al 1985). This was the first study showing a significant mortality reduction after screening with mammography as the only test method. It also showed a significant reduction of tumours in stage II+. The average follow-up time was 6 years.

Canadian National Breast Screening Study (CNBSS)

In this trial, the design differed for women 40-49 and 50-59 years and is therefore considered as two separate trials CNBSS I and CNBSS II, respectively. The women recruited were volunteers and signed a consent form. They were individually randomised to an intervention group or a control group. For women 40-49 year old, the intervention group received annual CBE and mammography while the control group had an initial CBE and annual follow-up. Women 50-59 years were individually randomised to undergo annual mammography and CBE or annual CBE only for four or five years. BSE was taught to all women. The first evaluation on breast cancer mortality showed a relative risks at 1.36 for CNBSS I (women 40-49 years at entry) (Miller et al 1992) and 0.97 for CNBSS II (women 50-59 years) (Miller et al 1992) after 7 years of follow-up. After 10.5 years, a rate ratio of 1.14 was found for CNBSS I (Miller et al 1997). In CNBSS II, the rate ratio was estimated at 1.02 after 13 years of follow-up (Miller et al 2000). The hypothesis tested in this trial is different from other trials, especially for women 50-59 years, because in this trial two screening methods were compared, mammography plus CBE and CBE only.

The Edinburgh trial

In one of the two screening districts in the non-randomised Trial of Early Detection of Breast Cancer (TEDBC) in the UK (Edinburgh), a randomised trial was performed. From 1978 to 1981, women aged 45-64 years were cluster randomised, with the 87 practices as clusters, to an intervention group or a control group (cohort 1) (Roberts et al 1984). The intervention group was invited to annual screening with mammography and CBE in year 1, 3, 5, and 7 and CBE only in year 2, 4, and 6. In 1982-83 and 1984-85, women aged 45-49 years old were randomised (cohort 2 and 3). A follow-up after 14 years (cohort 1) showed a relative risk at 0.87 (Alexander et al 1999). However,
there seemed to be some difference in socio-economic status (SES). After adjustment for SES, the result was 0.79 (95% CI 0.60-1.02). For women 45-49 years (all three cohorts), the RR adjusted for SES was somewhat lower (0.75).

**The Stockholm trial**

In this trial which started 1981, women 40-64 years were cluster randomised by date of birth, to an intervention group or a control group. The intervention group was invited to biennial screening with mammography. The attendance rate was 81-82% in the first rounds. The result after 11 years of follow-up showed a RR at 0.74 (95% CI 0.5-1.1) (Frisell et al 1997).

**The Göteborg trial**

Women 39-59 were randomised by cluster based on day of birth or individually to be invited to screening with 2-view mammography every 18 months or to a control group. Starting at the fifth round, the women in the control group were invited. Only results for women 39 to 49 years have been reported separately. A significant reduction in mortality from breast cancer was observed (RR=0.55) (Bjurstam et al 1997). However, the trial was included in the Swedish overview (Nyström et al 1993). The RR was 0.84 after 6 years of follow-up.

**Overview analysis of the trials**

In 1987, the Swedish Cancer Society initiated an overview of the four Swedish trials. The aim was to check the quality of follow-up by record linkage to the cancer register and the cause of death register. In addition, a blind review of cause of death by an independent endpoint committee was performed. The resulting breast cancer mortality showed a combined RR of 0.76 (95% CI 0.66-0.87) after a follow-up from 5 to 13 years (Nyström et al 1993). A recently published result after 16 years of follow-up showed a significant 21% reduction of breast cancer mortality (Nyström et al 2002).

A meta analysis was performed in 1993 at the European Society of Mastology Concensus Conference on Breast Cancer Screening (Wald et al 1994). All randomised trials (except the Canadian trials) were included. The reason was that the Canadian trials compared two screening regimens while the other trials compared a screened group with a control group without screening. The RR was estimated at 0.78 for women aged 40-74 years.

**Summary**

The Swedish RCTs have the same design in that they are population based and study the effect of mammography screening only. Trials performed outside Sweden evaluate the effect of mammography and CBE combined. Among these RCTs only the Edinburgh trial was population-based. The Canadian trials were fundamentally different from all other randomised mammography screening trials as the control group was screened (CBE) and all participants, study objects and controls, were
individualy selected and informed before inclusion and randomisation. The design was fairly similar to that used for randomised therapy studies. In Figure 7, results from the RCTs comparing women invited to mammography screening with a non-invited control group are shown.

**Figure 7. Results from RCTs comparing women invited to mammography screening with a non-invited control group.**

<table>
<thead>
<tr>
<th>Trial</th>
<th>Follow-up (y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIP</td>
<td>10</td>
</tr>
<tr>
<td>Malmö</td>
<td>12</td>
</tr>
<tr>
<td>Two-County</td>
<td>6</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>14</td>
</tr>
<tr>
<td>Stockholm</td>
<td>11</td>
</tr>
<tr>
<td>Göteborg</td>
<td>7</td>
</tr>
</tbody>
</table>

**On-going trials**

In Singapore, starting in 1994, 166,000 women aged 50-64 years were randomised to two-view mammography for one screening round (prevalence detection trial) or to a control group (Ng et al 1998). Out of the 67,656 eligible invited women only 42% responded. No results on mortality have been reported.

In Slovenia, 12,400 women 50-64 years old were randomly selected to be invited to single view mammography and CBE. The screening interval was 24 to 36 months. The women in the invited group were also instructed on BSE (Rudolf et al 1998).

Most of the trials recruited women 40-49 years old, sometimes in separate cohorts, but none of the trials were designed to answer the question if invitation from 40 years reduces breast cancer mortality compared to invitation from 50 years. There is one ongoing trial studying this question. In this national multi-centre study in the UK, that started in 1991, women aged 40-41 years old were randomised to a study group invited to annual screening with mammography (65,000 women) or to a control group with no
invitation (130,000 women). At age 50, both groups will join the national screening programme. [Moss, 1999 #787].

NON-RANDOMISED TRIAL

UK Trial of Early Detection of Breast Cancer

The Trial of Early Detection of Breast Cancer (TEDBC) was started 1979 in the UK; it was large multi-centre non-randomised trial (TEDBC 1981). The purpose was to investigate the effect of screening and education about BSE on breast cancer mortality. Eight centres participated from which two performed screening, two were invited to education on BSE, and four centres were used as controls. The screening centres used mammography and CBE in year 1, 3, 5, and 7 and CBE only in year 2, 4, and 6. In an evaluation after 16 years of follow-up the rate ratio for breast cancer mortality was estimated at 0.73 (95% CI 0.63-0.84) (TEDBC 1999) when the two centres performing screening was compared with the four control centres.

CASE-CONTROL DESIGN

Case-control studies have been used to evaluate the efficacy on breast cancer mortality especially in observational studies where no control group was defined. Cases are usually selected among the women who died from breast cancer and the controls among women who were alive or not dead from breast cancer within the same follow-up time as the cases. If exposure is defined as acceptance of screening, then the results will be an estimate of the efficacy of mammography examinations combined with possible self-selection bias. The definition of non-exposure can also have an impact on such bias.

The MMST I has been evaluated both as a cohort comparing the study group and the control group (Andersson et al 1988) and using case-control design (Gullberg et al 1991). The case-control approach compares women dead from breast cancer with living controls about taking part in the screening programme. The relative risk estimated from the cohort analysis was 0.96 while the estimated odds ratio (OR) using case-control design was 0.42. The difference was claimed as a combined effect of the effect of mammography examination and self-selection bias. Self-selection bias can also go in the other direction (see HIP below).

Exposure can also be defined as an invitation to screening. The results can be comparable to the “intention to treat” effect estimated in randomised trials. However, this exposure requires individual data on invitation to screening that is not always easy to obtain.

In addition, the HIP trial has been evaluated by a case-control approach. Women with cancer diagnosed within the first seven years who died from breast cancer were compared with age matched controls with equal or longer survival (Friedman and Dubin 1991). When exposure was measured as allocated versus not allocated to
screening the OR was 0.74 after 14 years of follow-up. The result was similar to the result of the cohort analysis where a RR= 0.77 was found after 14 years of follow-up (Shapiro et al 1982). Women who neglected invitation to mammography screening compared to those not allocated to screening showed an OR at 0.74. Thus, the self-selection bias for non-attendees seemed to go towards a lower risk for breast cancer death.

In an evaluation of the TEDBC trial, two different case-control studies were performed (Moss et al 1992). In one study (A), the risk of breast cancer death in women in a screening district was compared to the risk of those in a control district. In the second study (B), the risk of breast cancer death for screened women was compared to never screened in the screening district. Women who had died from breast cancer in the two districts diagnosed after entry of the trial up to 1986 were used as cases. Age matched controls were drawn from the pooled population of the two districts. They were alive at time of death for the cases. For study B, a new set of age matched controls were drawn from the screening district only. When measuring exposure as invited to screening or not (study A), an OR at 0.76 was found. This can be compared to the result from the cohort analysis (0.73 after 16 years of follow-up). In study B, exposure was measured as ever screened versus never screened in the screening district. An OR at 0.51 was found indicating a self-selection bias. The never screened in the screening district were also compared to the control district resulting in an OR at 1.17.

OBSERVATIONAL STUDIES

Florence

In 1970 in Florence Italy, a population-based screening programme started. Women aged 40-74 years old were invited to mammography every 2.5 years. The efficacy was estimated in a case-control study. Women who died from breast cancer were used as cases compared to a matched group of living controls. The comparison was based on whether they took part in the screening or not (Palli et al 1986). An OR at 0.53 was found. The mean interval between first invitation and death was 7 years.

The DOM project

In 1974 in Utrecht in the Netherlands, women aged 50-64 years old were invited to screening with two-view mammography and clinical examination. Only women attending in the first round were invited to the following rounds. The breast cancer mortality was evaluated in a case-control study were women who died from breast cancer were used as cases. Age-matched living women were used as controls. Only women with breast cancer after the start of the programme were considered. The resulting OR was 0.30 (Collette et al 1984). In a later follow-up, both a cohort and a case-control approach were used. The estimated RR and OR comparing screened versus not screened were similar, 0.54 and 0.52, respectively (Collette et al 1992).
Nijmegen

In 1975 in Nijmegen in the Netherlands, a screening programme was initiated. Women 35-65 years old were invited to single view mammography every two years. In round two through four, women born before 1940 were invited. The program was evaluated in 1981 in a case-control study (Verbeek et al 1984). The cases were women who died from breast cancer from 1975 through 1981 while controls were women who did not die from breast cancer. Only cases with breast cancer diagnosis after the start were considered. An OR at 0.48 was found for screened versus not screened.

The Breast Cancer Detection Demonstration Project (BCDDP)

In 1973, the Breast Cancer Detection Demonstration Project was introduced. The programme recruited volunteers in 29 centres in the United States (a total of 280,000 women). The women were examined yearly for five years with palpation, mammography, and thermography. Thermography was later discontinued because of its low sensitivity. Because of concern with radiation hazard, the use of mammography was reduced, beginning in 1977, in women younger than 50 years old. In total, 87% had two or more mammography examinations. The project was not designed to evaluate the efficacy of screening and a comparison group was not used. However, in 1988 an evaluation was made based on 55,000 white women aged 35-74 years old in the project. Breast cancer mortality was compared to an expected breast cancer mortality for women without breast cancer at the start of observation and for the same age distribution based on incidence and case fatality data (Morrison et al 1988). After 9 years follow-up the ratio was 0.80. However, since all participants were self-selected, a bias can not be excluded.
Table 4. Screening programmes (RCTs and observational programmes) evaluated on breast cancer mortality.

<table>
<thead>
<tr>
<th>Programme</th>
<th>Study</th>
<th>Country</th>
<th>Start year</th>
<th>Age group</th>
<th>No. women *1000</th>
<th>Test</th>
<th>Screening interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIP</td>
<td>RCT</td>
<td>USA</td>
<td>1963</td>
<td>40-64</td>
<td>62</td>
<td>M+C vs 0</td>
<td>12</td>
</tr>
<tr>
<td>Malmö I</td>
<td>RCT</td>
<td>Sweden</td>
<td>1976</td>
<td>45-70</td>
<td>42</td>
<td>M vs 0</td>
<td>18-24</td>
</tr>
<tr>
<td>Malmö II</td>
<td>RCT</td>
<td>Sweden</td>
<td>1978</td>
<td>43-49</td>
<td>18</td>
<td>M vs 0</td>
<td>18-24</td>
</tr>
<tr>
<td>Two-County</td>
<td>RCT</td>
<td>Sweden</td>
<td>1977</td>
<td>39-74</td>
<td>133</td>
<td>M vs 0</td>
<td>24 (39-49), 36</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>RCT</td>
<td>UK</td>
<td>1978</td>
<td>45-64</td>
<td>55</td>
<td>M+C vs 0</td>
<td>24</td>
</tr>
<tr>
<td>CNBSS I</td>
<td>RCT</td>
<td>Canada</td>
<td>1980</td>
<td>40-49</td>
<td>50</td>
<td>M+C vs IC*</td>
<td>12</td>
</tr>
<tr>
<td>CNBSS II</td>
<td>RCT</td>
<td>Canada</td>
<td>1980</td>
<td>50-59</td>
<td>39</td>
<td>M+C vs C*</td>
<td>12</td>
</tr>
<tr>
<td>Stockholm</td>
<td>RCT</td>
<td>Sweden</td>
<td>1981</td>
<td>39-65</td>
<td>60</td>
<td>M vs 0</td>
<td>24</td>
</tr>
<tr>
<td>Göteborg</td>
<td>RCT</td>
<td>Sweden</td>
<td>1982</td>
<td>39-59</td>
<td>50</td>
<td>M vs 0</td>
<td>18</td>
</tr>
<tr>
<td>TEDBC</td>
<td>NRT</td>
<td>UK</td>
<td>1979</td>
<td>45-64</td>
<td>236</td>
<td>M+C vs BSE vs 0</td>
<td>24</td>
</tr>
<tr>
<td>BCDDP</td>
<td>O</td>
<td>USA</td>
<td>1973</td>
<td>35-74</td>
<td>55</td>
<td>M+C+T*</td>
<td>12</td>
</tr>
<tr>
<td>DOM, Utrecht</td>
<td>O</td>
<td>Netherlands</td>
<td>1974</td>
<td>50-64</td>
<td>21</td>
<td>M+C</td>
<td>6-24</td>
</tr>
<tr>
<td>Nijmegen</td>
<td>O</td>
<td>Netherlands</td>
<td>1975</td>
<td>35-65</td>
<td>30</td>
<td>M</td>
<td>24</td>
</tr>
<tr>
<td>Florence</td>
<td>O</td>
<td>Italy</td>
<td>1970</td>
<td>40-74</td>
<td>25</td>
<td>M</td>
<td>30</td>
</tr>
</tbody>
</table>

RCT= Randomised controlled trial, NRT= Non-randomised trial, O= Observational study
M= Mammography, C= Clinical breast examination (CBE), IC= One initial CBE, T= Thermography, 0= Control group not invited
* Volunteers only
Table 5. Results from evaluation of several mammography screening trials/programmes on breast cancer mortality.

<table>
<thead>
<tr>
<th>Programme</th>
<th>Exposure/method</th>
<th>Study design</th>
<th>Follow-up (years)</th>
<th>RR/OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RCTs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIP (Shapiro et al 1982)*</td>
<td>I vs NI</td>
<td>Cohort</td>
<td>10</td>
<td>0.71</td>
<td>0.55-0.91</td>
</tr>
<tr>
<td>HIP (Friedman et al 1991)</td>
<td>I vs NI</td>
<td>CC</td>
<td>14</td>
<td>0.75</td>
<td>0.60-0.93</td>
</tr>
<tr>
<td>HIP (Friedman et al 1991)</td>
<td>A vs NA</td>
<td>CC</td>
<td>14</td>
<td>1.07</td>
<td>0.75-1.54</td>
</tr>
<tr>
<td>Malmö (Nyström et al 1993)</td>
<td>I vs NI</td>
<td>Cohort</td>
<td>12</td>
<td>0.81</td>
<td>0.62-1.07</td>
</tr>
<tr>
<td>Malmö (Gullberg et al 1991)</td>
<td>A vs NA</td>
<td>CC</td>
<td>-</td>
<td>0.42</td>
<td>0.22-0.78</td>
</tr>
<tr>
<td>Two-County (Tabár et al 1985)</td>
<td>I vs NI</td>
<td>Cohort</td>
<td>6</td>
<td>0.69</td>
<td>0.51-0.92</td>
</tr>
<tr>
<td>CNBSS I (Miller et al 1997)</td>
<td>I vs NI</td>
<td>Cohort</td>
<td>11</td>
<td>1.14</td>
<td>0.83-1.56</td>
</tr>
<tr>
<td>CNBSS II (Miller et al 2000)</td>
<td>I vs NI</td>
<td>Cohort</td>
<td>13</td>
<td>1.02</td>
<td>0.78-1.33</td>
</tr>
<tr>
<td>Edinburgh (Alexander et al 1999)</td>
<td>I vs NI</td>
<td>Cohort</td>
<td>14</td>
<td>0.79</td>
<td>0.60-1.02</td>
</tr>
<tr>
<td>Stockholm (Frisell et al 1997)</td>
<td>I vs NI</td>
<td>Cohort</td>
<td>11</td>
<td>0.74</td>
<td>0.5-1.1</td>
</tr>
<tr>
<td>Göteborg (Nyström et al 1993)</td>
<td>I vs NI</td>
<td>Cohort</td>
<td>7</td>
<td>0.84</td>
<td>0.51-1.39</td>
</tr>
<tr>
<td><strong>Observational studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEDBC (TEDBC 1999)</td>
<td>I vs NI</td>
<td>Cohort</td>
<td>16</td>
<td>0.73</td>
<td>0.63-0.84</td>
</tr>
<tr>
<td>BCDDP (Morrison et al 1988)*</td>
<td>I vs NI**</td>
<td>Cohort</td>
<td>9</td>
<td>0.8</td>
<td>0.72-0.87</td>
</tr>
<tr>
<td>DOM (Collette et al 1992)</td>
<td>A vs NA</td>
<td>CC</td>
<td>14</td>
<td>0.52</td>
<td>0.32-0.83</td>
</tr>
<tr>
<td>DOM (Collette et al 1992)</td>
<td>A vs NA</td>
<td>Cohort</td>
<td>14</td>
<td>0.52</td>
<td>0.34-0.79</td>
</tr>
<tr>
<td>Nijmegen (Verbeek et al 1984)</td>
<td>A vs NA</td>
<td>CC</td>
<td>7</td>
<td>0.48</td>
<td>0.23-1.00</td>
</tr>
<tr>
<td>Florence (Palli et al 1986)</td>
<td>A vs NA</td>
<td>CC</td>
<td>7</td>
<td>0.53</td>
<td>0.29-0.95</td>
</tr>
</tbody>
</table>

I = Invited to screening, NI = Not invited to screening, A= Attendees, NA= Non-attendees, CC= Case-control
* CI from Wald et al (1994)
** Expected breast cancer mortality calculated
Organised nation-wide screening programmes, so called service screening, have been implemented in Finland, Island, Luxembourg, the Netherlands, Sweden and the UK. Some of them have been evaluated.

In Finland the national screening programme covers women 50-59 years old and can be continued up to age 64. The screening interval is two years. Women born in even years were first invited and the invitation of women born in odd years was delayed 2-4 years. This facilitated an evaluation. A risk ratio of 0.76 (95% CI 0.53, 1.09) was found when comparing the first invited birth cohorts with those with delayed invitation (Hakama et al 1997).

In the UK, breast cancer mortality has fallen but the explanation cannot only be attributed to screening because the fall started before the effect of the nation-wide screening programme could be expected. In England and Wales, the NHS breast screening programme started in 1988 and the first round was completed in 1995. The programme invited women 50 to 64 years old every three years. In an evaluation, breast cancer mortality rates for women 50-79 years old were predicted (based on the years 1971-1989) by an age cohort model. The predicted values were compared with the observed mortality. The effect of screening was restricted only to age groups that could be affected. The other age groups were assumed to be affected by treatment effects. A mortality decrease of 6% was found to be due to screening in 1998 (Blanks et al 2000).

The Dutch nation-wide programme started in 1989 and invited women aged 50 to 69 years. It was near full implementation in 1997 (Fracheboud et al 2002). In 1998, it was extended to include women aged 70-75 years old. In 1999, the breast cancer mortality rate was significantly lower than expected. However, there can be other explanations for the mortality decrease than mammography screening only.

The Swedish nation-wide service screening programme was introduced gradually between 1986 and 1997. Women 50-69 years old were invited but in several counties the age limits were 40-74 years. The first attempt to make an evaluation utilised the coverage of screening in each county and estimated the breast cancer mortality rates in a Poisson model (Törnberg et al 1994). Thus, both geographical and historical comparisons were made. However, the follow-up was too short to evaluate the service screening. An effect (19% breast cancer mortality reduction) was observed mainly attributable to the randomised trials. In another study the cumulative number of breast cancer deaths among women 50-78 years in 1987 through 1996 was compared to the expected number based on the breast cancer mortality rates in 1970-1986 (Sjönell and Ståhle 1999). A mortality reduction of 1% was found but the result was diluted by the cases diagnosed both before invitation to screening and in age over the upper limit for invitation. However, mortality has decreased more in the four counties which first started mammography screening than in the others (Rosén et al 2000). In an evaluation of the two counties, where the Two-County trial was performed, three time periods
were compared: before the trial started, during the trial period, and after the end of the trial when service screening took over. Only breast cancer cases diagnosed after the start of the follow-up were used (refined mortality). A 50% reduction of mortality from breast cancer was found when comparing the latest period (service screening) with the first period (no screening) (Tabár et al 2001).
AIMS

The aims of this study include the following:

- To develop models evaluating the effects of the Swedish service screening with mammography on breast cancer mortality.

- To estimate the mortality effects of the Swedish service screening. The evaluation was made from a general point of view but also with respect to specific age groups. It was also of interest to compare the results with those of the previous randomised trials.

- To use data that were easily available such as the cancer register because there is no Swedish nation-wide screening register available.
MATERIAL AND METHODS (PAPER I-V)

SERVICE SCREENING IN SWEDEN

In the late 1970s, results from the HIP study concluded that mammography combined with clinical breast examination could reduce breast cancer mortality. A 30% lower breast cancer mortality was found after 9 years of follow-up (Shapiro 1977). In 1980 the Swedish parliament recommended that all women who suspected that they had a breast cancer or felt worried about the risk of the disease could, if they wished it be referred for mammography (SCC 1992). In addition, information about BSE is provided mainly within the routine health care and company health services. The first results from the Two-County study were published in 1985 (Tabár et al 1985) and showed a 31% reduction of breast cancer mortality in the group invited to mammography screening after 7 years of follow-up. Because the results were promising, the National Board of Health and Welfare (NBHW) in Sweden in the following year recommended nation-wide screening with mammography (NBHW 1986). The recommended minimum age interval for invitation was 50-69 years old, but with an option to increase the upper limit to 74 years due to local resources and decisions. The lower limit was also optional but was recommended not to be below 40 years old. The recommended screening interval was 18-24 months. Separate screening programmes were implemented and organised on the county level. However, in some counties the screening organisation was divided into two or more screening centres with separate programmes. In 1986, the first counties started to invite women to service screening. Eleven years later the last county in the country started to invite women to mammography screening.

The organised screening in Sweden is population based. All women are sent a personal invitation to screening based on population registers. However, depending on the distances within the counties the invitation procedure can differ. In counties with small geographical areas, the women can be invited according to date of birth. In large but sparse populated areas, the invitation has to be based on residence in order to minimise the travel distances. In such areas, the mammography personnel have to move around either between different hospitals or in mobile screening trailers. Although there are about 35 screening centres in the country, all the programmes are based on the recommendations from NBHW. This means that the programmes can differ in time they start and in the upper and lower age limits. Attendance rate and recall rate also can differ, but the differences in other characteristics are small (Olsson et al 2000).

There were service screening activities in Sweden the decade preceding 1986. In the county of Gävleborg, pilot activities for service screening started 1974 and covered the whole county by 1979 (Lundgren 1979). There has also been four randomised trials carried out in Sweden in two counties (the Two-County trial in the counties Östergötland and Kopparberg (Dalarna)) (Tabár et al 1985) and three cities (Malmö, Stockholm and Gothenbourg) (Andersson et al 1988, Frisell et al 1997, Nyström et al
The trials started between 1976 and 1982. When the trial time ended, after 4-6 years, the control groups were also invited. Exceptions were the Malmö trials where the control groups were not invited until 1992. Figure 8 shows the proportion of women 50-59 years old in Sweden covered by (invited to) mammography screening programmes. The coverage of the age group 60-69 is similar except for the trials in Göteborg and Stockholm in the early 1980s with upper age limits 59 and 65 respectively.

**Figure 8. Proportion of women aged 50-59 years old covered by screening programmes in Sweden, 1970-1997.**

**DATA COLLECTION**

To receive characteristics for 30 of the 35 different screening programmes (geographical areas involved in the randomised trials were not used) questionnaires were sent to persons responsible for the different programmes. Information about time of start of invitation, invited age groups, screening intervals, changes over time, etc was asked for. In cases where there were uncertainties about the answers or about the geographical coverage of the screening programmes in counties with more than one programme, a follow-up telephone call was made to collect these data. Individual data on time, age, and residence at breast cancer diagnosis were obtained from the national cancer register. We also used time and age at death and cause of death that was provided from the Swedish Cancer Registry by record linkage to the cause of death register. An exception was paper I where similar data were obtained from the Regional Cancer Registry. To compute person years, aggregated population data from Statistics Sweden was used. Hence, the only individual data used in the cohorts was data on the breast cancer cases.
**BREAST CANCER DEFINITION**

A breast cancer case was defined as a case in the Swedish cancer register with diagnosis code 170 according to International Classification of Diseases, Revision 7 (WHO 1957) and histo-pathological code 096 according to WHO/HS/CAN/24.1 (WHO 1956). A breast cancer death (UCD) was defined as a breast cancer case reported to the cause of death register with breast cancer as the underlying cause of death according to International Classification of Diseases, Revisions 7-10.

**RESIDENCE CODES**

Residence codes have been changed over time especially for the municipalities. For counties with administrative borders intact over time and with only one screening programme the follow-up could be based on the county level. The counties with two or more different screening programmes were divided into geographical areas. Some of the counties have also been merged with other counties to larger counties in 1996 and 1997. The follow-up of these geographical areas therefore had to be based on municipality codes. For definition of the geographical areas in the Stockholm municipality which was shared between four screening programmes, residence codes on parish level were used. The residence codes were translated into current codes.

**REFINED MORTALITY**

A common outcome measure for the cause specific mortality is the number of deaths due to the actual disease divided by the person years here referred to as total breast cancer mortality. Typically, it is measured over a certain time (e.g., a calendar year) and for a certain age interval. Usually there are no restrictions on time and age for diagnosis of the disease that have caused death. In Papers I-V, we have used a cohort approach and have taken time and age at diagnosis into consideration when determining the cause specific mortality. We have used refined breast cancer mortality. It is defined as mortality where death due to breast cancer is valid only if the diagnosis was established after a certain time point and above a certain age.

**COHORT DEFINITIONS**

Besides refined mortality, restrictions on upper age and time period for diagnosis were used for a breast cancer death to be valid and hence to be included in the mortality. The studies were based on cohorts illustrated by the example in Figure 9. The figure shows a Lexis diagram where a straight line with slope equal to one represents an individual. In the example January 1987 is the starting point, which for a cohort in a study group is the screening start. For a control cohort, the starting point can be defined arbitrarily. The cohort is defined as all women in a given population in a specified age interval (45-59 in the example) at the starting point. We also define an accrual interval with its lower limit as the starting point, January 1987. The upper limit chosen in the example is the end of December 1993. We also define an age interval (50 to 59 years in the example). Only deaths due to breast cancer diagnosed in
the area (accrual area) within these intervals are counted. The exemplified cohort was followed until December 1998. Some breast cancer cases are illustrated where the plus sign denotes the diagnosis and square denotes breast cancer death. Filled squares denote counted deaths and deaths marked with empty squares are not counted.

Figure 9. An example of definition of a cohort, age limits, accrual interval, and follow-up. Breast cancer cases are designated as follows: diagnosis (+), breast cancer death valid in the refined model (filled square), breast cancer death not valid (open square).

**MEAN INDIVIDUAL FOLLOW-UP**

The follow-up time was measured from the start of invitation to screening in the study cohorts. This is not equal to the mean individual follow-up times since invitation. Generally, in randomised trials, women were randomised and then after a short time invited to mammography. Thus, mean follow-up in a randomised trial is measured from invitation to screening. However, measured in a county the last woman was not invited until the first screening round ended. This amount of time is usually at least as long as the screening interval but often it is longer. Because we have had no access to individual data on invitation, the computation of mean individual follow-up time since invitation was approximated. Let us, in the earlier example of the definition of cohorts, assume a time elapse for the first round to be 24 months. Let us also assume a mean waiting time from start of invitation to be 12 months (waiting time uniformly distributed on (0, 24) months). For a woman aged 53 years at start the mean (or
expected) follow-up was one year shorter than the follow-up time for the cohort measured from start. A woman aged 47 at start had to wait 3 years to become 50 years old and then if there was no screening for women below 50 years, the mean wait was 12 month (half the screening interval) before being invited. Thus, the average wait time was 4 years. These age specific times were then added as a weighted sum for each age group, with the number of women in the age groups as weights, to calculate an approximate individual follow-up time in the study group.

MORE THAN ONE BREAST CANCER

The fact that some women were diagnosed with two breast cancers does not present any problems when using total breast cancer mortality because only information about time and age at death and cause of death is used – that is, at most one event per woman was used. However, using excess mortality or refined mortality information about the breast cancer diagnosis is also used. A woman with two cancers results in two records and indicates two breast cancer deaths if the women died from breast cancer. When counting the number of deaths due to breast cancer, it is obvious that only one death per woman is adequate. We generally omitted the second cancer if a woman had two breast cancers within the accrual interval during the reference period or the study period.

STUDY AND CONTROL GROUPS

In Paper I, two counties where screening started in 1990 were compared with two counties that did not start until 5 or 6 years later. In addition, in Papers III (Figure 10B) and V (Figure 10D) counties that started service screening early were compared with counties where screening started later. In Paper II (Figure 10A), counties that used 40 years as the lower age limit for invitation were compared to counties that used 50 years as the lower age limit.

The study of the age group 70-74 years old (Paper IV) raises two questions: 1) what is the effect of screening on breast cancer mortality on this age group compared to no screening; 2) what is the effect of extending the upper age limit in the screening programme from 69 to 74 year olds. Because all counties in Sweden invites women up to 69 years old the last question is the most relevant. Therefore, we designed the study to address the second question. The study group in Paper IV (Figure 10C) came from geographical areas where the upper age limit was 74 years old. As control group, geographical areas with the upper age limit 69 years old were used. Both groups were selected among geographical areas where screening started from 1986 through 1990.
MORTALITY MEASURES

In this thesis two measures of breast cancer mortality have been used – the underlying cause of death as reported in the cause of death register and excess mortality.
Underlying cause of death (UCD) is determined for each individual based on death certificates written by the doctor responsible for the patient and coded by staff on Statistics Sweden according to International Classification of Diseases (ICD). From 1952, versions 6 to 10 of the ICD have been used for defined periods in the cause of death register. The underlying cause is the cause that is the main cause of the death. Several contributing causes are also reported and registered in the cause of death register.

Excess mortality (EM) is based on the number of deaths in patients with a certain disease (breast cancer in this case). From this number, the expected number of deaths is subtracted (Lenner 1990). The expected number of deaths is derived by multiplying the person-years of the breast cancer cases and the population total mortality. The computations were made for each 5-year age group and calendar year. EM has earlier been applied to the Swedish overview of the trials and compared to the UCD (Larsson et al 1996). Excess mortality has also been used for example for prediction of cancer mortality in the Nordic countries (Engeland et al 1995).

Although both alternatives measures the number of deaths due to breast cancer there are some differences between the two measures. To use UCD, information about death and about the disease history for each individual is needed. In some cases, the decision will be difficult especially in older women where death from intercurrent diseases is more common. The quality of breast cancer as the underlying cause of death in the Swedish Cause of Death Registry have been compared to findings from an expert committee going through hospital records (Nyström et al 1995). The findings showed that there was good agreement between the cause of death register and the judgement by the committee.

Using excess mortality, individual information about cause of death is not needed. However, because the only individuals (at least theoretically) that can die from breast cancer are the breast cancer cases, it is necessary to know which individuals had breast cancer and when and at what age the disease was diagnosed. This requires access to a cancer register. It is also necessary to have information on date of death (but not the cause of death) for the breast cancer cases. Furthermore, population total mortality data are needed on an aggregated level. The method is objective in difference from UCD. However, EM is only useful for aggregated measurements and not on individuals. Excess mortality measures excess death which means not only patients that are killed directly by the cancer but also indirect death such as treatment induced deaths or suicide due to breast cancer.

**Cumulative Mortality and Relative Risk**

Cumulative breast cancer mortality (CM) was used as outcome measure for comparison of the study and the control groups. It was estimated at follow-up year $k$.
by $CM(k) = \frac{\sum_{i=1}^{k} O_i}{\sum_{i=1}^{k} N_i}$ where $O_i =$ number of breast cancer deaths follow-up year $i$ and $N_i =$ number of person-years year $i$.

The relative risk (RR) was estimated by $RR = \frac{\sum_{i=1}^{k} O_i}{\sum_{i=0}^{k} O_i}$, where the first index indicates follow-up year $i$ and the second index indicates study group ($=1$) or control group ($=0$). RR was equivalent to the ratio of $CM(k)$ for the study group and the control group when both groups were followed for $k$ years.

**ADJUSTMENT FOR REFERENCE PERIOD**

The time from the start to end of follow-up described above is designated *study period*. In Papers II-V, a *reference period* was also defined before the start of screening. The purpose was to enable an adjustment for possible differences in breast cancer mortality between the study group and the control group expected in the absence of screening. Cohorts were similarly defined in the reference period as in the study period with the same accrual interval, age limits, and follow-up except for the lag in time. The reference period was defined at least as many years earlier as the length of the accrual interval. This ensured no overlapping of the two accrual intervals so the same breast cancer cases should not be found in both periods. It was expected to be an advantage to define the reference period close in time to the study period. The differences between the groups should then be similar in both periods.

To adjust for a baseline difference in breast cancer mortality between the study group and the control group the RR in the study period (SP) was divided by the RR in the reference period (RP) i.e., $RR_{adj} = \frac{RR_{SP}}{RR_{RP}}$.

$RR_{adj}$ is the RR due to screening assuming a multiplicative relationship of period, group, and screening or not on breast cancer mortality. This can be written $\lambda_{ij} = \lambda_0 p_i g_j s_{ij}$ where $\lambda_{ij} =$ breast cancer mortality in period $i$ and group $j$ and the parameters are $\lambda_0 =$ constant

$p_i =$ study or reference period

$g_j =$ study or control group
\( s_y = s \) in the study group in the study period and = 1 elsewhere.

A disadvantage with this adjustment is that the variance increases. Because the data in the two periods can be assumed independent, the variance for \( \ln(RR) \) is approximately doubled. Hence, the confidence interval for \( \ln(RR_{adj}) \) is roughly 40\% wider than for \( \ln(RR_{sp}) \).

**ESTIMATION OF MORTALITY (UCD AND EM)**

Consider a population of women in a certain age group followed over a certain time period. Let

\( O_T = \) the total number of deaths in the population

\( N = \) the total number of person-years in the population

\( X = \) the number of women in the population with breast cancer diagnosed

\( P = \) the number of person-years for the \( X \) individuals

\( O_T \) can be assumed to be; \( Po(\lambda N) \) i.e., Poisson distributed with parameter \( \lambda N \), where \( \lambda = \) total mortality.

The assumption can be motivated by many individuals independent of each other each having a small probability of death.

**Underlying cause of death (UCD)**

\( O_C = \) the number of deaths in the population with breast cancer as underlying cause of death (UCD)

\( O_C \) is usually assumed to have distribution \( Po(\lambda_C N) \) where

the breast cancer mortality \( \lambda_C \) can be estimated by \( \lambda_C = \frac{O_C}{N} \).

**Excess mortality (EM)**

Let \( O = \) the number of deaths (all causes) among the \( X \) breast cancer cases.

Assume that \( O \) is distributed as \( Po(\lambda_E N + E) \)

where \( \lambda_E = \) excess mortality and
E = the expected number of deaths among the X individuals if they did not have breast cancer (E = P \lambda).

By assuming independence between the number of deaths from breast cancer and from all other causes the distribution is motivated as above and as a sum of two independent Poisson distributed variables. E is considered as constant. Hence, the distribution is conditioned on E.

Excess mortality can be estimated by \( \hat{\lambda}_E = \frac{O - E}{N} \).

**POISSON MODELS**

For both outcome measures, UCD (Papers II-V) and EM (Papers III-V), multiplicative Poisson models were fitted as generalised linear models (McCullagh and Nelder 1983) for the refined mortality data. The annual (triennial in Papers IV and V) number of breast cancer deaths was used as the dependent variable in the case of UCD. To estimate the EM, the total number of deaths in the breast cancer cases was used as the dependent variable. The expected number of deaths (E above) was included in the link function \( \ln(\frac{O}{E} - E) \). The logarithm of person years in the cohorts during the follow-up was used as offset i.e., \( E(O) = \lambda e^{\ln(N)} + E \).

As categorical covariates, we used follow-up year (sometimes in 3-year intervals), age at follow-up (attained age) aggregated into age groups (mostly 5-year groups), geographical area, and period (study or reference).

The estimated breast cancer mortality model can be written

\[ \lambda_{ijkl} = \lambda_0 y_i a_j g_k p_l \theta_{kl} \] where the parameters are

\[ \lambda_0 = \text{constant} \]

\[ y_i = \text{follow-up year} \]

\[ a_j = \text{age group} \]

\[ g_k = \text{geographical area} \]

\[ p_l = \text{period} \]

\[ \theta_{kl} = \theta \] for the study group in the study period and =1 elsewhere.

The restrictions \( y_i = a_j = g_k = p_l = 1 \) were used.

\( \theta \) is the RR due to invitation to screening.
Breast cancer mortality as well as any effect of screening is low in the first years. In Paper V, the first four years were excluded. In the same paper, adjustment for over-dispersion was made.

**INCLUSION BIAS**

Since we had no individual information on invitation to screening we included all women from start of the screening in each geographical area in the study group. Hence, an unknown number of women were included who received a breast cancer diagnosis after screening started but before they were invited. Because screening could not reduce breast cancer mortality among these women, an obvious dilution of the observed potential benefit of screening will occur (inclusion bias).

To estimate this bias we assumed that the cancers diagnosed before invitation were similar to cases diagnosed before starting screening. It was possible to use the earlier breast cancer incidence to estimate the number of these cases. We could also estimate the expected number of breast cancer deaths due to these cases. As an example, let us assume that the duration of the first screening round was 2 years. Figure 11 illustrates an example where the population (on the y-axis) have to wait for an invitation. This wait time can be assumed to follow a uniform distribution between 0 and 2 years. The time for diagnosis for the cases diagnosed before invitation (Figure 11, upper triangle) varies between 0 to 2 years. Incident breast cancer cases in a two-year period before screening were then given an “invitation time” simulated from a uniform (0, 2) distribution. Women diagnosed before an “invitation” were followed from the start of the period and the cumulative breast cancer mortality was calculated. This simulation was based on a large population. The expected number of deaths was calculated separately for each geographical area dependent on the length of follow-up and in proportion to the size. They were then added to a total number of expected breast cancer deaths due to cases diagnosed before invitation in the study group ($\psi$).

The relative risk, adjusted for reference period or not, can be formulated

$$RR = \frac{O}{E} = \frac{O}{O}$$

where $O$ is the number of breast cancer deaths in the study group and $E$ is the expected number of breast cancer deaths based on the control group (and on RR in the reference period). A straightforward adjustment of the relative risk for the inclusion bias is then $RR_a = \frac{O - \psi}{RR - \psi}$.

In Papers II-IV, we used the mean length of the first round and made the simulation in two periods of that length, each with 200 replicates. We then used the mean of the estimated cumulative mortality in the two simulations. We used the study population and the control population combined for the simulations. The number of deaths from
breast cancer in the cases diagnosed before invitation was about 25% of the total number of breast cancer deaths.

In Paper V, this problem was simplified because the study group included only one county and we knew the start time and end time of invitation in the first round for all the 13 geographical sub-areas of the county. In the analysis, we used each sub-area as a cohort. It was not necessary to do simulations because the time between the start and the end of the invitation was short (average=3.6 months) and the follow-up was long (22 years). We only had to estimate the cumulative mortality, which in this case was based on the whole country and proceed as above. The number of cases invited before invitation was small. Hence, also the number of breast cancer deaths was small, only 2% of the total number of breast cancer deaths. This meant that inclusion bias was insignificant in this study.

Figure 11. Time for invitation to screening during the first round in a population, along the y-axis, with a duration of two years for the first screening round.

**LEAD TIME BIAS**

The purpose of screening is early diagnosis of breast cancer. Undoubtedly, many cancers are diagnosed at an earlier state in a screened population (Tabár et al 1985); i.e., there is a lead time for the screening detected cancers. In Papers I-IV, we used age limits at diagnosis for the inclusion of cases. This might have introduced a bias. If a woman in the control group was diagnosed in the age over the upper limit and died from breast cancer during the follow-up, then the death was not counted as an event. A similar case belonging to the study group with the breast cancer detected earlier by screening could have been diagnosed below the age limit and thus the death was
counted (Figure 12). Hence, if there was a lead time for the women who died from breast cancer during the follow-up, then we have a lead time bias toward an underestimation of the effect (Papers I-IV). If women below the age limit also are invited to screening (in the study group but not in the control group), then we can have a lead time bias in the opposite direction (Papers I, III and IV). In addition, lead time bias can occur when the accrual interval is shorter than the follow-up time. This bias could underestimate the effect. This last variant of lead time bias might be less important if the remaining follow-up after the end of the accrual interval is short (Paper III) but it can be substantial if this time is long (Paper V).

In practice, we estimated the mean time to death for the deceased women who died from breast cancer (UCD). For women with two cancers in the follow-up period we excluded the second one. We then compared the mean values in the study population and the control population. If this lead time was longer among women in the screened population, then we used the difference \( x \). Let us call the age limit \( y \). We then checked how many breast cancer deaths were diagnosed in the age interval \( y-x \) to \( y \). The observed number of breast cancer deaths in the RR was then adjusted. We assumed that the lead time was the same in all cases who died from breast cancer. The adjustments used in Papers I-V are shown in Table 6.

Figure 12. Example of time to death from breast cancer (filled square) measured from diagnosis in the control group (circle) and in the study group (plus sign).
Table 6. Adjustment of RR for lead time bias.

<table>
<thead>
<tr>
<th>Paper</th>
<th>Reduction of RR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>None*</td>
</tr>
<tr>
<td>II</td>
<td>0.4</td>
</tr>
<tr>
<td>III</td>
<td>0.0</td>
</tr>
<tr>
<td>IV</td>
<td>1.7</td>
</tr>
<tr>
<td>V</td>
<td>4.2</td>
</tr>
</tbody>
</table>

* Estimated lead time for cases who died from breast cancer was longer in the control group than in the study group

**COMPARISON OF VARIANCES**

**Mortality**

Variances for the two mortality estimates are as follows:

$$ V(\lambda_E) = \frac{V(O)}{N^2} = \frac{E(O)}{N^2} = \frac{\lambda_E}{N} + \frac{E}{N^2} $$

$$ V(\lambda_C) = \frac{V(O_C)}{N^2} = \frac{E(O_C)}{N^2} = \frac{\lambda_C}{N} $$

Where $O_C$ = the number of deaths in the population with breast cancer as underlying cause of death (UCD) and $O$ = the number of deaths (all causes) among the breast cancer cases.

To compare the variances of the estimates let us assume that

$$ \lambda_C = \lambda_E, \text{ i.e. } E(O_C) = E(O) - E $$

The relative variance $\frac{V(\lambda_E)}{V(\lambda_C)}$ can then approximately be estimated by $\frac{O}{(O - E)}$.

However, this ratio will vary with age. In Table 7, the relative variances are calculated for different ages at randomisation based on data from the control group in the overview of the Swedish trials (Larsson et al 1996). The variance for EM in the age group 70-74 years old is twice as big as for UCD. For the younger women the difference was only 10%.
Table 7. Observed, expected, and excess number of deaths for EM and UCD for different age at randomisation for the control groups in the Swedish randomised trials (Larsson et al 1996). Relative variance for EM and UCD is calculated.

<table>
<thead>
<tr>
<th>Age at randomisation</th>
<th>Observed number of deaths (O)</th>
<th>Expected number of deaths (E)</th>
<th>Excess number of deaths (O-E)</th>
<th>$\frac{\hat{V}(\hat{\kappa}_E)}{\hat{V}(\hat{\kappa}_C)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>40-49</td>
<td>82</td>
<td>8.3</td>
<td>73.7</td>
<td>1.1</td>
</tr>
<tr>
<td>50-59</td>
<td>215</td>
<td>39.3</td>
<td>175.7</td>
<td>1.2</td>
</tr>
<tr>
<td>60-69</td>
<td>220</td>
<td>86.9</td>
<td>133.1</td>
<td>1.7</td>
</tr>
<tr>
<td>70-74</td>
<td>84</td>
<td>44.0</td>
<td>40.0</td>
<td>2.1</td>
</tr>
<tr>
<td>40-74</td>
<td>601</td>
<td>178.5</td>
<td>422.5</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Relative risk

It can be shown that the variance for the logarithm of RR using EM can be estimated by

$$\hat{V}(\ln(RR_{EM})) = \frac{O_s}{(O_s - E_s)^2} + \frac{O_c}{(O_c - E_c)^2}$$

and for UCD

$$\hat{V}(\ln(RR_{UCD})) = \frac{1}{O_{cs}} + \frac{1}{O_{cc}}$$

where s and c stand for study group and control group

If we assume that $O_{cs} = O_{cc}$, $O_s = O_c$ and $O_{ci} = O_i - E_i$ ($i=s$ or c) we get

$$\frac{\hat{V}(\ln(RR_{EM}))}{\hat{V}(\ln(RR_{UCD}))} = \frac{O_c}{(O_c - E_c)^2}.$$ That is, it produces the same relative difference as above.

Hence, in the age group 70-74 year old, the confidence intervals for $\ln(RR_{EM})$ are roughly 40% wider than for $\ln(RR_{UCD})$ (Table 7).
MATERIAL AND METHODS (PAPER VI)

In Paper VI, no individual data was used. Published aggregated data on the number of women invited and number of breast cancers detected at screening were used. Data from three randomised Swedish trials (invited group) and service screening in six Swedish counties were used (Table 8). Incidence and population data aggregated on the county level was retrieved from the Swedish Cancer Registry and from Statistics Sweden.

The ratio (R) between the cancer detection rate in the first screening round (prevalence) and the expected incidence was compared for different age groups. The expected incidence was estimated by the 5-year incidence in the respective county preceding the screening start. In a Poisson model, adjustments were made for age and geographical area to estimate the age specific R. Estimation was made both for invasive cancer and invasive cancer and CIS together.

The model used was

\[ \lambda_{ijk} = \lambda_0 a_i g_j p_k a_p g_p \]

where

\[ \lambda_{ijk} = \text{prevalence or incidence for a certain age and geographical area} \]

and the parameters are

\[ \lambda_0 = \text{constant} \]

\[ a_i = \text{age group } i, \ i=1,...,7 \]

\[ g_j = \text{geographical area } j, \ j=1,...,10 \]

\[ p_k = \text{incidence } (k=1), \text{ prevalence } (k=2) \]

\[ a_p = \text{interaction terms, age and prevalence} \]

\[ g_p = \text{interaction terms, geographical area and prevalence} \]

The common restrictions for the parameters have been used

\[ a_i = g_i = p_1 = a_p = g_p = 1. \]

Thus

\[ R_y = p_2 a_p g_p \]

Differences in attendance rate, and screening "quality" were adjusted by the interaction term \( g_p \). The age specific R relative to the age group 40-44 years was estimated by \( a_p \).
Table 8. Start year, age limits, and number of women invited in the first screening round for the randomised trials and service screening programmes included in Paper VI.

<table>
<thead>
<tr>
<th>Geographical area</th>
<th>Trial/ service screening</th>
<th>Year for start</th>
<th>Age limits for invitation</th>
<th>No. of women invited</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malmö</td>
<td>Trial</td>
<td>1976</td>
<td>45-69</td>
<td>21242</td>
</tr>
<tr>
<td>Kopparberg§</td>
<td>Trial</td>
<td>1977</td>
<td>40-74</td>
<td>39051</td>
</tr>
<tr>
<td>Östergötland</td>
<td>Trial</td>
<td>1978</td>
<td>40-74</td>
<td>39034</td>
</tr>
<tr>
<td>Stockholm</td>
<td>Trial</td>
<td>1981</td>
<td>40-65</td>
<td>40318</td>
</tr>
<tr>
<td>Västmanland</td>
<td>Service</td>
<td>1986</td>
<td>40-69</td>
<td>46910</td>
</tr>
<tr>
<td>Bohuslän</td>
<td>Service</td>
<td>1986</td>
<td>50-74</td>
<td>36258</td>
</tr>
<tr>
<td>Uppsala</td>
<td>Service</td>
<td>1988</td>
<td>40-74</td>
<td>48517</td>
</tr>
<tr>
<td>Norrbotten</td>
<td>Service</td>
<td>1989</td>
<td>40-74</td>
<td>52734</td>
</tr>
<tr>
<td>Västernorrland</td>
<td>Service</td>
<td>1990</td>
<td>40-74</td>
<td>55158</td>
</tr>
<tr>
<td>Västerbotten</td>
<td>Service</td>
<td>1995</td>
<td>50-69</td>
<td>27282</td>
</tr>
</tbody>
</table>

§ Currently Dalarna
RESULTS

The number of women in the population in the studied age intervals is given in Table 9. However, in Papers I-IV also women younger than the studied age intervals were included in the study and control groups because women who at start were younger than the lower age limit were included if they reached the lower age limit within the accrual interval. In contrast, in Paper V birth cohorts were used so the same women were followed during the follow-up. In addition, person-years, follow-up time, number of geographical areas and the number of breast cancer deaths from breast cancer (UCD) are given in the table. Table 10 summarises the results.

In Paper I, excess mortality was used as the only outcome measure. In 1994, the number of women aged 40-74 years old was 109,000 and 78,000 in the study population and the control population, respectively. The follow-up time was 6 years in both the study group and the control group. The total excess number of deaths was 108.4 in the study group and 107.7 in the control group. This resulted in a RR of 0.72 (95% CI 0.53-0.99). The data were also analysed for separate age groups 40-49, 50-69, and 70-74 years old at diagnosis. In the age group 50-69 years, a 33% mortality reduction was found (RR=0.67, 95% CI 0.46-0.99). In the other age groups, non-significant mortality reductions of 17% were found for both 40-49 years old and 70-74 years old.

In Paper II women 40-49 years old at diagnosis were studied. In 1988, the number of women aged 40-49 were 202,000 and 237,000 in the study population and the reference population, respectively. The follow-up time varied for the geographical areas in the study group from 3 to 10 years (average of 8.0 years). The follow-up in the control group was 10 years for all cohorts. In this paper, UCD was used as an outcome measure. We found a 9% breast cancer mortality reduction, (RR= 0.91, 95% CI 0.72-1.15) adjusted for differences in the reference period. Analyses produced the same result – RR= 0.91 (95% CI 0.72-1.15) – for annual breast cancer mortality in a Poisson model adjusted for follow-up year, geographical area and period. After adjustment for lead time bias and inclusion bias, the RR decreased to 0.88. There were three geographical areas in the study group that were followed for 10 years. A separate analysis for these areas as a study group compared to the control group resulted in a RR (unadjusted for bias) at 0.86 (95% CI 0.58-1.27).

In the target population for Paper III, the number of women 50-69 years old in 1987 were 162,000 and 99,000 in the study population and the control population, respectively. The cohorts in the study group were followed on average 10.6 years and the cohorts in the control group were all followed for 11 years. The number of breast cancer deaths (UCD) observed were 451 and 318 in the study group and the control group, respectively. Breast cancer mortality was measured using UCD and EM. Using EM, crude RR adjusted for reference period was estimated at 0.84 (95% CI 0.67-1.05). Adjustment for lead time bias and inclusion bias reduced the RR to 0.80. In a Poisson
model, we adjusted for attained age, follow-up year, geographical area, and period. The estimated RR was 0.86. Using UCD, the estimates were somewhat higher; the crude RR adjusted for reference period was 0.90 (95% CI 0.74-1.10). After adjustment for bias, the RR was 0.87. Refined breast cancer mortality in this model was compared to total breast cancer mortality for age 50-79 at death. Only 27% of the deaths in the total breast cancer mortality from 1987 through 1997 were counted in the refined model. The corresponding figures for age 50-59, 60-69, and 70-79 were 33, 43 and 12%.

In Paper IV, the number of women 70-74 years old, living in the study population and the control population in 1987 were 83,800 and 41,600, respectively. The cohorts were followed 10.1 and 9.3 years on average. The follow-up generated 1.3 and 0.6 million person-years for the study group and the control group, respectively. The number of breast cancer deaths (UCD) observed were 325 and 146 in the study group and the control group, respectively. The estimated crude RR adjusted for reference period was 0.82 (95% CI 0.57-1.19) using EM. After adjustment for bias, the RR was reduced to 0.76. The estimation using a Poisson model was 0.84. The estimates using UCD was higher. The crude RR adjusted for the reference period was 0.97 (95% CI 0.73-1.28), which after adjustment for bias was 0.94. The Poisson analysis resulted in a RR at the same level.

The aim of Paper V was to estimate the long-term effect of an early pilot screening project without restriction to a certain age group. Hence, the same target population was followed (e.g., a birth cohort). The number of women 40-65 years old in 1977 in the study population was 43,749. The corresponding number of women was 127,632 and 618,342 in control population I and control population II. The follow-up was 22 years in both the study group and the two control populations. Using EM, we estimated the crude RR during the study period at 0.84 (95% CI 0.71-1.00) when the study group was compared to control group I. After adjustment for lead time bias, the RR was estimated at 0.80. The adjustment for reference period decreased the RR to 0.82 but made the 95% CI wider (0.65-1.03). Adjustment for lead time bias led to the estimate 0.79. The crude RR adjusted for the reference period using UCD was 0.86 (95% CI 0.71-1.05). After adjustment for lead time bias, RR was 0.82. The comparisons to control group II resulted in a lower estimated mortality reduction. In the Poisson analysis, the first four years of follow-up was excluded. The results could be expected to result in a larger estimated mortality reduction, which was true for EM but not for UCD. Comparing to control group I RR for EM was 0.78 (95% CI 0.76-0.99), which after adjustment for lead time bias became 0.75. For UCD, the corresponding result was 0.88 and 0.84 after adjustment for bias.

In Paper VI, the number of screen detected invasive breast cancers and invasive cancer and carcinoma in situ (CIS) together in the first round (prevalence) was compared to the expected incidence. The comparisons were made for 5-year age groups from 40 to 74 years old. Both incidence and prevalence increased with age. Even the ratio between prevalence and incidence increased by age. In a Poisson model adjusted for geographical differences, the age-specific ratio was estimated. The increase was more pronounced for invasive cancer than for invasive cancer and CIS together. Relative to
the age group 40-44, the ratio for age 50-54 was 2.3 (95% CI 1.5-3.4) for invasive cancer only. For invasive cancer and CIS together, the corresponding ratio was 1.9 (95% CI 1.4-2.6).
Table 9. Age, population, years of follow-up, number of geographical areas, number of person-years, and number of breast cancer deaths (UCD) in the study group (SG) and the control group (CG), Papers I-V.

<table>
<thead>
<tr>
<th>Paper</th>
<th>Age</th>
<th>Year</th>
<th>Population *1000</th>
<th>No. of counties/ geographical areas</th>
<th>Person-years*1000</th>
<th>No. of bc deaths (UCD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>SG</td>
<td>CG</td>
<td>SG</td>
<td>CG</td>
</tr>
<tr>
<td>I</td>
<td>40-74</td>
<td>1994</td>
<td>109</td>
<td>78</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>III</td>
<td>50-69</td>
<td>1987</td>
<td>162</td>
<td>99</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>IV</td>
<td>70-74</td>
<td>1987</td>
<td>84</td>
<td>42</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td>V, CG I</td>
<td>40-65</td>
<td>1977</td>
<td>44</td>
<td>128</td>
<td>1§</td>
<td>4</td>
</tr>
<tr>
<td>V, CG II</td>
<td>40-65</td>
<td>1977</td>
<td>-</td>
<td>618</td>
<td>-</td>
<td>18</td>
</tr>
</tbody>
</table>

§ Divided into 13 sub-areas
Table 10. Results (Papers I-V). Age at diagnosis, outcome measure, cumulative or Poisson estimated breast cancer mortality, relative risk (RR) adjusted for reference period (RP), 95% confidence interval for RR, mean follow-up time, RR adjusted for inclusion and lead time bias.

<table>
<thead>
<tr>
<th>Paper</th>
<th>Subgroups</th>
<th>Age at diagnosis</th>
<th>Outcome measure</th>
<th>Cumulative or Poisson</th>
<th>RR adj. for RP</th>
<th>95% CI</th>
<th>Mean follow-up (years)</th>
<th>RR adjusted for bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>40-74</td>
<td>EM</td>
<td>Cum.</td>
<td>0.72</td>
<td>0.53-0.99</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age 40-49</td>
<td></td>
<td></td>
<td>0.83</td>
<td>0.46-1.50</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age 50-69</td>
<td></td>
<td></td>
<td>0.67</td>
<td>0.46-0.99</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age 70-74</td>
<td></td>
<td></td>
<td>0.83</td>
<td>0.34-1.98</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>II</td>
<td></td>
<td>40-49</td>
<td>UCD</td>
<td>Cum.</td>
<td>0.91</td>
<td>0.72-1.15</td>
<td>8.0 (5.2)*</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UCD</td>
<td>Poisson</td>
<td></td>
<td>0.91</td>
<td>0.72-1.15</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 year f-u</td>
<td></td>
<td></td>
<td>0.86</td>
<td>0.58-1.27</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>III</td>
<td></td>
<td>50-69</td>
<td>EM</td>
<td>Cum.</td>
<td>0.84</td>
<td>0.67-1.05</td>
<td>10.6 (8.7)*</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EM</td>
<td>Poisson</td>
<td></td>
<td>0.86</td>
<td>0.69-1.07</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UCD</td>
<td>Cum.</td>
<td></td>
<td>0.90</td>
<td>0.74-1.10</td>
<td></td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UCD</td>
<td>Poisson</td>
<td></td>
<td>0.91</td>
<td>0.74-1.10</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>IV</td>
<td></td>
<td>70-74</td>
<td>EM</td>
<td>Cum.</td>
<td>0.82</td>
<td>0.57-1.19</td>
<td>10.1 (8.1)*</td>
<td>9.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EM</td>
<td>Poisson</td>
<td></td>
<td>0.84</td>
<td>0.59-1.19</td>
<td></td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UCD</td>
<td>Cum.</td>
<td></td>
<td>0.97</td>
<td>0.73-1.28</td>
<td></td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UCD</td>
<td>Poisson</td>
<td></td>
<td>0.96</td>
<td>0.73-1.28</td>
<td></td>
<td>0.93</td>
</tr>
<tr>
<td>V</td>
<td>CG I</td>
<td>40-64</td>
<td>EM</td>
<td>Cum.</td>
<td>0.82</td>
<td>0.65-1.03</td>
<td>22.0</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UCD</td>
<td>Cum.</td>
<td></td>
<td>0.86</td>
<td>0.71-1.05</td>
<td></td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EM</td>
<td>Poisson**</td>
<td></td>
<td>0.78</td>
<td>0.61-0.99</td>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UCD</td>
<td>Poisson**</td>
<td></td>
<td>0.88</td>
<td>0.70-1.09</td>
<td></td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>CG II</td>
<td></td>
<td>EM</td>
<td>Cum.</td>
<td>0.90</td>
<td>0.73-1.09</td>
<td></td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UCD</td>
<td>Cum.</td>
<td></td>
<td>0.93</td>
<td>0.77-1.11</td>
<td></td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EM</td>
<td>Poisson**</td>
<td></td>
<td>0.88</td>
<td>0.70-1.09</td>
<td></td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UCD</td>
<td>Poisson**</td>
<td></td>
<td>0.94</td>
<td>0.77-1.16</td>
<td></td>
<td>0.90</td>
</tr>
</tbody>
</table>

CG I control group I (neighbouring counties), CGII control group II (Sweden)
* Individual mean follow-up time from invitation to screening
** From year 5
DISCUSSION

EFFECT OF INTERVENTION

What is usually meant by the effectiveness of mammography screening on breast cancer mortality is actually the effect of invitation to screening. Hence, since the attendance rate is below 100%, 81% in 1995-96 in Sweden (Olsson et al 2000), the real effect of reducing mortality by mammography examinations is higher (Tabár et al 2001). To evaluate this effect, those that declined to take part should be excluded. This is not without problem, as noted earlier, since to attend or not is based on a self-selection. If this selection is related to breast cancer mortality, an exclusion of the refusers will lead to a biased estimate. In this thesis the effect of invitation to screening has been considered.

NON-CONSTANT RELATIVE RISK

In this thesis, as well as in other studies, an estimated RR was usually given for the whole follow-up period. Thus, the estimate is an average over the follow-up time. The effect from screening on breast cancer mortality is delayed due to the fact that the early detected cases in the study group that should benefit from screening must be followed at least to the time where a sufficient number of corresponding cancers in the control group will be both diagnosed and will have died from breast cancer. Since the 5-year relative survival in Sweden is approximately 80% (Stenbeck 1995), a certain amount of time has to pass until we can expect to see an effect on mortality in a target population. Ten years have been proposed (Day et al 1989). There will also be a number of breast cancer cases in late stages in the first screening round, which cannot be helped with screening. These cases will dilute the observed effect. The RR will be close to one for the first years of follow-up and then decrease with follow-up time and finally converge to a "true" RR. By "true" RR is meant the effectiveness on the population we can expect after a long follow-up with continuous screening. In for example the randomised trials, the trial time was at most 8 years (except for the Malmö trials where it was longer), which might be too short for the observed "mean" RR to ever reach the "true" RR.

The estimated cumulative mortality in Paper II and III is shown in Figure 13. The study group was compared to an expected cumulative mortality based on the control group and the results in the reference period. The curves diverge after 8-9 years. Compared to randomised trials the divergence point was expected to appear later, as noted earlier, depending on the difference between the individual follow-up and the follow-up time in the cohorts. The mortality was low for the first three years and then rather constant over time.
Figure 13. Cumulative breast cancer mortality in the study group (S) compared to expected breast cancer mortality without screening (E). A) paper II (UCD) and B) paper III (EM).
MISCLASSIFICATION IN EXPOSURE OF INVITATION TO SCREENING

Our information of the invitation of women, such as age limits and start time, in each geographical area is based on information from the questionnaire. Any error due to exact definition of the areas (when a screening programme covered only a part of a county) as well as temporary or permanent changes in age intervals for invitation outside our knowledge will tend to underestimate the effect of screening.

There might be special routines for invitation of women in the age near the lower or upper limit for invitation and this also might be different in the different screening programmes. For example, women can be invited due to birth year. However, we had no knowledge about that and have used models based on the exact age for the computations. To some degree, this will cause a misclassification error that might lead to dilution of the mortality reduction.

In Paper I, the material was analysed as if the screening started in 1990 in the study group. However, in one of the two counties, Norrbotten, the screening started earlier in the city of Luleå where it started in March 1989. The population of women 40-74 in Luleå constituted about 12% of the study group. The number of women invited between March 1989 and December 1989 was 5,800, which is 11% of the invited population. Due to lead time bias, it was therefore possible to overestimate the mortality reduction. However, since we did not find any difference in lead time among the women who died from breast cancer between the study group and the control group, we did not consider this bias.

MAMMOGRAPHY OUTSIDE THE SERVICE SCREENING

The increased use of mammography as a tool for diagnosis may lead to misclassification. Because of reported positive results from randomised trials and other studies and because of the recommendation of the NBHW of nation-wide service screening, the use of mammography as a diagnostic tool has increased. Mammography probably increased the greatest in parts of the country where screening started late or in age groups not invited. This may be because women have felt something in their breasts and were referred to mammography by their physician or women who had their breasts examined by mammography as a health control. Organised local initiatives (sometimes called opportunistic screening) also contribute to an increased use of mammography. However, this phenomenon rarely is seen outside large cities (Törnberg 1994) and is mostly addressed to younger women. Both these phenomena may dilute the observed effect of screening because the control group is not unexposed to screening mammography.

MISCLASSIFICATION IN CAUSE OF DEATH

Using individual data, it can be difficult in many cases to decide whether breast cancer is an underlying cause of death. Since mortality due to intercurrent causes/diseases increases by age, the determination becomes more complicated in older women
Excess mortality, in contrast, compares observed and expected mortality among breast cancer patients without regard to cause (see Material and Methods). In the overview of the Swedish trials the two measures were compared (Larsson et al 1996). In the age group 70-74 years at randomisation, a 16% non-significant EM reduction was seen. Using UCD based on the cause of death register, the breast cancer mortality was 5% higher in the invited group, thus a difference of 21%. For the age group 50-69 the corresponding difference was 3-4% lower for EM than for UCD.

One difference when comparing a screened population with a non-screened control population is the higher detected prevalence of breast cancer in the screened population, especially in the first screening round. If, there were difficulties in deciding whether the underlying cause of death was breast cancer, the decision by the clinician would be breast cancer for some of these cases. On the other hand, if the population had been screened, the number of detected breast cancer cases would have been larger. If some of these cause of death determinations among the deceased cases were the result of incorrect determinations, it is possible that a differential misclassification bias could occur leading to overestimation of the number of breast cancer deaths in the study group.

Using data from the paper mentioned above (Larsson et al 1996), we calculated the ratio between the breast cancer mortality measured with UCD and EM for the invited group and the control group. In Figure 14 the ratio is plotted for different age groups. Breast cancer mortality was somewhat higher in the invited group but the difference was substantial in the oldest age group, 70-74 years. All results where both the EM and UCD methods were used (Papers III-V) as well as the comparison in the overview of the Swedish randomised trials (Larsson et al 1996) showed a larger mortality reduction using EM than using UCD. The largest difference was found for the oldest age group.

In conclusion, much of the evidence indicates that the misclassification bias due to the increased number of cancer cases described above is a reality, especially among older women.
TOTAL MORTALITY

If breast cancer mortality is reduced, total mortality (all causes) is also expected to be reduced. However, breast cancer as cause of death is unusual among older women but for younger women it is more common (18% and 5% for women aged 45-49 and 70-74 years, respectively, Table 2). In the overview of the Swedish RCTs where women 40-74 years old at randomisation were included 3% of the deaths were due to breast cancer (UCD) (Nyström et al 1996). Using this proportion and assuming that the RR for breast cancer mortality is 0.8 the total mortality reduction due to screening becomes 0.6% (RR= 0.97+0.03*0.8). If breast cancer mortality had been 10 % of the total, the corresponding total mortality reduction would have been 2%.

REFINED BREAST CANCER MORTALITY COMPARED TO TOTAL BREAST CANCER MORTALITY

When using total breast cancer mortality rates as outcome measures for the effectiveness of mammography screening, there are two problems to deal with. First, breast cancer cases with diagnosis before the start of screening or in age below the lower limit for invitation are included. Because screening cannot reduce mortality for such non-invited cases the result will be diluted. These dilution problems will decrease with the length of the follow-up. Second, the age interval that should be studied is unclear because the screened women will be older during follow-up. None of these problems occur when the refined mortality model is used. We found (Paper III) that a minority (27% of the breast cancer deaths, aged 50-79 years at death) of women who died from breast cancer during the follow-up was relevant for evaluation of screening;
i.e., they had a breast cancer diagnosis after screening start and within the age limits for invitation.

**CUMULATIVE MORTALITY**

The crude estimate of cumulative mortality used (described in Material and Methods) is not the only alternative. Another alternative is $CM_2(k) = \sum_{i=1}^{k} \frac{O_i}{N_i}$. The alternatives are equivalent if the person years are constant, i.e. $N_i = N$ for all $i$ (i= year of follow-up). If the mortality is assumed to be constant over follow-up time, both alternatives are unbiased and the alternative used has the minimum variance of all such estimates (Lehmann 1983). If the mortality cannot be assumed to be constant, only $CM_2(.)$ is an unbiased estimate of $\sum_{i=1}^{k} \lambda_i$, where $\lambda_i$ is the mortality. The alternative estimate $CM_2(.)$ however corresponds to the estimation of RR in the Poisson model (used in Paper II-V) where $\lambda_i$ is assumed to vary over time.

**ADJUSTMENT FOR REFERENCE PERIOD**

If the results in the study period are not adjusted for the reference period, then there might be a bias in the observed RR. Bias might occur because a retrospective material was used, which was not randomised to either a study group or a control group. Because the breast cancer incidence can differ geographically (SOC 1995), there can be differences in breast cancer mortality between geographical areas (Rosén et al 1983). We have no guarantee that the study group and the control group are comparable with respect to breast cancer mortality except for exposure to screening.

Although there is an obvious advantage to adjust for the reference period to avoid bias, there is also a disadvantage to adjust if the study and the control groups are similar in baseline breast cancer mortality. As shown above (Material and Methods), the variance for $\ln(RR)$ was doubled when adjustments for the reference period was made. Hence, for example, in paper V, using EM and control group I, the RR in the study period was 0.84 while it was 1.03 in the reference period. The adjusted RR was 0.82. Although the RR decreased, the 95% confidence interval was extended from 0.71-1.00 to 0.65-1.03. The adjustment in this case seems not to be justified: What criteria for adjustment should be used? One alternative is to adjust only when the RR in the reference period was statistically significant on the 5% level. If this criterion had been followed, no adjustments had been made in Papers III-V. Because adjustment resulted in conservative confidence intervals, no false significant results should have been obtained. We used adjustment in most of the situations.
The confidence interval is dependent on the variance of $\ln(RR)$ which for UCD is a sum of the inverses of the numbers of breast cancer deaths (EM is dependent on the total number of deaths). The power, here defined as the probability, given a true RR, to having a 95% two-sided confidence interval with upper end point below one is shown in Figure 15. It was calculated using normal approximation for $\ln(RR)$ with variance as given in the Material and Methods section. The number of breast cancer deaths in the study group during the study period was recalculated to the expected number of deaths due to the null hypothesis; no screening effect. The power depends on certain factors. The size of the study group and the control group will determine the number of breast cancer deaths. Since breast cancer mortality increases with age, another factor of importance is the age group studied.

Also two methodological factors are of importance; outcome measure for breast cancer mortality and adjustment for difference in baseline breast cancer mortality using a reference period. The variances for $\ln(RR)$ using UCD and EM have been compared in the Material and Methods section. The difference was largest for old women (twice as large for EM than for UCD) while it was almost equal for young women. Hence, if the method were chosen only based on the power, the EM model should be avoided for older women. However, there are also other factors to consider, such as misclassification of cause of death. When adjusting for differences in baseline mortality using the estimated $RR$ in the reference period, the variance for $\ln(RR)$ is doubled assuming the same breast cancer mortality in both periods (Figure 15). The power for the studies in Papers II-V assuming a true RR of 0.8 is calculated in Table 11.

Table 11. Power (as defined above) for Papers II-V, assuming a true RR of 0.8 with outcome measures EM, and UCD, with or without adjustment for the reference period (RP).

<table>
<thead>
<tr>
<th>Paper</th>
<th>Study period</th>
<th>Adj. for RP</th>
<th>Study period</th>
<th>Adj. for RP</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>0.81</td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>0.86</td>
<td>0.62</td>
<td>0.77</td>
<td>0.53</td>
</tr>
<tr>
<td>IV</td>
<td>0.61</td>
<td>0.35</td>
<td>0.41</td>
<td>0.24</td>
</tr>
<tr>
<td>V</td>
<td>0.88</td>
<td>0.62</td>
<td>0.80</td>
<td>0.53</td>
</tr>
</tbody>
</table>
Figure 15. Power for a significant breast cancer mortality reduction (the probability, given a true RR, to having a 95% two-sided confidence interval with upper end point below one). The curves show power for the study period (upper curve) and with adjustment for the reference period (lower curve). Results for Papers II-V and both outcome measures (EM and UCD).

A Paper II, UCD

B Paper III, UCD

C Paper III EM
INCLUSION BIAS

The inclusion bias in the first round, as described earlier, is a dilution bias. This means that the influence of the bias depends on the estimated effect. Hence, the adjustment becomes larger if the RR is small. In addition, the influence is larger if the follow-up time is short. For example, in Paper I we could expect a substantial adjustment for inclusion bias because the follow-up was only 6 years. However, in Paper I we did not adjust for this bias.

Assuming the time for the first round to be 2 years, the average time to diagnosis for the women who got a diagnosis after start but before invitation was 0.67 years. However, if the first round had been performed very fast, most of those should have been detected by screening. The average lead time for those cancers should then have been almost 0.67 years. Thus, the gain from screening for these cases might be less than for screening detected cancers in general. The estimated mean sojourn time, i.e., the maximum lead time, have been estimated to 2.4, 3.7, and 4.2 years for the age groups 40-49, 50-59, and 60-69, respectively (Duffy et al 1997).

There is also an inclusion bias problem due to women who are too young to be included during the first round. They were included later on when they reached the lower age limit for invitation. When they reached that age they also had to wait in average a half screening interval to be invited. However, this assumption may not always be the truth. See “Misclassification in exposure of invitation to screening” above. In Paper II, we probably have a bias towards an underestimation of the effect due to this inclusion bias for young women. We could not use the estimated adjustment for inclusion bias in the first round because we here have a follow-up that varies between zero and six years. However, the knowledge of exact invitation routines of women around 40 years was vague and we did not adjust for this bias. In Paper III, we did not allow for this bias either, but all counties in the study group except one also invited women 40-49 years to screening (81%). In Paper III, the bias probably was small.

LEAD TIME

As described above, lead time can introduce a bias due to those individuals who died from breast cancer. The estimates used for adjustment were the mean survival among those who died from breast cancer. Then the estimated lead time was compared between the study group and the control group. The estimates of this mean lead time based on data from Paper V is shown in Figure 16 for the study group and the control group. Cases were diagnosed 0-5 years and 5-10 years after screening started and the follow-up ranged from 1 to 12 years.

Lead time for those who died from breast cancer became larger with longer follow-up time. This is natural because a short follow-up time includes only cases with short survival times. During the first years a number of breast cancer cases will be screening detected in late stage in the first screening round or symptomatic cases diagnosed
outside screening. However, the cases diagnosed the first five years since start did not show any difference in lead time (Figure 16A) while the cases diagnosed 5-10 years after start showed a longer lead time in the study group after at least 8 years of follow-up (Figure 16B). Thus in cases diagnosed during the incidence screening there might be a lead time for those who died from breast cancer.

Figure 16. Estimation of mean survival time for those who died from breast cancer (UCD) in Paper V. On the x-axis is the maximum years of follow-up from diagnosis and on the y-axis the mean survival time in the study group and control group II. The estimation is based on cohorts aged 40-65 years at start. A) Diagnosis within 0-5 years from start, B) Diagnosis within 5-10 years from start.
COMMENT TO PAPER I-V

In Paper I-V there are a lot of results on the estimated breast cancer mortality reduction due to population-based mammography screening in Sweden. There are results for two outcome measures (EM, UCD) mostly adjusted for the reference period. The estimates were also adjusted for lead time bias and inclusion bias. Although most of the results were not statistically significant, all of the results showed an estimated reduction of breast cancer mortality. The follow-up time in the studied cohorts was not comparable to randomised trials since the women had to wait a certain time, from the start of invitation, to be invited. With respect to the length of the follow-up time, the results are well comparable to the results seen in the Swedish randomised trials.

COMMENT TO PAPER VI

The ratio of the screen detection rate in the prevalence round and the expected incidence was larger among old women than among young women. Thus, the ability to achieve good efficiency with mammography screening seemed to be better among old women than among young women. There are at least two possible explanations for this. 1) Mammography as diagnostic method works better on old women’s breasts since young women generally have more dense breasts. 2) It is a difference in the average growth rate for breast cancer which becomes manifest at a younger age than those occurring later in life.

Which alternative that is predominant is not clear but if the same screening routines are used for both old and young women, the result (measured as reduced mortality from breast cancer) will surely be better for old women. However, certain screening programmes already have different routines, e.g. screening intervals and number of views.
CONCLUSIONS

- The variation in start and age limits for invitation allowed the evaluation of the effect of the Swedish service screening in the age groups 40-49, 50-69 and 70-74 years.

- The evaluation was possible to do without access to individual data in the whole population. Only individual data for the breast cancer cases was necessary.

- Adjustment for lead time and inclusion bias increased the mortality reduction.

- Refined breast cancer mortality was a more appropriate outcome measure than total breast cancer mortality.

- All results, although most of them not significant, showed an estimated breast cancer mortality reduction due to service screening with mammography.

- The size of the reduction was in accordance with the results from the Swedish randomised trials.

- Using the same screening routines for both young and old women would probably give a better result among the old women.

- In all studies, where the two methods were compared, a larger mortality reduction was observed with EM than with UCD as outcome measure. A plausible explanation is misclassification of death from breast cancer due to a larger number of breast cancer cases in the study group.

- Excess mortality, as outcome measure, is a valuable complement to the traditional underlying cause of death, especially for older age groups where underlying cause of death can be biased.

Effekterna av mammografiscreening har studerats i sju vetenskapliga studier varav fyra är svenska. Dessa har visat att mammografiscreening kan sänka bröstcancerdödligheten.


I denna avhandling utvärderas hur den allmänna mammografiscreeningen i Sverige har påverkat bröstcancerdödligheten. Metoder har utvecklats för att möjliggöra en sådan utvärdering. Två olika sätt att bestämma dödsorsak har använts, dels den individuellt bestämda underliggande dödsorsaken som finns i dödsorsaksregistret och dels genom att mäta överdödligheten hos bröstcancerpatienterna.

I tre studier har åldersgrupperna 40-49, 50-69 och 70-74 år studerats där områden som inbjudit till mammografiscreening jämförts med områden som inte har haft screening. Efter en uppföljning på 8-10 år var bröstcancerdödligheten i områden med mammografiscreening 12% lägre för den yngsta åldersgruppen jämfört med områden utan screening. Motsvarande siffror för åldersgrupperna 50-69 och 70-74 år var 20% respektive 24%.

En studie genomfördes i Norra sjukvårdsregionen. Resultaten visade på en 28% lägre dödlighet i bröstcancer för kvinnor 40-74 år i de län som inbjöd till mammografiscreening än i de övriga länen efter 6 års uppföljning. I Gävleborgs län påbörjades allmän mammografiscreening som ett pilotprojekt redan 1974. Detta län
har jämförts dels med sina grannlän och dels med övriga Sverige under en 22 år lång uppföljning. Jämförelsen med grannlänen visade på en 21% lägre bröstcancerdödlighet i Gävleborgs län för kvinnor 40-64 år vid första inbjudan till screening.


**Slutsatser**

- Samtliga resultat visar på att allmän bröstcancerscreening med mammografi i ålder 40-74 år sänker bröstcancerdödligheten.

- Resultaten ligger i nivå med resultaten från tidigare studier.
ACKNOWLEDGEMENTS

This study was performed at the Department of Radiation Sciences and at the Oncological Centre, Umeå University. I want to express my gratitude to the many people who contributed to this work.

In particular, I want to thank the following persons:

My family, Britt-Inger – without you it would have been impossible – Erika, Frida and David, for being patient with me when I didn’t have time for you while at the same time I was monopolising the computer.

My parents Stig and Ingegärd, for always standing by me and supporting me.

My supervisor Per Lenner, for introducing me to the field of mammography screening, and because you succeeded, through the use of mild persistence, in finally convincing me to write a thesis on the subject. Thank you for all the discussions we have had and for your patience in trying to answer both easy and difficult questions. Thanks also for your constant support and for believing in me.

My associate supervisors and co-authors Lennarth Nyström and Sven Törnberg, for your enthusiasm and for your extensive knowledge about mammography screening. And thanks for the many years of meetings and writing.

My co-author Lars-Gunnar Larsson, for your inspiration and your untiring enthusiasm that rubs off on others. Thanks for taking the time to give me guidance instead of enjoying your well-deserved retirement. I consider it a privilege to have been able to share in your great knowledge of screening as well as of other areas.

My co-author Bengt Lundgren, for your interest and your extensive knowledge of mammography screening.

My chief and co-worker of many years Lena Damber, for your interest in and your support of my research. Thank you for the fruitful discussions and valuable viewpoints. And thanks for making it possible to adapt my work hours to allow time for my research.

Roger Henriksson and Sven-Ola Hietala, the Director and Assistant Director of the department, for your support.

Henrik Grönberg and Stefan Emdin, for valuable viewpoints at the midway-seminar.

Matti Hakama, who gave me crucial support and encouragement when I needed it most.

Pål Bordas, for your positive opinion of my research and for your useful viewpoints. Thanks also to all the staff at the mammography unit at Sunderby Hospital for taking such good care of me at my visit.
Ola Gardfjell and Ulf Hultin, for the world’s most dependable computer system, and for always being available and standing by me when I needed your help. Special thanks to Ola for your thoughtfulness in never crossing the goal-line before me when we have gone jogging.

Fredrik Wiklund, for valuable discussions, particularly those dealing with mortality and S-plus.

Åsa Sundberg, for help with the layout and editing as well as organising the photo, and not least for expert tips regarding Word.

Gunilla Andersson, Karin Andersson, Eva Hägglund, Lena Lundkvist, for assistance with a variety of things.

Carina Ahlgren, Gun-Britt Ivarsson and Anna Wernblom, for administrative assistance.

The staff at the Oncologal Centre, for all the ordinary days together especially the “fruitful” discussions around the coffee table – how trivial life would seem without you.

The Swedish Cancer Society, the European Commission, the Gävle Cancer Foundation, and the Lions’ Cancer Research Foundation, for financial support.
REFERENCES


