SCREEN-DETECTED BREAST CANCER AND PROGNOSIS

Tiina Lehtimäki

Department of Oncology, Helsinki University Central Hospital and
Institute for Molecular Medicine Finland

Academic dissertation

To be publicly discussed with the permission of the Faculty of Medicine of
the University of Helsinki, at Biomedicum Helsinki 1 (Haartmaninkatu 8), lecture hall 3,
on November 8th, 2013, at 12 noon.

Helsinki 2013
SUPERVISORS:
Docent, MD, PhD Johan Lundin
Institute for Molecular Medicine Finland
Biomedicum Helsinki 2U
University of Helsinki
Helsinki
Finland

Academy Professor Heikki Joensuu
Department of Oncology,
Helsinki University Central Hospital
Helsinki
Finland

REVIEWERS:
Professor Nea Malila
School of Health Sciences
University of Tampere
Tampere
Finland

Docent, MD, PhD Arja Jukkola-Vuorinen
Department of Oncology,
Oulu University Central Hospital
Oulu
Finland

OPPONENT:
Docent, MD, PhD Minna Tanner
Department of Oncology,
Tampere University Central Hospital
Tampere
Finland

©Tiina Lehtimäki 2013
Helsinki University Print (Unigrafia)
Helsinki 2013
Contents

1 ABSTRACT .......................................................................................................................... 6

2 ORIGINAL PUBLICATIONS .............................................................................................. 8

3 ABBREVIATIONS............................................................................................................... 9

4 INTRODUCTION ............................................................................................................... 10

5 REVIEW OF LITERATURE ................................................................................................. 12

5.1 Breast cancer ................................................................................................................ 12

5.1.1 Epidemiology and risk factors ................................................................................... 12

5.1.2 Treatment, recurrence and prognosis ...................................................................... 13

5.1.3 Prognostic factors ..................................................................................................... 15

5.1.3.1 Tumor-associated factors .................................................................................... 15

5.1.3.1.1 TNM-classification and staging ...................................................................... 15

5.1.3.1.2 Histological type ............................................................................................. 19

5.1.3.1.3 Grade ............................................................................................................... 20

5.1.3.2 Biological factors .................................................................................................. 20

5.1.3.2.1 Ki-67 ................................................................................................................. 20

5.1.3.2.2 p53 ................................................................................................................... 21

5.1.3.2.3 Estrogen and progesterone receptors .............................................................. 21

5.1.3.2.4 HER2 ............................................................................................................... 22

5.1.3.2.5 Molecular subtypes ........................................................................................ 23

5.1.3.3 Nottingham prognostic index and St.Gallen criteria ........................................... 24

5.1.3.3.1 Nottingham Prognostic Index (NPI) ................................................................. 24

5.1.3.3.2 St. Gallen criteria ............................................................................................. 24

5.1.4 Generalizability of prognostic factors .................................................................... 25

5.2 Mammography screening programs .......................................................................... 26

5.2.1 Overview ................................................................................................................... 26
5.2.2 Mammography screening in Finland ................................................................. 27
5.2.3 Mammography screening and tumor biology .................................................... 28
5.2.4 Mammography screening and prognosis .......................................................... 31
5.2.5 Biases associated with the interpretation of mammography screening results 34

6 AIMS OF THE STUDY ................................................................................................. 37

7 PATIENTS AND METHODS..................................................................................... 38

7.1 Patients .................................................................................................................. 38
7.1.1 The Finprog study cohort ................................................................................... 38
7.1.2 Definition of screen-detected and non-screen-detected breast cancers .......... 41
7.1.3 The SEER database ............................................................................................ 41

7.2 Methods .................................................................................................................. 42
7.2.1 Tissue microarray (TMA) .................................................................................... 42
7.2.2 Immunohistochemistry ....................................................................................... 43
7.2.3 Chromogenic in situ hybridization ..................................................................... 44
7.2.4 Statistical methods .............................................................................................. 44

8 PERMISSIONS .......................................................................................................... 46

9 RESULTS .................................................................................................................... 47

9.1 Prognostic factors in screen-detected breast cancer (Studies I-III) ....................... 47
9.2 Molecular subtypes in screen-detected breast cancer (Study I) ............................. 50
9.3 Prognosis of patients with screen-detected breast cancer (Studies I, II, III) ......... 50
9.4 Comparison of the FinProg and SEER cohorts (Study IV) ..................................... 51

10 DISCUSSION ............................................................................................................ 53

10.1 General remarks ................................................................................................... 53
10.2 Strengths and limitations of the study ................................................................. 53
10.3 Tumor biological profile ....................................................................................... 56
10.4 Prognosis .............................................................................................................. 57
1 ABSTRACT

The incidence of breast cancer has increased world-wide over recent decades and currently (in 2011) stands in Finland at 96.6 per 100,000 women. The survival rate of breast cancer patients has improved mainly due to innovative therapies and mammography screening programs. Despite improvements in survival, breast cancer is still the leading cause of cancer death among women.

Mammography screening programs were started in Finland in 1987. Since this year, the incidence of breast cancer has continued to increase and the survival rates of breast cancer patients have gradually improved. Breast cancers detected at screening mammography are, in general, smaller at the time of the diagnosis and less aggressive than cancers found by the woman herself, and are reported to be associated with generally favorable survival. It is, however, unclear whether this survival advantage can be explained by early detection and favorable tumor biology only, or whether other unknown factors might account for the survival advantage. The purpose of this study was to explore the method of detection as a prognostic factor for patient outcome. The aim was also to study the long-term prognosis and biological profiles of screen-detected breast cancers as compared to non-screen-detected cancers. The generalizability and geographic transportability of breast cancer survival estimates were investigated by comparing the current study series, the FinProg database, with a large dataset from the United States.

The FinProg database consists of 2,936 breast cancer patients diagnosed within five University Hospital regions in Finland in 1991-1992, representing about 50% of all breast cancers diagnosed in Finland during these years. Clinicopathological information was collected from the hospital case records. In addition, survival data were updated from the files of the Finnish Cancer Registry and Statistics Finland. Histological tumor samples from each patient’s cancerous tissue were collected for a tissue microarray and analyzed for a number of novel and established biomarkers. This allowed assessment of the tumor biological profiles, i.e. the breast cancer molecular subtypes.

The distribution of the molecular subtypes of screen-detected breast cancers differed from those of cancers detected by other methods. The molecular subtypes associated with favorable survival rates were more frequent in screen-detected breast cancers than in non-screen-detected cancers. The differences in molecular subtypes between mammography screen-detected cancers and cancers found otherwise could not, however, explain all of the survival advantage associated with breast cancer detection at screening, and detection of cancer at screening had independent influence on survival in multivariate analyses.

Breast cancers detected by the mammography screening program had a significantly lower 10-year risk for distant recurrence than cancers of similar size diagnosed by other methods. Patients with screen-detected breast cancer had also significantly better breast cancer-specific survival compared with patients with non-screen-detected cancer. This advantage persisted over a
15-year follow-up period. There was no significant difference in the risk for contralateral breast cancer between these two groups.

The generalizability and geographic transportability of the survival estimates obtained in the Finnish cohort were verified by comparing the FinProg database to a database from the SEER registry (Surveillance, Epidemiology and End Result, USA). Despite differences in demographic variables, use of adjuvant therapies and screening mammography between series, the prognostic factors produced close to overlapping survival curves. When commonly used prognostic factors, such as tumor size, nodal status and histological grade are evaluated, survival estimates were similar in these large unselected cohorts of breast cancer patients.

The current observations suggest that patients with screen-detected breast cancer often have a prognostically more favorable cancer biological profile than patients with a breast cancer detected by other methods. The biological factors investigated in the current study could not explain entirely the survival advantage associated with cancer detection at mammography screening, suggesting that there are other biological factors that may also explain the difference, which is in line with some previous studies. The risk for recurrence may be overestimated for women diagnosed by mammography screening unless the method of detection is taken into account in risk estimations, which may lead to overtreatment of patients with screen-detected breast cancer.
2 ORIGINAL PUBLICATIONS

This thesis is based on the following original publications, which are referred by their Roman numerals.


*=equal contribution
3 ABBREVIATIONS

AJCC  American Joint Committee on Cancer
BCSS  breast cancer-specific survival
CISH  chromogen in situ hybridization
CNBSS Canadian National Breast Cancer Screening Studies
CK    cytokeratin
DCIS  ductal carcinoma in situ
DDFS  distant disease free-survival
EGFR  epidermal growth factor receptor
ER    estrogen receptor
erbB2 HER2, human epidermal growth factor receptor-2
FISH  fluorescence in situ hybridization
HER2  human epidermal growth factor receptor-2
HIP   Health Insurance Plan
IHC   immunochemistry
LCIS  lobular carcinoma in situ
NPI   Nottingham Prognostic Index
NST   no specific type
OS    overall survival
p53   protein coded by the TP53 tumor supressor gene
PR/PgR progesterone receptor
RR    risk ratio
SEER  Surveillance, Epidemiology and End Result Program
TMA   tumor microarray
TNBC  triple negative breast cancer (ER-, PR- and HER2-)
UICC  the International Union Against Cancer
WHO   World Health Organization
4 INTRODUCTION

Breast cancer is the most common cancer in women worldwide. In the past decades its global incidence has increased (Weaver et al. 2006) (Glass, Hoover 1990). The incidence has increased also in Finland being 96.6 per 100,000 women in 2011, which relates to approximately 4,700 new cases diagnosed annually. The relative five-year survival estimate for a patient with breast cancer (of any stage) is currently approximately 89 % in Finland (Finnish Cancer Registry, Statistics from 2002-2009). Despite improvements in treatment and diagnostic methods, approximately 800 to 900 women die of breast cancer annually in Finland (Paajanen 2006) (www.cancerregistry.fi).

As expected, prognosis after the diagnosis and treatment of early stage tumors is more favorable than that of more advanced disease (Boyle 2003)(Tubiana, Koscielny 1999). According to recent statistics from the Surveillance, Epidemiology and End Results (SEER) database managed by the National Cancer Institute of the United States, the five-year relative survival rate of a patient with localized breast cancer is 98.4%, while it is only 23.8% in patients with advanced disease (www.seer.cancer.gov).

Mammography screening programs, launched in several countries in late 1980’s and early 1990’s, have been successful in detecting breast cancer early, and the proportion of early stage tumors has increased since then (Esserman, Shieh & Thompson 2009) (Paci et al. 2002). Mortality from breast cancer has decreased, which is likely mainly due to both earlier diagnosis and improved therapies (Duffy et al. 2002)(Esserman, Shieh & Thompson 2009, Wishart et al. 2008)(Hakama et al. 2008)(Sainsbury 2013).

Screen-detected breast cancers are reported to have a more favorable biological profile in general as compared with cancers detected outside of screening (Palka et al. 2008). The tumors found in screening tend to be better differentiated and more often hormone-receptor positive, have lower mitotic activity, less frequently HER-2 gene amplification and protein expression, and a low S-phase fraction (Dawson et al. 2009)(Nagtegaal et al. 2011)(Groenendijk et al. 2000)(Porter et al. 1999) (Klemi et al. 1992). Survival of patients whose tumor is detected at mammography screening is more favorable as compared with those women whose cancer is found outside of screening (Paajanen et al. 2006)(Mook et al. 2011b) (Allgood et al. 2011)(Otto et al. 2012)(Jonsson et al. 2003).

The stage shift (i.e. tumors are detected at an earlier stage in screening) and the biological factors assessed thus far attributed to cancers found at screening, have failed to explain all of the survival advantage of screen-detected cancers as compared with similarly treated cancers detected outside of screening. It has been estimated that only about half of the survival advantage can be explained by the observed differences in tumor size, stage, and grade (Gill et al. 2004). Another study suggests that 72% of the survival benefit is dependent on age and stage shift and the remaining 28% on other factors (Wishart et al. 2008). Some survival benefit remains even after adjustments for the commonly used prognostic factors (Olsson et al. 2012).
Although screen-detected breast cancers are in general smaller and associated with a prognostically more favorable clinicopathological and tumor biological profile than cancers detected outside of screening, very few studies are available where the survival estimates have been adjusted for a detailed prognostic factor profile between these two groups. Also, little is known about the long-term outcome (beyond the first ten years after the diagnosis) of patients with screen-detected breast cancer. Previous studies with long-term follow-up show that survival differences between subsets of breast cancer patients with different prognostic factors can decrease over time (Arriagada et al. 2006) (Warwick et al. 2004) (Railo et al. 2007) (Pichon et al. 1996), but whether this phenomenon also applies to the method of tumor detection as a prognostic factor is not known.

The purpose of the present study was to compare the clinical outcome of patients with screen-detected cancer to those with non-screen-detected tumor when the outcome is adjusted for a series of commonly used prognostic factors. The long-term influence of the method of breast cancer detection on survival was investigated. This study also explores the distribution of the molecular subtypes of the tumor between screen-detected and non-screen-detected breast cancers. To assess the representativeness of the current patient series (the FinProg series), we studied the generalizability and geographical transportability of the breast cancer survival estimates by comparing the FinProg population-based breast cancer series with a large series from another geographical area (the Surveillance, Epidemiology and End Result Program Registry data, the USA).
5 REVIEW OF LITERATURE

5.1 Breast cancer

5.1.1 Epidemiology and risk factors

In the past decades the global incidence of breast cancer has increased (Weaver et al. 2006) and it is the most common female cancer in the world comprising approximately 23% of all female cancers in the world (the World Cancer Research Fund International, www.wrcf.fi). The increasing number of women are diagnosed with breast cancer also in Finland (Moller et al. 2005). In the early 1980’s there were under 2,000 women diagnosed with breast cancer annually while the annual incidence in the late 2000’s was approximately 4,500 (www.cancerregistry.fi). The increase in incidence could be attributed to mammography screening programs (White, Lee & Kristal 1990)(Moller et al. 2005). The increase in breast cancer incidence, however, is observed also in unscreened women (White, Lee & Kristal 1990)(Harmer, Staples & Kavanagh 1999)(Parvinen et al. 2006), which may reflect time effects such as population ageing, better awareness of the disease and modern and more accurate diagnostic facilities (Harmer, Staples & Kavanagh 1999)(Jatoi, Miller 2003)(Gotzsche et al. 2012). Although relative 5-year survival of breast cancer patient is 89% in Finland, 800-900 women die of the disease annually (www.cancerregistry.fi).

Breast cancer in men is rare accounting only under 1% of all breast cancers (White et al. 2011). Only 10-20 male breast cancers are diagnosed annually in Finland (www.cancerregistry.fi)

Risk of breast cancer is low before the age of 30 but a moderate increase in incidence is observed already in the age group of 30 to 45. A clear increase in the incidence occurs from the age 45 onwards to 65-70 (www.cancerregistry.fi).

A third of breast cancer patients have at least one breast cancer patient among her relatives and the positive family history has been observed to increase the risk for breast cancer. The extent of this risk varies according to the nature of the family history for example according to the relation to the affected individual and number of affected relatives (Pharoah et al. 1997). Hereditary mutations cause approximately 5-10% of all breast cancers the proportion being probably larger in breast cancer patients aged under 30 (Nevanlinna H et Kallioniemi O-P 1999). Several gene mutations are identified of which highly penetrant gene mutations in tumor suppressor genes BRCA1 and BRCA2 are the most important (Nevanlinna H et Kallioniemi O-P 1999)(Turnbull et al. 2010)(Kuusisto et al. 2011)(Nasir et al. 2009).

Several other factors are also reported to be associated with the risk of breast cancer, most being related to the tissue levels of estrogens (nulliparity, late first childbirth, early menarche, late menopause, hormone replacement therapy and postmenopausal obesity) (Clemons M
A prior history of usual or atypical breast hyperplasia and other neoplastic breast disease (previous invasive breast carcinoma, LCIS, DCIS) are considered remarkable risk factors for breast cancer (Colditz, Rosner 2000) (Singletary 2003) as well as previous other cancers such as ovarian or uterine cancers (Arts-de Jong et al. 2013).

Radiation exposure from any source or purpose (i.e. diagnostic imaging, radiation therapy) especially to the chest area increases the risk for carcinogenesis and breast cancer growth. Association of alcohol consumption, smoking and physical inactivity with breast cancer have also been documented (Colditz, Rosner 2000) (Singletary 2003) (Singletary, Gapstur 2001) (Luo et al. 2011). Also disruption of circadian rhythm and light at nights have been associated with increased risk for breast cancer probably due to disturbance of melatonin secretion (Hansen 2001). The above life style and environmental risk factors, however, present a relatively modest or small increase in the relative risk. Physical activity, normal body mass index and breast feeding are reported to decrease the risk for breast cancer (Pieta et al. 2009).

The incidence of breast cancer have been observed to vary between races and social status. It is lower among the American black people than among whites probably reflecting lower risk for breast cancer among black people. The incidence of breast cancer rises with improving social status indicating increased risk for breast cancer among women with high economical status. The association between cultural factors and these differences is currently unclear (Vainshtein 2008).

The individual risk estimation of invasive breast cancer can be estimated using e.g. the freely available Gail model, which considers the most important risk factors such as age at menarche, age at first live birth, number of previous breast biopsies and the presence of atypical hyperplasia in a biopsy specimen, number of first-degree relatives with breast cancer and age (www.cancer.gov) (Gail et al. 1989) (Meads, Ahmed & Riley 2012) (Rockhill et al. 2001). Approximately 75% of women with breast cancer, however, have no observable risk factors for the disease (Berg 2010).

5.1.2 Treatment, recurrence and prognosis

Most breast cancer patients undergo breast surgery, and the type of the surgical procedure (breast-conserving or mastectomy) is often determined by the extent of the disease, but also other factors such as patient preference and the expected cosmetic outcome. Patient’s age and menopausal status, tumor biological type and stage, and the type of surgery are key factors that define the need for radiation therapy, hormonal therapy and chemotherapy. Tumor biological factors are used in decision-making on targeted therapies, such as expression of hormone receptors that guide for hormonal treatments (e.g. tamoxifen and aromatase inhibitors) and expression of HER-2 that is used to guide for the use of HER2-targeted monoclonal antibodies (e.g.
trastutsumab) and tyrosine kinase inhibitors (e.g. lapatinib). In general, systemic adjuvant therapy with such agents reduces the risk of recurrence substantially and improves survival (Early Breast Cancer Trialists' Collaborative Group (EBCTCG) 2005)(Early Breast Cancer Trialists' Collaborative Group (EBCTCG) et al. 2008)(Arriagada et al. 2006)(Jatoi et al. 2011) (Smith 2000)(Early Breast Cancer Trialists' Collaborative Group (EBCTCG) et al. 2011)(Hortobagyi 2012).

Breast cancer relapse may occur even twenty or more years after the primary diagnosis (Joensuu, Pylkkanen & Toikkanen 1999)(Brenner, Hakulinen 2004). Recurrences in the ipsilateral regional area (breast and/or lymph nodes) are classed as a locoregional recurrence while spread beyond regional lymph nodes is classed as distant recurrence. Breast tumor biological subtypes have been associated with preferential tumor spread to specific organs (Sihto et al. 2011).

In parallel with the increase in the incidence of breast cancer in recent years, mortality has decreased (Tabar et al. 2001)(Nystrom et al. 2002)(Tabar et al. 2003)(van Schoor et al. 2011), the reduction estimated as 20-35% amongst women at the screening age (50-69 years) (Hakama et al. 2008). The improvement in survival of breast cancer is considered to be mainly due to mammography screening programs and more effective local and adjuvant therapies (Berry et al. 2005)(Cronin et al. 2006)(Schopper, de Wolf 2009). The relative contributions of these factors are subject to statistical modeling and are likely to differ between study settings. In several studies it is estimated that mammography screening programs and adjuvant therapies each account for a 20-50% reduction of the breast cancer mortality(Kalager et al. 2009)(Berry et al. 2005)(Vilaprinyo, Puig & Rue 2012). A study from Norway suggests a 28% reduction in breast cancer mortality after implementation of screening mammography program of which 18% regarding as a time effect (breast cancer awareness, improvements in therapies, better diagnostic management) the rest (10%) being associated with the screening mammography program. In the same study the reduction in breast cancer mortality among older women (70 to 84 years, not screened) was approximately the same as among the screened women (50 to 69 years) (Kalager et al. 2010). Increased awareness of the disease, better breast imaging techniques and screening mammography programs lead to earlier diagnoses, and breast cancers are in general detected at an earlier stage (Harmer, Staples & Kavanagh 1999)(Jatoi, Miller 2003)(Gotzsche et al. 2012).

Breast cancer is a heterogenous disease and its natural course is variable. In old observational studies a small fraction of patients have survived for 10-15 years even without treatment, which could be used as an argument for indolence of few breast cancers or even rare regression of cancerous tissue under certain circumstances (Larsen, Rose 1999).
5.1.3 Prognostic factors

The aim when making therapy choices is to avoid overtreatment without compromising an effective cancer care. The decisions on an individual level depend on the risk of recurrence. The risk estimations are based on several prognostic and predictive factors including both patient-related and tumor-related factors. To select the appropriate individual therapy for each patient, prognostic and predictive factors should be comprehensively evaluated (Generali et al. 2011).

A prognostic factor is a characteristic of a patient or a disease that can be used to estimate the chance of recovery from that disease or the risk of the disease recurring (Mohsin et al. 2004)(McShane et al. 2006). In breast cancer they are used to estimate the risk of recurrence and death from breast cancer. The risk of recurrence influences the decision-making on treatment in order to choose the individual therapy (the strategy for surgery, radiotherapy and adjuvant therapy) for each patient. Tumor size, nodal status, stage, histological type, histological grade, hormone receptor (ER, PR) status, proliferation antigen Ki-67, tumor suppressor gene p53, HER2 expression and HER2 gene amplification, invasion to lymph or blood vessels and age at diagnosis are commonly used prognostic factors in breast cancer (O'Malley,F.P., Pinder,S.E. 2006)(Page, Jensen & Simpson 1998).

A predictive factor is defined as any measurement associated with the response to a particular therapy (Mohsin et al. 2004)(McShane et al. 2006), and some of the prognostic factors in breast cancer are also predictive. Factors that are targets for systemic therapy, like ER, PR and HER2 for example, may predict also the response to a particular systemic therapeutic agent (hormonal treatment, trastutsumab) and are thus both prognostic and predictive.

5.1.3.1 Tumor-associated factors

5.1.3.1.1 TNM-classification and staging

TNM-classification defined by the American Joint Committee on Cancer (AJCC) jointly with the TNM Committee of the International Union against Cancer (UICC) serves as an international system for anatomical staging of malignancy, which measures three major parameters: T for the diameter of the tumor, N for lymph node involvement and M for the presence or absence of distant metastases. The current TNM classification in 2009 is the seventh modification of the classification (Sobin LH, Gospodarowicz MK, Wittekind C ).

Pathologically defined tumor size (pT) reflects the potential for dissemination (Tubiana, Koscielny 1991)(Tubiana, Koscielny 1990) (Tubiana, Koscielny 1999)(Tubiana, Koscielny 1999) and is highly correlated with the nodal status and histological grade (Tabar et al. 1999)(Koscielny et al. 1984).
While growing, tumor may lose its differentiation (Duffy et al. 1991)(Tubiana, Koscielny 1991)(Henson et al. 1991) and may become more aggressive. Tumor size is a strong and generally accepted prognostic factor especially for node-negative patients and in patients with less than four positive nodes (Quiet et al. 1995)(Quiet et al. 1996)(Michaelson et al. 2002).

Screen-detected tumors are on average smaller than non-screen-detected tumors (Porter et al. 1999)(Dong et al. 2008) (Burrell et al. 1996) (Palka et al. 2008), and the proportion of smaller tumors has increased in parallel with the increase in breast cancer incidence and screening programs (Esserman, Shieh & Thompson 2009)(Paci et al. 2002) (Gilliland et al. 2000)(McCann, Stockton & Day 1998). The classification of T-status according to AJCC and UICC is shown in table 1.

<table>
<thead>
<tr>
<th>TXX</th>
<th>Primary tumor cannot be assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>T0</td>
<td>No evidence of primary tumor</td>
</tr>
<tr>
<td>Tis</td>
<td>Carcinoma in situ (DCIS*, LCIS**, Paget)</td>
</tr>
<tr>
<td>T1</td>
<td>Tumor ≤ 20 mm</td>
</tr>
<tr>
<td>T1mi</td>
<td>Tumor ≤ 1 mm</td>
</tr>
<tr>
<td>T1a</td>
<td>Tumor &gt; 1 mm but ≤ 5 mm</td>
</tr>
<tr>
<td>T1b</td>
<td>Tumor &gt; 5 mm but ≤ 10 mm</td>
</tr>
<tr>
<td>T1c</td>
<td>Tumor &gt; 10 mm but ≤ 20 mm</td>
</tr>
<tr>
<td>T2</td>
<td>Tumor &gt; 20 but ≤ 50 mm</td>
</tr>
<tr>
<td>T3</td>
<td>Tumor &gt; 50 mm</td>
</tr>
<tr>
<td>T4</td>
<td>Tumor of any size with direct extension to the chest wall and/or to the skin</td>
</tr>
<tr>
<td>T4a</td>
<td>Extension to the chest wall, not including only pectoralis muscle</td>
</tr>
<tr>
<td>T4b</td>
<td>Ulceration and/or ipsilateral satellite nodules and/or edema of the skin, which do not meet the criteria for inflammatory ca</td>
</tr>
<tr>
<td>T4c</td>
<td>Both T4a and T4b</td>
</tr>
<tr>
<td>T4d</td>
<td>Inflammatory carcinoma</td>
</tr>
</tbody>
</table>

*DCIS=ductal carcinoma in situ
**LCIS=lobular carcinoma in situ

Table 1. TNM classification according to AJCC.
(All measures in greatest dimension.)
The axillary nodal status correlates with the tumor size (Tabar et al. 1999)(Koscielny et al. 1984), and in general early nodal involvement predicts rapid tumor growth (Tubiana, Koscielny 1991). The presence of lymph node metastases and the number of involved nodes are generally considered to be the most important prognostic factors of recurrence and survival (Toikkanen, Joensuu 1990). Most screen-detected cancers are node negative (Crosier et al. 1999). The classification of nodal status according to AJCC and UICC is shown in table 2.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NX</td>
<td>Regional lymph nodes cannot be assessed (e.g., previously removed).</td>
</tr>
<tr>
<td>N0</td>
<td>No regional lymph node metastases.</td>
</tr>
<tr>
<td>N1</td>
<td>Metastases to movable ipsilateral level I, II axillary lymph node(s).</td>
</tr>
<tr>
<td>N2</td>
<td>Metastases in ipsilateral level I, II axillary lymph nodes that are clinically fixed or matted. OR Metastases in clinically detected bilateral mammary lymph nodes in the absence of clinically evident axillary lymph node metastases. OR Metastases in ipsilateral level I, II axillary lymph nodes fixed to one another (matted) or to other structures.</td>
</tr>
<tr>
<td>N2a</td>
<td>Metastases only in clinically detected bilateral mammary lymph nodes and in the absence of clinically evident level I, II axillary lymph node metastases.</td>
</tr>
<tr>
<td>N2b</td>
<td>Metastases only in clinically detected bilateral mammary lymph nodes and in the absence of clinically evident level I, II axillary lymph node metastases.</td>
</tr>
<tr>
<td>N3</td>
<td>Metastases in ipsilateral infraclavicular (level III axillary) lymph node(s) with or without level I, II axillary lymph node involvement. OR Metastases in clinically detected bilateral mammary lymph node(s) with clinically evident level I, II axillary lymph node metastases. OR Metastases in ipsilateral supraclavicular lymph node(s) with or without axillary or internal mammary lymph node involvement.</td>
</tr>
<tr>
<td>N3a</td>
<td>Metastases in ipsilateral infraclavicular lymph node(s).</td>
</tr>
<tr>
<td>N3b</td>
<td>Metastases in ipsilateral internal mammary lymph node(s) and axillary lymph node(s).</td>
</tr>
<tr>
<td>N3c</td>
<td>Metastases in ipsilateral supraclavicular lymph node(s).</td>
</tr>
</tbody>
</table>

Table 2. N-status classification according to AJCC.
Distant disease free survival is commonly used as an end point in survival analyses. In that sense, M-status is an end point parameter rather than a prognostic factor. Postsurgical staging of the disease is performed by combining information from tumor size, the nodal status and presence of distant spread according to the TNM-classification (AJCC Cancer Staging Manual). Stage is an established prognostic factor in breast cancer as well as in most other solid cancers (Sant et al. 2003). The postsurgical staging is shown in table 3.

<table>
<thead>
<tr>
<th>Stage</th>
<th>T</th>
<th>N</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Tis</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>IA</td>
<td>T1b</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>IB</td>
<td>T0</td>
<td>N1mi</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T1b</td>
<td>N1mi</td>
<td>M0</td>
</tr>
<tr>
<td>II A</td>
<td>T0</td>
<td>N1c</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T1b</td>
<td>N1c</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>IIB</td>
<td>T2</td>
<td>N1</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>IIIA</td>
<td>T0</td>
<td>N2</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T1b</td>
<td>N2</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>N2</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td>N1</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td>N2</td>
<td>M0</td>
</tr>
<tr>
<td>IIIB</td>
<td>T4</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T4</td>
<td>N1</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T4</td>
<td>N2</td>
<td>M0</td>
</tr>
<tr>
<td>IIIC</td>
<td>Any T</td>
<td>N3</td>
<td>M0</td>
</tr>
<tr>
<td>IV</td>
<td>Any T</td>
<td>Any N</td>
<td>M1</td>
</tr>
</tbody>
</table>

b=T1 includes Tmi
c=T0 and T1 tumors with nodal micrometastases only are excluded from Stage IIA and classified Stage IIB

Table 3. Staging of breast cancer according to AJCC Cancer Staging Manual.
5.1.3.1.2 Histological type

Histological typing is often performed according to the World Health Organization (WHO) classification criteria (Elston, Ellis 1991) and the histologic types are grouped into the following entities: in situ carcinomas (DCIS and LCIS), invasive ductal carcinoma (NST, no specific type), invasive lobular carcinoma and breast carcinoma of a special type.

Breast cancer cells are derived from epithelial cells lining mammalian ducts and lobules. In situ carcinomas are neoplastic proliferations of cancer cells within the ducts or lobules without invasion through the epithelium to the surrounding tissue. In contrast, invasive carcinomas invade through the epithelium and the surrounding tissue and are potential for lymphatic or hematogenic spread. Ductal carcinoma in situ comprises approximately 7% and lobular carcinoma in situ 1% of all breast cancers in Finland (FinProg, unpublished data).

Ductal carcinoma is the most common type of invasive breast carcinoma (50-80%) (Eheman et al. 2009) (Weigelt, Geyer & Reis-Filho 2010). In Finland two thirds of all breast carcinomas are of ductal type (FinProg, unpublished data). Due to its unspecified nature, the classification of ductal carcinoma is based on exclusion of other tumor types. An increased frequency of ductal carcinomas is seen in patients aged less than 35 years (67%) compared to older patients (53%) (O'Malley, F.P., Pinder, S.E. 2006). About 70-80% of ductal carcinomas express ERs and 15-20% are HER2 positive (Yamamoto et al. 2009) (Ladjemi et al. 2010).

Invasive lobular carcinoma comprises 5-15% of all invasive breast carcinomas (Weigelt, Geyer & Reis-Filho 2010) (O'Malley, F.P., Pinder, S.E. 2006) (Lopez, Bassett 2009) and in Finland the proportion of lobular cancer is approximately 16% (FinProg, unpublished data). There is an increased proportion of lobular – as well as special type – carcinomas among older patients and it is hypothesized that the use of hormone replacement therapy for menopausal women increases the amount of invasive lobular carcinoma more than the other histological types (Li et al. 2003). Infiltrating lobular carcinoma is distinguished from infiltrating ductal carcinoma by its small cells and diffuse, single cell file growth pattern (Howell, Harris 1985). It is radiographically more elusive and has lower density and thus may remain occult in mammography (Li et al. 2003) (Adler, Engel 1990) (Holland, Hendriks & Mravunac 1983) (Garnett, Wallis & Morgan 2009). Furthermore, it is also less likely to have microcalcifications than tumors of ductal morphology (O'Malley, F.P., Pinder, S.E. 2006) (Michael, Garzoli & Reiner 2009). Several subtypes of lobular carcinomas have been identified (e.g. classical, tubulo-lobular, alveolar, solid and mixed) with distinctive properties. In general, the prognosis of a lobular carcinoma is similar to or slightly better than that of ductal carcinoma of the same stage (Fritz et al. 2010) (Dian et al. 2009) (Toikkanen, Pylkkänen & Joensuu 1997) (Dian et al. 2009). Contralateral invasive breast carcinoma, however, is reported to be more frequent in patients with a history of invasive lobular carcinoma (Rai et al. 2011) (Chen et al. 1999) (Horn, Thompson 1988). Hence, patients with lobular carcinoma should be intensively followed up for the risk of cancer in the contralateral breast.

Special type carcinomas are histologically a heterogeneous group of tumors (Weigelt, Geyer & Reis-Filho 2010), which includes types such as papillary, medullary, tubular and mucinous
carcinoma. The proportion of special type carcinomas is slightly greater in older patients than younger ones (O’Malley, F.P., Pinder, S.E. 2006). In general, survival of patients with special type of carcinomas is more favorable than that of women with ductal or lobular carcinoma at the same stage (Toikkanen, Pylkkanen & Joensuu 1997) (Simpson, Page 1992), but some types of special carcinoma, like metaplastic carcinoma, are associated with unfavorable outcomes (Toumi et al. 2011) (Luini et al. 2007).

5.1.3.1.3 Grade

The histological evaluation of differentiation of a tumor is generally performed according to the Bloom – Richardson method of grading (Bloom, Richardson 1957) (WHO classification) based on three components; the mitotic count, nuclear pleomorphism and tubule formation (Elston, Ellis 1991). The most commonly used grading system is the three-grade scale, but in some pathological units a four-grade scale is being used (like in the SEER database described later) (Henson et al. 1991).

Histological grade is an independent prognostic factor in several studies and correlates significantly with survival (Lundin et al. 2001) (Rakha et al. 2008). The prognosis of patients with a low grade (grade 1) tumor is more favourable than women with a high grade (grade 3) tumor. Grade 2 tumors show an intermediate prognosis at least in the early years of follow-up although having less favourable survival in long term follow-up. Histological grade is highly correlated to tumor size, and tumors have been hypothesized to become less differentiated during growth (Tabar et al. 1999) (Koscielny et al. 1984) (Rakha et al. 2010).

5.1.3.2 Biological factor

5.1.3.2.1 Ki-67

Tumor growth is determined by the balance between tumor cell proliferation, cell senescence (aging), tumor dormancy (latency) and apoptosis. The Ki-67 protein is associated with cell proliferation in cancerous and normal tissues (Gerdes et al. 1984). Ki-67 is a nuclear protein which is encoded by the MK167 gene located in chromosome 10 (10q25) (Fonatsch et al. 1991) and is expressed in all phases of the cell cycle (G1, S, G2, mitosis) except in G0. Thus, immunostaining for Ki-67 is capable of distinguishing proliferating cells from resting cells (Yerushalmi et al. 2010). In malignant tumor tissues Ki-67 production is activated and serves as a marker for cancer growth,
and increased expression of the protein is usually associated with a poor prognosis (Railo et al. 1993)(de Azambuja et al. 2007), even though Ki-67 positive tumors tend to respond better to cytotoxic agents (Keam et al. 2011). Expression of Ki-67 correlates with other markers of proliferation such as the S-phase fraction, the mitotic index, and expression of cyclins and cell proliferation-associated kinases (Nakano, Oka 1993)(Gasparini et al. 1991)(Dedic Plavetic et al. 2013)(Ermiah et al. 2012)(Cance, Liu 1995). The Ki-67 protein is usually assessed either from fresh frozen or paraffin-embedded tissue samples using immunohistochemistry (IHC).

There has been an increasing interest in Ki-67 in recent years (Yerushalmi et al. 2010), although there is no consensus on how to assess and evaluate the results of Ki-67 staining. In a meta-analysis of 38 studies that examined the results obtained by two different antibodies (anti-Ki-67 and anti-MIB-1), the cut-off values for positivity varied widely from 3.5 to 34% (de Azambuja et al. 2007).

5.1.3.2.2 p53

The first tumor suppressor gene described in several human cancers was p53 (Gasco, Shami & Crook 2002). In contrast to the oncogenes which are activated by mutation or by other mechanisms, tumor suppressor genes are attributed to tumorigenesis due to the loss of function, for example by mutation (Weinberg 1991). A key function of the p53 protein is to sense and drive repair of DNA damage, and to inhibit proliferation of mutated cells suppressing the malignant transformation of cells also by inducing apoptosis of damaged cells (Symonds et al. 1994). Mutation of TP53 may result in abrogation of its growth regulatory functions leading to tumorigenesis. It is estimated that TP53 mutations contribute to the development of up to 50% of all cancers (Gasco, Shami & Crook 2002) (Levine 2012). Expression of p53 in cancerous tissue is generally associated with poor prognosis (Thor et al. 1992)(Kronqvist et al. 2004).

5.1.3.2.3 Estrogen and progesterone receptors

Estrogen receptor (ER) is a nuclear transcription factor that is involved in breast development, growth, differentiation and tumorigenesis. ER binds estradiol and regulates the transcription of several genes like progesterone receptor (PR) (O'Malley, F.P., Pinder, S.E. 2006)(Oesterreich, Lee & Davidson 2010). Progesterone receptor expression is induced by ER and serves as a surrogate marker for ER activity and as a predictive factor for effect of hormonal therapy. Hence, the primary goal when assessing ER and PR is to find patients who may benefit from endocrine therapy (Mohsin et al. 2004)(Oesterreich, Lee & Davidson 2010). The incidence of hormone
receptor-positive tumors has increased over the past decades (Li, Daling & Malone 2003)(Pujol et al. 1994)(Glass, Hoover 1990). Of all breast cancers 70-80% are estrogen receptor-positive and 60-70% progesterone receptor-positive (Li, Daling & Malone 2003). Most of all breast cancers are positive for both estrogen and progesterone receptors while only approximately 20% are negative for both (ER- and PR-) (Bauer, Parise & Caggiano 2010).

The hormone receptor status of a histological tumor sample is mainly assessed by immunohistochemical methods which are superior to ligand-binding assays (Mohsin et al. 2004)(Hammond et al. 2010) (Harvey et al. 1999). Because up to 20% of the analyses are estimated to be inaccurate, an international Expert Panel convened by The American Society of Clinical Oncology and the College of American Pathologists developed a new recommendation for optimal hormone receptor testing and interpretation procedures in 2010. Patients with ≥1% positive tumor cells may benefit from endocrine therapy (Harvey et al. 1999) and thus, according to the recommendation, the threshold for positivity was lowered from ≥10% to ≥1% (Hammond et al. 2010).

ER and PR positivity correlates with a low histological grade (a more differentiated tumor), low cell proliferation rate, HER-2 negativity and p53 negativity. Tumor hormone receptor positivity is associated with favorable short-term survival (Li, Daling & Malone 2003). However, several studies show that the ER status is a time-dependent prognostic variable; the advantage of hormone receptor positivity diminishes and even disappears along with a long patient follow-up time (Anderson et al. 2006)(Railo et al. 2007)(Takeuchi et al. 2005)(Jatoi et al. 2011).

5.1.3.2.4 HER2

The proto-oncogene HER2 (synonyms c-erbB2, HER-2/neu) is located on chromosome 17q and encodes a transmembrane tyrosine kinase (Ross, Fletcher 1998). The name is derived from the Human Epidermal Growth Factor Receptor as it features homology with the epidermal growth factor receptor (EGFR)(Slamon et al. 1987). The protein product that the gene encodes appears in low levels in normal breast tissue (Hoff et al. 2002) but may be over-expressed by malignant cells. When amplified and/or over-expressed by the tumor cells it increases cell proliferation as well as invasiveness and dissemination of tumor cells. Hence, over-expression of HER2 is associated with greater tumor aggressiveness and poorer prognosis (Purdie et al. 2010)(Toikkanen et al. 1992)(Joensuu et al. 2003). HER2 is assessed using quantitative polymerase chain reaction (qPCR), immunohistochemistry (IHC) or in situ hybridization techniques (FISH and CISH), of which the in situ hybridization methods may be considered most reliable for clinical purposes (Wolff et al. 2007)(Koninki et al. 2009).

Of all breast cancers 15-20% are considered HER2 positive (Yamamoto et al. 2009)(Ladjemi et al. 2010), although a decline in its proportional incidence has been observed probably due to mammography screening programs and the increase in the use of postmenopausal hormone

5.1.3.2.5 Molecular subtypes

DNA profiling studies have identified clinically relevant molecular subtypes of breast cancer (Perou et al. 2000)(Prat, Perou 2010) (Sorlie et al. 2001). Differences in gene-expression reflect variability in tumor biology and patterns of growth, and are associated with response to treatment and clinical outcome (Voduc et al. 2010)(Parise et al. 2009)(Vallejos et al. 2010)(Dawood et al. 2011). Hormone receptors (ER and PR) as well as HER2 are widely accepted prognostic and predictive factors in breast cancer (Bauer, Parise & Caggiano 2010)(Chacon, Costanzo 2010). Both hormone receptor-positive and HER2-positive breast cancers constitute effective molecular targets for adjuvant therapy. A minority (10-15%) of breast cancers expresses neither hormone receptors nor HER2 protein (Yamamoto et al. 2009) and lack therapeutic molecular targets for endocrine therapy and trastuzumab (Foulkes, Smith & Reis-Filho 2010). Microarray-based expression analysis methods have identified distinctive subtypes within this triple negative breast cancer (TNBCs; ER-, PR-, HER2-) with different clinical outcomes (Nielsen et al. 2004)(Seal, Chia 2010)(Constantinidou, Jones & Reis-Filho 2010)(Fornier, Fumoleau 2012). Basal-like breast cancer is characterized by expression of genes including basal cytokeratins CK5/6 and CK14, EGFRs (van de Rijn et al. 2002)(Nielsen et al. 2004) and other characteristics associated with poor prognosis. However, several studies show that 16-47% of TNBCs do not express these basal markers and are associated with a more favorable prognosis as compared to basal-like breast cancers (Yamamoto et al. 2009)(Yamamoto, Iwase 2010). The intrinsic subtypes of breast cancer include luminal A (usually approximated with ER+ and/or PR+, HER2-), luminal B (ER+ and/or PR+, HER2+), HER2 enriched (ER-, PR-, HER2+), TNBCs with positive basal markers (basal-like) and nonexpressor type (negative for all key classifiers).

Due to the genetic diversity of breast cancer, patients with the same stage of disease may have markedly different treatment responses and overall outcomes. Gene microarray technologies such as OncotypeDX, PAM50 and Mammaprint may help in assessing the likelihood of disease recurrence and response to certain types of adjuvant therapies (Nature 2012 Vol 486)(Goldhirsch et al. 2011).
5.1.3.3 Nottingham prognostic index and St.Gallen criteria

5.1.3.3.1 Nottingham Prognostic Index (NPI)

The Nottingham Prognostic Index (NPI) integrates three independent prognostic factors (grade, regional lymph node involvement and tumor size), and was first presented in 1982 (Haybittle et al. 1982) and revised ten years later (Galea et al. 1992)(Suen, Chow 2006). It was originally derived from a multivariate regression analysis where the three factors were found to provide significant and independent prognostic information. The original formula was further simplified for day-to-day use. The Nottingham prognostic index is calculated as follows:

\[
NPI = \text{grade}* + \text{lymph node stage}** + (0.2 \times \text{size}***)
\]

* grade: scores 1-3 based on the modified Bloom-Richardson grading
** lymph node stage:
  1= no nodal involvement
  2=1-3 positive nodes or positive node in the internal mammary chain
  3=>3 positive nodes or involvement of both axillary and internal mammary lymph nodes
*** size: maximum diameter in centimeters

The Nottingham Prognostic Index can be categorized into excellent or good prognosis with the index value ≤3.40, moderate prognosis with values of 3.41-5.40, and poor prognosis with the NPI score >5.41 (Galea et al. 1992). The NPI index correlates well with survival and is able to identify different prognostic subgroups (Blamey et al. 2010)(Todd et al. 1987)(Page, Jensen & Simpson 1998), which can be divided further into smaller subgroups by the index (Blamey et al. 2007). Ten-year survival of patients with NPI 2.00-2.49 was reported to be 96% and with NPI 6.41-7.00 only 37% (Blamey et al. 2007).

5.1.3.3.2 St. Gallen criteria

The St. Gallen risk categories were designed for making treatment decisions for early breast cancer based on literature evidence and clinical expertise (Garassino et al. 2009). The first St. Gallen consensus statements were made in 1988 (Glick 1988)(Bauer, Parise & Caggiano 2010) and the criteria and therapy recommendations are updated every two years. According to the algorithm,
therapy decisions are recommended to be made based on cancer biological properties and the stage of the disease (Goldhirsch et al. 2009). The 2011 St. Gallen consensus meeting focused on therapy recommendations for different molecular subtypes of early breast cancer (Gnant, Harbeck & Thomssen 2011).

5.1.4 Generalizability of prognostic factors

A large number of prognostic factors have been proposed in breast cancer, but only few can be considered as established enough to be used in the clinical routine. This has partly been attributed to deficiencies in reporting of prognostic factor analysis results (McShane et al. 2005). The generalizability and transportability of prognostic factor results between institutions and populations is claimed to be addressed insufficiently (McShane et al. 2005). Results of generalizability have typically been obtained from studies that focus on validation of prognostic models (Blamey et al. 1979)(Haybittle et al. 1982)(Todd et al. 1987)(Olivotto et al. 2005)(Hajage et al. 2011). Prognostic models combine information about multiple prognostic factors and are sometimes available as open-access web-based calculators (Ravdin et al. 2001)(Lundin et al. 2004)(Michaelson et al. 2011). These prognostic tools are proposed as aids for the clinicians in the decision-making process for the selection of treatment of individual patients by modeling complex relationships between prognostic factors and outcomes. The generalizability and transportability of the quantitative survival estimates across populations has been shown to be accurate in some studies (Olivotto et al. 2005), but also less well calibrated in others (Campbell et al. 2009)(Mook et al. 2009)(Hajage et al. 2011).

Comparisons between countries have demonstrated marked variations in breast cancer survival outcomes, and variations also within the continents (Walters et al. 2013). The survival rates for U.S. breast cancer patients may in general be more favorable as compared to European women (Rosso et al. 2010)(Gatta et al. 2000)(Sant et al. 2004), but no nationwide statistics are available from the U.S., and the U.S. regions with less favorable results may have been underrepresented in the analyses. Regional differences in survival, however, are observed within Europe with the highest survival rates in the Nordic countries and the lowest ones in countries of the Eastern Europe (Sant et al. 2003)(Quinn, Martinez-Garcia & Berrino 1998)(Sant et al. 1998). Representativeness of patient cohorts and validation of methodological issues such as the cause of death registration need to be considered in the interpretation of these differences (Gatta et al. 2000). In general, breast cancers may be diagnosed at an earlier stage in the U.S. than in Europe (Allemani et al. 2013). The comparisons between the SEER (U.S.) and the EUROCARE (European) datasets also show that the average number of axillary lymph nodes examined is significantly higher in the U.S. patients compared with European breast cancer patients (Gatta et al. 2000),
which suggests a more accurate staging of the disease, but might also lead to stage shift. Also differences in the availability of modern treatments, frequency of early diagnosis, and general awareness of the disease in the population may relate to variations in survival between countries (Sant et al. 2004).

5.2 Mammography screening programs

5.2.1 Overview

The main objective of screening is to detect the disease at a stage when it has not disseminated and is curable, consequently potentially improving the prognosis of the disease and reducing mortality. The screened disease should be an important health problem and efficient therapy should be available. The screening itself should be cost-effective, risks associated with the screening test itself minimal and the benefits greater than the risks (Wilson, Jungner 1968) (Kopans, Monsees & Feig 2003).

The aim of breast cancer screening programs is to reduce breast cancer deaths by early detection when treatment is still less harmful and potentially curative. Breast cancer is often symptomless especially in the early stages and may become mammographically detectable during this preclinical phase. The principal aim of mammography screening programs is to detect breast cancers during this sojourn time, and prior to cancer dissemination (Warner 2011).

The first breast cancer screening programs were launched in the late 1980’s and early 1990’s, mainly in developed countries. The efficacy of mass screening was first evaluated by Tabar et al, who reported a 31% reduction in breast cancer mortality in a randomized controlled trial of 162,981 women aged 40-74 (Tabar et al. 1985). These results were supported by other studies (Andersson 1981)(Andersson et al. 1988).

The policies for mammography screening programs are still inconsistent, and the optimal screening interval is under continuous debate (Michaelson, Halpern & Kopans 1999)(White et al. 2004). The sojourn time is associated with age being generally longer in elderly people than patients aged 40-49 (van Oortmarssen et al. 1990)(Tabar et al. 1995) and varies from one tumor to another (Berry 2008). Therefore, the estimations of sojourn times are crude and vary from 1 year to 5 years. The sojourn time in women at the screening age (50-69 years) is estimated to be approximately 3.5 years and biennial screening is considered sufficient, the costs and subsequent examinations increasing with more frequent intervals, and less frequent screening resulting in increasing numbers of interval cancers (Kerlikowske et al. 1995). Annual screening, however, may be more effective among younger (40-49 years) women (White et al. 2004)(Wai et al. 2005) with
an estimated shorter sojourn time (McPherson et al. 1997). Although annual screening may minimize the risk for advanced disease among younger (40-49 years), the risk for false-positive mammography is higher with annual screening (Kerlikowske et al. 2013). The other study, however, suggests that there is no difference in the incidence or incidence-based mortality between patients aged 40-49 years who were invited annually or triennially (Parvinen et al. 2011). According to a recent recommendation, women aged 50-74 years should be screened biennially and women aged 40-49 years only when risk-level is considered increased (US Preventive Services Task Force 2009).

Mammography screening is carried out with two-view or one-view X-rays. The sensitivity of two-view mammography is reported to be higher (Burrell et al. 1996)(Young et al. 1997) and to reduce the recall rates (Blanks et al. 2005), although a meta-analysis suggested that in women older than 50 years the sensitivity of one-view mammography is similar to that of two-view mammography and that there is no mortality difference between these imaging procedures (Kerlikowske et al. 1995).

According to a study based on the Breast Cancer Surveillance Consortium data from the U.S., 1.6% of screening mammograms lead to a subsequent biopsy and 0.43% to the diagnosis of invasive breast carcinoma in 1996-2001 (Weaver et al. 2006). The number of patients that need to be screened to avoid one death is estimated to be 470 (Duffy 2005)(Swedish Organised Service Screening Evaluation Group 2006).

Despite the well-documented effect of screening, it has certain limitations. Ten to twenty percent of breast carcinomas are not detectable by mammography (Gilliland et al. 2000)(Evans 2012). A failure to detect cancer may be attributed to the testing facilities, image interpretation problems as well as the breast or tumor characteristics (Gilliland et al. 2000). According to the observations by Saarenmaa, the sensitivity of mammogram as a method of tumor detection was 92% in defined regions in Finland (Saarenmaa 2001). The sensitivity of a mammogram increases with age as the breast tissue changes from dense to fatty (Kerlikowske et al. 1996). False negatives reduce the sensitivity and the efficacy of mass screening and the sensitivity is less when there is only one reader for the mammogram (Hukkanen, Kivisaari & Vehmas 2006)(Burrell et al. 1996).

One cause for criticism of screening is the false positive findings and the subsequent anxiety (Jorgensen 2013)(Elmore et al. 2005). According to studies on the psychological consequences following a false positive mammogram, however, the increase in the distress level has been only moderate (Aro et al. 2000)(Schwartz et al. 2000), and false positive findings are considered to be an acceptable consequence of mammography screening (Woloshin, Schwartz 2010).

5.2.2 Mammography screening in Finland

The nationwide mammography screening program was started in Finland in 1987 (Elovainio et al. 1989)(Hakama et al. 1997). In 1991-92 the municipalities in Finland offered biennial screening for
women aged 50-59 years, although a few municipalities offered mammography screening also for other age cohorts (40-49 years, 60-69 and/or ≤ 70 years). Currently (2012), in accordance with the Finnish legislation based on the Government Decree on screenings (1339/2006) (www.cancer.fi), a mammography screening program is provided by the municipalities for women aged 50-69 years biennially.

The breast cancer incidence rates increased markedly in Finland after the introduction of the screening program (Anttila et al. 2008)(Moller et al. 2005). Of all breast cancers 20-25% are detected by mammography screening, which means approximately 1,000 annual cases in Finland (Paajanen 2006)(Paajanen et al. 2006). In the subset of women who are at the screening age, 65% of cancers are detected by screening, while 35% are interval cancers (Suomen Lääkärilehti 2010).

In 2006 the European Commission published the 4th edition of the European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis which defines the quality standards for national breast cancer screening programmes in Europe. The coverage of a national breast cancer screening program should be at least 95%. The guidelines also suggest that the proportion of women invited who attend for screening should exceed 70% the desirable proportion being >75%. The desirable level of recalls for further examination is 3% and should not exceed 5%. The coverage of breast cancer screening program in women aged 50-59 is nearly 100% in Finland. The attendance rate during early 1990’s was 88-89% (Suomen Lääkärilehti 2010)(Dean, Pamilo 1999) with a slight decrease to 87.9% in 1992 and to 85.5% in 2009 (www.cancerregistry.fi). In 2009 recall rate for further examinations was 2.61% and in 0.56% malignancy was detected (www.cancer.fi). Currently, screening program in Finland meets the acceptable standard according to the EU-guidelines.

The specificity in the Finnish mammography screening program was reported to be 97% (Dean, Pamilo 1999) and the sensitivity 58% (Anttila, Koskela & Hakama 2002). Among true participants, however, the sensitivity was estimated to be 70% (Anttila, Koskela & Hakama 2002). According to a study by the Finnish Cancer Registry, screening reduced mortality from breast cancer among the invited women by 22% and among the true participants 28%, but there was no effect on overall mortality (Sarkeala, Heinavaara & Anttila 2008b).

5.2.3 Mammography screening and tumor biology

Mammography screening programs are considered to be successful in detecting of more early stage tumors, but they are also associated with an increase in breast cancer incidence in general (Esserman, Shieh & Thompson 2009) (Paci et al. 2002)(Malmgren et al. 2012)(Gabriel, Wilson & Helvie 1997). In most of the previous studies tumor size of screen-detected cancers was significantly smaller as compared with cancers found outside of screening (Redondo et al 2012)(Dawson et al. 2009)(Crispo et al. 2013). In only one study tumor size did not differ
significantly between the two groups, but in this study screen-detected tumors were compared to interval cancers rather than to all non-screened cancers (Gilliland et al. 2000). Screen-detected breast cancers are also more often node-negative (Weaver et al. 2006), and only a few studies report no significant difference in the nodal status between screen-detected and non-screen-detected groups (Crosier et al. 1999)(Collett et al. 2005)(Chuwa et al. 2009)(Porter et al. 1999)(Table 4).

Tubular histology and a large *in situ* component are also associated with screen-detected breast cancer (Porter et al. 1999). In line with the observations that the histological grade may worsen while the tumor grows, screen-detected tumors are generally better differentiated than non-screen-detected tumors (Table 4).

Screen-detected tumors have also a more favorable biomarker profile. Tumor growth depends on the balance between cell proliferation and apoptosis. Biological markers indicating cell proliferation, like mitotic count, S-phase fraction and Ki67 expression, are lower in screen-detected cancers. Tumor suppressor gene *TP53* mutations and *HER-2* gene amplification and protein expression are also more common in non-screen-detected tumors (Anttinen et al. 2003)(Soomro et al. 1993). In general, ER expression is more common in screen-detected tumors suggesting favorable prognosis and responsiveness to hormonal therapy. With regard to PR expression, however, the observations have been more variable (Table 4).
<table>
<thead>
<tr>
<th>paper</th>
<th>size</th>
<th>N+</th>
<th>grade</th>
<th>histology</th>
<th>ER</th>
<th>PR</th>
<th>HER2</th>
<th>Ki-67</th>
<th>p53</th>
<th>MI</th>
<th>groups compared</th>
<th>Total N</th>
<th>age (yrs)</th>
<th>dg</th>
<th>in situ/inv.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crosier et al, 1999</td>
<td>↓***</td>
<td>NS</td>
<td>↓***</td>
<td>↑***</td>
<td>NS</td>
<td>↓***</td>
<td>↓***</td>
<td>↓***</td>
<td>scrs vs. intervals</td>
<td>135</td>
<td>50-65</td>
<td>invasive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Porter et al, 1999</td>
<td>↓***</td>
<td>NS</td>
<td>↓</td>
<td>more tubular**</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>↓</td>
<td>scrs vs. intervals</td>
<td>429</td>
<td>40-&gt;</td>
<td>1988-1993</td>
<td>invasive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dawson et al, 2009</td>
<td>↓***</td>
<td>less***</td>
<td>↓***</td>
<td>NS</td>
<td>↑***</td>
<td>↑</td>
<td>NS</td>
<td>↓***</td>
<td>scrs vs symptomatic</td>
<td>1,379</td>
<td>50-70</td>
<td>1991-2004</td>
<td>invasive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dong et al, 2008</td>
<td>↓***</td>
<td>less***</td>
<td>↓***</td>
<td>↑***</td>
<td>NS</td>
<td>↓</td>
<td>↓***</td>
<td>scrs vs symptomatic</td>
<td>5,481</td>
<td>&gt;40</td>
<td>1997-2005</td>
<td>invasive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gilliland et al, 2000</td>
<td>NS</td>
<td>↓</td>
<td>NS</td>
<td>NS</td>
<td>↓</td>
<td>↓***</td>
<td>scrs vs. intervals</td>
<td>127</td>
<td>any</td>
<td>1991-1993</td>
<td>both</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ernst et al, 2002</td>
<td>↓***</td>
<td>less***</td>
<td>↑</td>
<td>↑</td>
<td>NS</td>
<td>↓</td>
<td>↓***</td>
<td>scrs vs. clinically detected</td>
<td>1,198</td>
<td>50-69</td>
<td>1996-1999</td>
<td>invasive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collett et al, 2005</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>↑</td>
<td>↑</td>
<td>NS</td>
<td>↓</td>
<td>↓***</td>
<td>scrs vs. intervals</td>
<td>190</td>
<td>any</td>
<td>1996-2001</td>
<td>invasive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groenendijk et al, 2003</td>
<td>↓***</td>
<td>less***</td>
<td>↑***</td>
<td>NS</td>
<td>↓***</td>
<td>scrs vs. intervals vs. symptomatic</td>
<td>581</td>
<td>any</td>
<td>1991-1999</td>
<td>invasive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klemi et al, 1992</td>
<td>↓***</td>
<td>less***</td>
<td>↓</td>
<td>less likely to be ductal</td>
<td>↑</td>
<td>↓</td>
<td>scr (prevalence screen) vs. symptomatic</td>
<td>251</td>
<td>40-79</td>
<td>1987-1990</td>
<td>invasive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagtegaal et al, 2011</td>
<td>↓***</td>
<td>less***</td>
<td>↓***</td>
<td>more tubular***</td>
<td>↑***</td>
<td>↑***</td>
<td>scrs vs. intervals vs. unexposed</td>
<td>21,382</td>
<td>50-64</td>
<td>1998-2001</td>
<td>both</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweeney et al, 2007</td>
<td>NS</td>
<td>NS</td>
<td>↓</td>
<td>NS</td>
<td>↑***</td>
<td>scrs vs. prevalence screen vs. symptomatic</td>
<td>437</td>
<td>50-64</td>
<td>2000-2002</td>
<td>both</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pålka et al, 2008</td>
<td>↑***</td>
<td>less***</td>
<td>↓</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>scrs vs. symptomatic vs. intervals</td>
<td>569</td>
<td>45-65</td>
<td>2004-2007</td>
<td>invasive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immonen-Räihä et al, 2005</td>
<td>↓***</td>
<td>less***</td>
<td>↓***</td>
<td>less likely to be ductal**</td>
<td>scrs including interval carcinomas vs. cancers</td>
<td>527</td>
<td>40-74</td>
<td>1987-1993</td>
<td>invasive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mook et al, 2011</td>
<td>↓***</td>
<td>less***</td>
<td>↓***</td>
<td>NS</td>
<td>↑***</td>
<td>scrs vs. nonscreening-related</td>
<td>2,962</td>
<td>50-69</td>
<td>1990-2000</td>
<td>invasive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chuwa et al, 2009</td>
<td>↓***</td>
<td>NS</td>
<td>↓</td>
<td>more DCIS***</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>scrs vs. symptomatic</td>
<td>767</td>
<td>40-65</td>
<td>2002-2003</td>
<td>both</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olsson et al, 2012</td>
<td>↓***</td>
<td>less**</td>
<td>↓***</td>
<td>NS*</td>
<td>NS*</td>
<td>NS*</td>
<td>NS*</td>
<td>scrs vs. non-screen-detected</td>
<td>466</td>
<td>47-75</td>
<td>1991-1996</td>
<td>invasive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crispo et al, 2013</td>
<td>↓***</td>
<td>less**</td>
<td>↓***</td>
<td>NS</td>
<td>↑</td>
<td>NS</td>
<td>↓</td>
<td>scrs (opportunistics incl.) vs. symptomatic</td>
<td>448</td>
<td>≥18</td>
<td>2004-2006</td>
<td>invasive</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05  
** p ≤ 0.01  
*** p ≤ 0.001  
NS = not significant difference  
¥ adjusted for tumor size and age  
° adjusted for age, body mass index, menopausal status and hormone replacement therapy and tumor size.  
scr = screen-detected  
N+ = node positivity  
MI = mitotic count  

Table 4. Tumor biological characteristics and screen-detected breast cancer (compared to non-screen-detected cancers)
5.2.4 Mammography screening and prognosis

The efficacy of mammography screening programs in reducing mortality from breast cancer is dependent on the coverage of a screening program as well as the quality and efficacy of treatment. The risk of death from breast cancer is smaller for screening participants than for non-participants (Sarkeala, Heinavaara & Anttila 2008b). Mammography screening has been found to reduce mortality from breast cancer 20%-39% among women aged 50 years or older (Kerlikowske et al. 1995)(Andersson et al. 1988)(Tabar et al. 2011) and somewhat less among younger women (19% in women aged 40-49) (World Health Organization. 2002). In Finland, the reduction in breast cancer mortality after mammography screening program has been 20-25% (Sarkeala, Heinavaara & Anttila 2008b)(Sarkeala, Heinavaara & Anttila 2008a).

The risk for breast cancer recurrence is significantly lower with screen-detected tumors than non-screen-detected ones (Paajanen 2006)(Immonen-Raiha et al. 2005). Post-recurrence prognosis, however, is similar in both groups (Immonen-Raiha et al. 2005).

Although tumors detected by mammography screening are smaller in general and have a more favorable biological profile, several studies show that screen-detected cancers have a better prognosis even after adjustment for these differences (Moody-Ayers, Wells & Feinstein 2000)(Mook et al. 2011a)(Crispo et al. 2013). The effects of certain prognostic markers, such as ER, diminish with the follow-up time, but according to four studies (Duffy et al. 1991)(Shen et al. 2005)(Mook et al. 2011b)(Immonen-Raiha et al. 2005) screening remained as an independent prognostic factor even after ten years of follow-up. In the study with the longest follow-up time (median, 16 years), screening remained as an independent prognostic factor after adjustments for tumor size, the nodal status and disease stage (Shen et al. 2005). This study, however, dates back to a time-period when the diagnostic procedures and the treatment regimens differed from those that are used today, and in situ carcinomas were also included in the cohort with the longest follow-up time. A more recent study showed that screen-detection is a favorable prognostic factor independent of stage migration (Mook et al. 2011b). This study compared also the risk estimations made with and without the data on the method of detection using the Adjuvant!-online prognostic tool. The Adjuvant!-online is a computer model based on several datasets (including the SEER) to assist in decision making about adjuvant therapy in clinical practice (Ravdin et al. 2001). Both overall survival and breast cancer-specific survival rates were underestimated for those whose cancer was detected in a screening program in individual risk estimations.

McPherson with her colleagues compared the survival of patients with screen-detected and non-screen-detected breast cancer after adjustment for lead time bias. To reduce the lead time effect they excluded the first year of follow-up from survival analysis from patients with screen-detected cancer. The exclusion was based on estimation of 2-year sojourn time for women aged 40-49 in the cohort and application Feig’s rationale meaning that lead time is one-half the duration of the sojourn time. The survival advantage of women with screen-detected cancer persisted even after this correction for lead time bias (McPherson et al. 1997). On the contrast, in the other study applying the correction model for lead time, size, nodal status and lead time explained almost all of the survival advantage. In that study, the lead time adjustment was performed applying the
method of Duffy et al (2008) which uses the complex equations to calculate the lead time for patients with screen-detected cancer. These estimated lead times are then subtracted from each patient’s survival time. This method, however, is based on average sojourn time and does not take patient’s age into account (Duffy et al. 2008b).

The effect of mammography screening on early detection of breast cancer may be underestimated when a study includes invitees not willing to attend screening and who are not true participants (noncompliance) (Berg 2010). The effect may also be underestimated if opportunistic breast imaging occurs in the screened cohort (contamination) (Evans 2012). A summary of the literature on mammography screening and prognosis is provided in Table 5.
<table>
<thead>
<tr>
<th>Paper</th>
<th>Advantage</th>
<th>Adjusted</th>
<th>Adjusted HR (95% CI)</th>
<th>End-point</th>
<th>Follow-up Time</th>
<th>Groups Compared</th>
<th>Age (years)</th>
<th>N</th>
<th>dg</th>
<th>Invasive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allgood et al., 2011</td>
<td>Observed</td>
<td>Size, nodal status, grade, NPI, lead time*</td>
<td>0.70 (0.63-0.78)#</td>
<td>OS</td>
<td>10 yrs</td>
<td>Sr vs. symptomatic</td>
<td>50-64</td>
<td>19,411</td>
<td>1986-2004</td>
<td>Invasive</td>
</tr>
<tr>
<td>Dawson et al., 2009</td>
<td>Observed</td>
<td>Size, nodal status, grade, NPI</td>
<td>0.43-0.69 ***</td>
<td>OS</td>
<td>7 yrs</td>
<td>Sr vs. symptomatic</td>
<td>50-70</td>
<td>1,379</td>
<td>1991-2004</td>
<td>Invasive</td>
</tr>
<tr>
<td>Wishart et al., 2003</td>
<td>Observed</td>
<td>Age, NPI</td>
<td>0.79 (0.63-0.99)</td>
<td>OS</td>
<td>5 yrs</td>
<td>Sr vs. symptomatic</td>
<td>50-64</td>
<td>5,694</td>
<td>1969-2003</td>
<td>Invasive</td>
</tr>
<tr>
<td>McPherson et al., 1997</td>
<td>Observed</td>
<td>Size, nodal status</td>
<td>0.53-0.67</td>
<td>OS</td>
<td>4.7 yrs</td>
<td>Mammogram vs. BSE vs. CBE vs. PI</td>
<td>40-48</td>
<td>971</td>
<td>1966-2002</td>
<td>Invasive</td>
</tr>
<tr>
<td>Shen et al., 2005</td>
<td>Observed</td>
<td>Stage, trial effect</td>
<td>0.65 vs. intervals/indications</td>
<td>BCSS</td>
<td>16 yrs (HP)</td>
<td>Sr vs. controls vs. non-attendees</td>
<td>0.49 (CNBSS-1)</td>
<td>608 (HP)</td>
<td>1969-2002</td>
<td>Both</td>
</tr>
<tr>
<td>Duvall et al., 2001</td>
<td>Observed**</td>
<td>Size, nodal status, grade, age, county</td>
<td>0.57 (0.38-0.85)#</td>
<td>BCSS</td>
<td>11 yrs</td>
<td>Sr vs. prevalence screen vs. later screens vs. intervals vs. refusers</td>
<td>40-69</td>
<td>1,562</td>
<td>1971-2004</td>
<td>Invasive</td>
</tr>
<tr>
<td>Mork et al., 2011</td>
<td>Observed</td>
<td>Age, size, grade, nodal status, ER, adjuvant systemic therapy</td>
<td>0.62 (0.50-0.78)</td>
<td>BCSS</td>
<td>11 yrs</td>
<td>Sr vs. non-screening-related</td>
<td>50-49</td>
<td>2,592</td>
<td>1990-2000</td>
<td>Invasive</td>
</tr>
<tr>
<td>Robinson et al., 2006</td>
<td>Not observed</td>
<td>Age, size, nodal status</td>
<td>0.79 (0.40-1.55)=</td>
<td>BCSS</td>
<td>10 yrs</td>
<td>Sr vs. intervals vs. unexposed</td>
<td>50-64</td>
<td>3238</td>
<td>1969-2002</td>
<td>Invasive</td>
</tr>
<tr>
<td>Nagtegaal et al., 2011</td>
<td>Observed</td>
<td>Age, size, nodal status, histology, grade</td>
<td>0.65 (0.59-0.77)</td>
<td>BCSS</td>
<td>10 yrs</td>
<td>Sr vs. clinically-detected</td>
<td>47-75</td>
<td>466</td>
<td>1991-1996</td>
<td>Invasive</td>
</tr>
<tr>
<td>Olsson et al., 2005</td>
<td>Observed</td>
<td>Age, size, nodal status, grade</td>
<td>0.40 (0.22-0.72)</td>
<td>BCSS</td>
<td>8.1 yrs</td>
<td>Sr vs. clinically-detected</td>
<td>40-74</td>
<td>527</td>
<td>1991-1993</td>
<td>Invasive</td>
</tr>
<tr>
<td>Immonen-Räihä et al., 2007</td>
<td>Observed</td>
<td>Size, nodal status, grade, histology, treatment, age</td>
<td>0.72 (0.40-1.55)</td>
<td>BCSS</td>
<td>10.0 yrs (surv-det)</td>
<td>Sr vs. non-screen-detected</td>
<td>50-64</td>
<td>767</td>
<td>2002-2003</td>
<td>Both</td>
</tr>
<tr>
<td>Chuwa et al., 2009</td>
<td>Not observed</td>
<td>Size, nodal status, grade, LVI, HER-2, stage</td>
<td>0.53 (0.08-3.43)</td>
<td>RFS</td>
<td>3.3 yrs</td>
<td>Sr vs. symptomatic</td>
<td>40-45</td>
<td>372</td>
<td>2002-2003</td>
<td>Both</td>
</tr>
<tr>
<td>Dong et al., 2009</td>
<td>Observed</td>
<td>ER, PR, HER-2/neu, Ki-67, and treatment</td>
<td>0.63-0.85 ***</td>
<td>RFS</td>
<td>3.2 yrs</td>
<td>Sr vs. symptomatic</td>
<td>&gt;40</td>
<td>5,481</td>
<td>1991-2005</td>
<td>Invasive</td>
</tr>
<tr>
<td>Critto et al., 2013</td>
<td>Observed</td>
<td>Age, size, nodal status, molecular subtypes</td>
<td>0.50 (0.24-1.03)</td>
<td>RFS and OS</td>
<td>5.15 yrs</td>
<td>Sr vs. symptomatic</td>
<td>&gt;18</td>
<td>448</td>
<td>2014-2016</td>
<td>Invasive</td>
</tr>
</tbody>
</table>

NPI = Nottingham Prognostic index
OS = Overall survival
RFS = Recurrence-free survival
BCSS = Breast cancer specific survival
HP = Health insurance Plan of New York trial
CNBSS = Canadian National Breast Screening Studies
BC = Breast cancer
* = adjustment for lead time bias using the method of Duffy et al. (2008)
** = in prevalence screen only
*** = with different adjustments
BSE = Breast self-examination
CBE = Clinical breast examination
PI = Patient incident
= grades 1-2, 0.63 (0.33-1.21)
# = first screening round vs. controls
### = first screening round and follow-up time

Table 5. Prognosis and screen-detected breast cancer (compared to non-screen-detected breast cancers)
5.2.5 Biases associated with the interpretation of mammography screening results

The observed survival advantage related to screen-detected breast cancer is confounded by several biases, and the size of the absolute benefit of mammography screening programs remains a subject of discussion. Cancers detected by mammography screening are identified earlier in their natural history, which is also the purpose of any effective screening program. The **lead time** is the time period that the diagnosis is advanced by the application of the mammography screening (Figure 1). (Duffy et al. 2008a). The time period from the diagnosis to an outcome measure is longer for screen-detected tumors than non-screened tumors even if screening may not prolong the absolute duration of life at all (Figure 1).

![Lead Time Diagram]

**Figure 1. Lead time bias.**

The pre-clinical, but mammographically detectable, phase of tumor progression is referred to as 'the sojourn time’. Its duration varies according to the natural course of the tumor. A slowly growing tumor may be asymptomatic, but mammographically detectable, for a longer time period and more susceptible to detection in a mammography screening program as compared to a fast-growing tumor, which causes the **length time bias** (Figure 2). (Zahl, Strand & Maehlen 2004). Because of the slow pace of progression, patients with such a cancer have long survival times in general.
The lead and length time biases associated with survival analyses may be addressed in several ways. Adjusting for known prognostic factors and excluding the most indolent tumors, *in situ* carcinomas, reduces the effect of the lead time and length time biases. The models for leadtime corrections have been developed. They are, however, based on approximations of the rate of transition from asymptomatic to symptomatic disease disregarding the age of the patient (Duffy et al. 2008a)(Feig 1995).

Mammography screening programs increase the incidence of early stage tumors (Malmgren et al. 2012). However, the incidence of advanced disease has not decreased respectively (Autier et al. 2011)(Nederend et al. 2012)(Bleyer, Welch 2012). Some cancers with a low malignant potential may not even manifest during a patient’s lifetime without screening or might not progress even if left untreated. This may lead to the extreme form of the length time bias, overdiagnosis (Jorgensen 2013). Estimates of overdiagnosis range from 1% to 30% (Duffy et al. 2005)(Zahl, Strand & Maehlen 2004). Approximately two lives are saved by breast screening mammography for one overdiagnosed case (Duffy et al. 2010).
Women who attend a screening program are a self-selected and possibly health-conscious group of the general population, and may differ in their overall hazard of dying from those who do not attend screening. This may cause a **selection bias**.

Some types of breast cancers are more easily detected by mammography than others. Potentially less aggressive, spiculated or calcified tumors are relative easy to detect as compared to lobular carcinomas, which may be reflected in the distribution of the histologic types among interval carcinomas (Porter et al. 1999)(Garnett, Wallis & Morgan 2009)(Holland, Hendriks & Mravunac 1983), and may lead to a **detection bias** (Ikeda et al. 1992).
6 AIMS OF THE STUDY

The purpose of this thesis was to compare the clinicopathological characteristics, biologic profiles and outcome of breast cancers detected at mammography screening with those found by other means.

The specific aims of the study were:

1. To investigate the risk of recurrence and long-term breast cancer-specific survival of women diagnosed with screen-detected breast cancer versus those with non-screen-detected breast cancer.

2. To describe screen-detected and non-screen-detected breast cancers with regard to their clinicopathological characteristics and tumor biological profiles.

3. To analyze the generalizability of survival estimates obtained in a nationwide breast cancer cohort (FinProg, Finland) and another cohort (SEER - Surveillance, Epidemiology and End-Results program) from another geographical region (the USA).

4. To evaluate the method of detection as a factor associated with prognosis adjusting for an extensive series of known prognostic factors.
7 PATIENTS AND METHODS

7.1 Patients

7.1.1 The Finprog study cohort

The Finprog cohort consists of breast cancer patients diagnosed in Finland in 1991-92. The patients were identified from the files of the Finnish Cancer Registry. Five geographically defined regions were selected, including all five University Hospitals in Finland (the Helsinki and Uusimaa region, Pirkanmaa, Varsinais-Suomi, Northern-Savo, and Northern Bothnia and Lapland) (Figure 3).

Figure 3. The coverage of the FinProg study. The proportions of breast cancer patients diagnosed within each region and the time period of 1991-2 and who were included in the study are shown.
The FinProg-series consists of 2,936 breast cancer patients, which comprises 53% of the 5,551 women who were diagnosed with breast cancer in Finland 1991-92. About 50 patient and tumor characteristics as well as the vital status were captured from each patient’s hospital records. Male breast cancer patients were excluded from the study. Survival data were completed and updated also from the files of the Finnish Cancer Registry and Statistics Finland.

Fourteen (0.5%) patients had zero follow-up time (the diagnosis was made at autopsy, or the patient died during or soon after breast surgery) and were excluded. Each of these fourteen patients had also another criterion for exclusion from the studies, consisting of presence of distant metastases at the time of the primary diagnosis (n=7), and/or surgery was not performed (n=10), and/or incorrect breast cancer diagnosis (n=1). During data extraction 46 (1.6%) patients were discovered not to have breast cancer diagnosed and were excluded from the analyses. After these exclusions, 2,883 patients remained in the FinProg series.

For the final analyses we excluded also patients with no invasive breast cancer. These consisted of 191 (6.6%) patients with ductal carcinoma in situ and 18 (0.6%) with lobular carcinoma in situ. Patients who presented with distant metastases (had stage IV breast cancer; n=131, 4.5%) were excluded from the analyses. Patients who were not surgically treated (n=101, 3.5%) were also excluded. All patients with missing data regarding these exclusion criteria were excluded, except for a missing M-status, since staging examinations were often not routinely carried out for small tumors and exclusion of these patients would thus have resulted in a substantial bias. Any one particular patient could have been excluded for several reasons. After exclusions, there were 2,449 patients left in the study cohort.

For each study we used additional exclusion criteria applied to the specific study setting. In patients with bilateral breast cancer or a malignancy other than breast cancer, the true origin of potential distant metastases is difficult to determine. Patients with bilateral breast cancer or other malignancy (except for basal cell carcinoma or cervical carcinoma in situ) were excluded in studies I, II and IV. In study III, a main analysis was performed excluding all patients with bilateral breast cancer and patients with a malignancy other than breast cancer (except for those diagnosed before 1985). In the first sensitivity analysis, exclusions were performed according to the time of the diagnosis of the potential contralateral carcinoma or malignancy other than breast cancer. The second sensitivity analysis was performed including all patients with bilateral breast cancer or other malignancy. The distributions of molecular subtypes in screen-detected and non-screen-detected breast cancer were investigated in study I, and all patients with missing data in one or more of the classification markers were excluded. The exclusions applied to each study are shown in table 6.
Table 6. Study exclusion criteria.

Systemic adjuvant therapy was administered to 21% of patients with screen-detected cancer compared with 41% of those who had non-screen-detected cancer. Adjuvant therapy usually consisted of tamoxifen or CMF (cyclophosphamide, methotrexate and fluorouracil). Hormonal therapy was administered to 14% of patients with cancer detected in mammography screening and to 25% of those with non-screen-detected cancer. The proportions for patients who received chemotherapy were 8% and 15%, respectively.

The median follow-up time for DDFS was 9.5 years and for BCSS 15.4 years.
7.1.2 Definition of screen-detected and non-screen-detected breast cancers

In this study we refer to tumors detected in a mammography screening program as “screen-detected” breast cancers. Those breast cancers that were not diagnosed by mammography screening are referred to as “non-screen-detected” breast cancers. The non-screen-detected cancers include patients who were under or above the screening age, interval cancers and patients invited to screening but who decided not to attend. Asymptomatic breast cancers that were not detected in screening were considered as non-screen-detected cancers regardless of their detection at mammography outside of a screening program. The screening rounds were not investigated in the current study.

7.1.3 The SEER database

The Survival, Epidemiologic, End Results (SEER) program, organized by the National Cancer Institute, is an authoritative source of information on cancer incidence and surveillance in the United States. SEER is a public-use database which collects clinicopathological information from patients with various types of cancer. It collects cancer incidence and survival rates from public-based cancer registries covering approximately 28% of the U.S. population (www.seer.cancer.gov). To compare the SEER database with the FinProg database, all female breast cancer patients diagnosed in 1991-92 were selected from the SEER database. From the SEER dataset, a subset of 43,238 women diagnosed with breast cancer in 1991 or 1992 was extracted (National Cancer Institute released April 2004, based on the November 2003 submission.). Male patients, patients with bilateral carcinoma, carcinoma in situ or other carcinoma (except for basal cell carcinoma and cervical carcinoma in situ) and patients with distant metastasis at the time of diagnosis were excluded, leaving 25,753 patients for the comparisons. The median follow-up for the patients from the SEER database was 9.7 years.

There were a few remarkable differences in the data content between the SEER and the FinProg databases. Firstly, in the SEER data the grade of cancer differentiation is coded either using the three-grade or the four-grade scale. In the FinProg database the grade of differentiation is coded using the three-grade scale from grade 1 (well differentiated) to grade 3 (poorly differentiated). In the comparisons between the FinProg and SEER series, grades 3 and 4 in the SEER data were combined resulting to differences in the distribution of grades 1-3 between the two datasets. Secondly, the SEER database does not contain information about adjuvant breast cancer treatment. Thirdly, the method of detection or mammography screening is not recorded in the SEER database. Therefore, we made estimations using data from other sources with respect of the frequency and type of adjuvant therapy (Harlan et al. 2002)(Mariotto et al. 2002) and screening (Breen et al. 2001).
7.2 Methods

7.2.1 Tissue microarray (TMA)

When available, histological tumor samples from each patient’s cancerous tissue was collected. Formalin-fixed, paraffin-embedded tumor samples were used for tissue microarrays (Gillett et al. 2000)(Kallioniemi et al. 2001) (Camp, Charette & Rimm 2000)(Nocito et al. 2001)(Kononen et al. 1998)(Selvarajan et al. 2006). Representative tumor regions were first defined from hematoxylin and eosin stained sections and marked. Tumor tissue array (TMA) blocks were obtained by punching a 0.6-mm tissue cylinder through a histologically representative area of each “donor” tumor block. The cylindrical sample of the tumor was inserted into an empty “recipient” tissue array paraffin block using a specific instrument (Figure 4).

Figure 4. The process of preparing a tumor microarray. A representative tumor area is marked on an original tumor slide (a). Cylindrical core of cancerous tissue was punched from the tissue block (b) and then transferred (c) and inserted (d) into a tissue array block, which was sectioned for biomarker analysis (e).

Three cores were cut from each donor breast cancer specimen for the TMA. From the tumor samples available, 19 tissue array blocks were prepared, each containing 50–144 tumor samples.
Sections of 5 µm were cut and processed for immunohistochemistry and chromogenic in situ hybridization. Evaluation of the tissue array slides was aided by the use of a computer-controlled and motorized specimen stage (EcoDrive; Maérzhauser, Inc., Wetzlar, Germany) installed on an Olympus BX50 (City, Country) microscope (Study I-IV).

### 7.2.2 Immunohistochemistry

The TMAs were utilized to determinate ER, PR, HER1, HER2, p53, Ki-67, GATA-3, CK5, CK18 and cyclooxygenase-2 protein expression using immunohistochemistry and the respective monoclonal antibodies for their detection (Table 7). Immunostaining for ER and PR was considered as positive when ≥ 10% of the cancer cells showed staining and as negative when < 10% of tumor cells showed staining (studies I, II and IV). According to recent recommendations (Hammond et al. 2010), the 1% cut-off point was used in Study III, and thus breast cancer tumor was classified as hormone receptor negative when < 1% of the cancer cells stained and as positive when ≥ 1% of cells showed staining. Immunostaining for Ki-67 was considered as high when ≥ 10% of the tumor cells showed staining, moderate when 1-10% of the cells showed staining and low when < 1% of the cells were positive. Because patients whose tumors expressed either low or moderate Ki-67 protein levels had similar survival, we grouped Ki-67 staining as either low-moderate or high.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Antibody</th>
<th>Producer</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER</td>
<td>6F11</td>
<td>Novocastra Laboratories Ltd, Newcastle, UK</td>
<td>1:500</td>
</tr>
<tr>
<td>PR</td>
<td>312</td>
<td>Novocastra Laboratories Ltd, Newcastle, UK</td>
<td>1:500</td>
</tr>
<tr>
<td>HER1</td>
<td>NCL-EGFR</td>
<td>Novocastra Laboratories Ltd, Newcastle, UK</td>
<td>1:500</td>
</tr>
<tr>
<td>HER2</td>
<td>CB11,</td>
<td>Novocastra Laboratories Ltd, Newcastle, UK</td>
<td>1:200</td>
</tr>
<tr>
<td>p53</td>
<td>D07</td>
<td>Novocastra Laboratories Ltd, Newcastle, UK</td>
<td>1:500</td>
</tr>
<tr>
<td>Ki-67</td>
<td>MM-1</td>
<td>Novocastra Laboratories Ltd, Newcastle, UK</td>
<td>1:1000</td>
</tr>
<tr>
<td>GATA-3</td>
<td>SC-268</td>
<td>Santa Cruz Biotechnology, Santa Cruz, U.S.</td>
<td>1:300</td>
</tr>
<tr>
<td>CK5</td>
<td>M7237</td>
<td>DakoCytomation, Glostrup, Denmark</td>
<td>1:25</td>
</tr>
<tr>
<td>CK18</td>
<td>NCL-CK18</td>
<td>Novocastra Laboratories Ltd, Newcastle, UK</td>
<td>1:20</td>
</tr>
<tr>
<td>cyclooxygenase-2</td>
<td>160112</td>
<td>Cayman Chemical Company, Michigan, U.S.</td>
<td>1:200</td>
</tr>
</tbody>
</table>

Table 7. Monoclonal antibodies
7.2.3 Chromogenic in situ hybridization

HER-2 was assessed by chromogenic in situ hybridization. The microarray slides were deparaffinized and incubated in 0.1 M Tris-HCl (pH 7.3) at 92°C for 10 min, followed by cooling for 20 min at room temperature. Enzymatic digestion was done by applying 100 µl of digestion enzyme onto the slides (Digest-All III solution; Zymed, Inc., San Francisco, CA). After dehydration, a ready-to-use digoxigenin-labeled HER-2/neu DNA probe (Zymed) was applied on the slides. The sections were denatured on a thermal plate, and hybridization was carried out overnight at 37°C. The HER-2/neu probe was detected by means of sequential incubations with mouse antidigoxigenin (diluted 1:300; Roche Biochemicals, Mannheim, Germany), antimouse-peroxidase polymer (Powervision+; Immunovision, Inc. Daly City, CA), and diaminobenzidine as a chromogen. The tissue sections were lightly counterstained with hematoxylin and embedded. Positive and negative control samples (tumors with and without HER-2 amplification visualized by fluorescence in situ hybridization) were included in every hybridization batch. The sections were evaluated using a microscope with a 40 x dry objective. Amplification was defined as more than or equal to six signals per nucleus in >50% of cancer cells or when large gene copy clusters were identified (Tanner et al. 2000).

7.2.4 Statistical methods

Frequency tables were analyzed using the χ² test. Life-tables were calculated according to the Kaplan-Meier method. Distant disease-free survival (DDFS) was defined as the time from the date of the diagnosis to the date of first distant recurrence of cancer. Patients who did not have distant metastases were censored on the last date of follow-up or the date of death, if death was considered to result from an intercurrent cause. Overall survival (OS) was defined as the time from the date of the diagnosis to the date of death censoring patients who were alive on the last date of follow-up. Breast cancer specific-survival (BCSS) was defined as the time from the date of the diagnosis to the date of death due to breast cancer, censoring patients who were alive or who died from a competing cause on the last date of follow-up.

The time of primary diagnosis and the time when each patient entered the study were defined as time zero. By definition, all the patients are alive at the time of diagnosis. Thus, according to general statistical procedure, patients with zero or close to zero follow-up time were excluded from the study (patients who were diagnosed at autopsy or who died during or soon after surgery).

The log-rank test was used to assess the statistical significance of the difference in survival between two groups. Relative survival was calculated according to the Ederer II method (Ederer F 1959). Multivariate survival analysis was calculated using the Cox proportional hazards model. P-
values were two-tailed and values < 0.05 were considered significant. Statview (version 5.0, SAS Institute Inc, Cary, NC) and STATA (StataCorp.2007.Stata Statistical Software: Release 10. College Station, TX: StataCorp LP) statistical software were used in statistical calculations. A web-based system for individual survival estimation, a case-match calculation (Lundin et al. 2003), based on Kaplan-Meier product-limit survival analyses, was utilized to calculate the survival of different subgroups within the FinProg patient series.
8 PERMISSIONS

We had permissions for this study from each of the five University Hospitals in Finland (HUS 9.2.2004 and 11.2.2010). Permission to use tissue samples for research purposes was provided by the Ministry of Social Affairs and Health, Finland (permission 15.12.1997 and No.123/08/97). The study was approved by the Ethical Committee of Surgery of the Hospital Districts of Helsinki and Uusimaa.
9 RESULTS

9.1 Prognostic factors in screen-detected breast cancer (Studies I-III)

In 1991-92 mammography screening was mainly offered to women aged 50-59 years, although some municipalities offered mammography screening also for women beyond that age group. Therefore, the proportion of screen-detected breast cancers in the age group of 50-69 differed from that of other age groups. Among women who were at the screening age, 41% of breast cancers were detected by screening (Study II and III). Due to some differences in the study exclusion criteria this percentage was slightly lower in Study I (37%). When women of any age were included, the proportion of screen-detected tumors was 22% (Studies II and III) or 20% (Study I, again depending on the study inclusion criteria).

Screen-detected breast cancers were smaller as compared to non-screen-detected breast cancers. Of all screen-detected breast cancers 80-83% belonged to the pT1 category (≤20mm in diameter), whereas 52-54% of the non-screen-detected breast cancers belonged to this category. Screen-detected tumors were more often node-negative as compared to non-screen-detected tumors (77-78% vs. 58-62%, p<0.001).

There were slight differences in the histological types between screen-detected and non-screen-detected subsets. The majority of screen-detected tumors were ductal carcinomas (70-75%) while the proportion of ductal carcinomas in the non-screen-detected group was (75-77%). Screen-detected cancers were more often of one of the special histological types (13-14% vs. 9%, p<0.01) as compared to non-screen-detected tumors (Studies I-III). However, there was no significant association in the frequency of lobular carcinomas according to the method of detection (11-16% vs.14-16%). Similar differences in cancer histological types were seen also in the 50- to 69-year-old age cohort (Studies II-III).

Of all screen-detected tumors 31-35% was classified as grade 1 tumors, while in the subset of non-screen-detected tumors the proportion of grade 1 cancers was 16% (Studies II and III) and 21% (Study I). This difference was highly significant in all of these analyses (p<0.001). The association with the method of detection and tumor grade, however, weakened with increasing tumor size (Study III).

The recommended threshold for hormone receptor positivity was lowered from ≥10% to ≥1% in 2010. The differences in the association between the method of detection and hormone receptor status between the may be explained by this change. There were significantly lower proportion of ER-negative tumors in the screen-detected group compared with non-screen-detected group in women at screening age (9% vs. 19%, p=0.002) when lowered threshold was used (Study III). Similar differences were seen also according to PR status (p<0.001). The association between the
method of detection and hormone receptor status was weaker when a higher threshold was used (Studies I-II).

The proportion of HER2-positive tumors did not differ significantly between the screen-detected and non-screen-detected cohorts regardless of the method of HER2 assessment (either IHC or CISH, Studies I-III).

Descriptive statistics of FinProg study with main exclusions (zero follow-up time, wrong diagnosis, distant metastases at diagnosis, not operated, DCIS and LCIS) is shown in table 8.
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n (%)</th>
<th>n(screen-detected)</th>
<th>n(non-screen-detected)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total*</td>
<td>2,883</td>
<td>607</td>
<td>2,146</td>
</tr>
<tr>
<td>Main exclusion criteria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distant metastases at primary dg</td>
<td>131 (4,5)</td>
<td>7 (1,2)</td>
<td>119 (5,5)</td>
</tr>
<tr>
<td>not operated</td>
<td>101 (3,5)</td>
<td>3 (0,5)</td>
<td>48 (2,2)</td>
</tr>
<tr>
<td>DCIS</td>
<td>191 (6,6)</td>
<td>82 (13,5)</td>
<td>100 (4,7)</td>
</tr>
<tr>
<td>LCIS</td>
<td>18 (0,6)</td>
<td>9 (1,5)</td>
<td>7 (0,3)</td>
</tr>
<tr>
<td>Total (after main exclusions)</td>
<td><strong>2,449</strong></td>
<td><strong>507</strong></td>
<td><strong>1,876</strong></td>
</tr>
<tr>
<td>Age at diagnosis (years)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 39</td>
<td>142 (5,8)</td>
<td>4 (0,8)</td>
<td>137 (7,3)</td>
</tr>
<tr>
<td>40-49</td>
<td>507 (20,7)</td>
<td>55 (10,8)</td>
<td>441 (23,5)</td>
</tr>
<tr>
<td>50-59</td>
<td>618 (25,2)</td>
<td>315 (62,1)</td>
<td>299 (15,4)</td>
</tr>
<tr>
<td>60-69</td>
<td>496 (20,3)</td>
<td>116 (22,9)</td>
<td>366 (19,5)</td>
</tr>
<tr>
<td>≥70</td>
<td>684 (27,9)</td>
<td>17 (3,4)</td>
<td>643 (34,3)</td>
</tr>
<tr>
<td>Tumor size (mm)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤10</td>
<td>463 (18,9)</td>
<td>195 (38,5)</td>
<td>265 (14,1)</td>
</tr>
<tr>
<td>11-20</td>
<td>989 (40,4)</td>
<td>221 (43,6)</td>
<td>753 (40,1)</td>
</tr>
<tr>
<td>21-50</td>
<td>796 (32,5)</td>
<td>72 (14,2)</td>
<td>710 (37,8)</td>
</tr>
<tr>
<td>&gt; 50</td>
<td>90 (3,7)</td>
<td>2 (0,4)</td>
<td>85 (4,5)</td>
</tr>
<tr>
<td>NA</td>
<td>112 (4,6)</td>
<td>17 (3,4)</td>
<td>93 (3,4)</td>
</tr>
<tr>
<td>Nodal status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>negative</td>
<td>1,547 (63,2)</td>
<td>396 (78,5)</td>
<td>1,151 (59,5)</td>
</tr>
<tr>
<td>positive</td>
<td>811 (33,1)</td>
<td>103 (20,3)</td>
<td>688 (36,6)</td>
</tr>
<tr>
<td>NA</td>
<td>91 (3,7)</td>
<td>6 (1,2)</td>
<td>74 (3,9)</td>
</tr>
<tr>
<td>Histological type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ductal</td>
<td>1,799 (73,5)</td>
<td>358 (70,6)</td>
<td>1,388 (74,0)</td>
</tr>
<tr>
<td>lobular</td>
<td>407 (16,6)</td>
<td>80 (15,8)</td>
<td>327 (16,9)</td>
</tr>
<tr>
<td>special</td>
<td>241 (9,8)</td>
<td>69 (13,6)</td>
<td>169 (9,0)</td>
</tr>
<tr>
<td>NA</td>
<td>2 (0,1)</td>
<td>2 (0,1)</td>
<td>2 (0,1)</td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grade I</td>
<td>471 (19,2)</td>
<td>159 (31,2)</td>
<td>312 (16,4)</td>
</tr>
<tr>
<td>grade II</td>
<td>873 (35,6)</td>
<td>182 (35,9)</td>
<td>683 (36,5)</td>
</tr>
<tr>
<td>grade III</td>
<td>465 (19,0)</td>
<td>67 (13,2)</td>
<td>388 (20,7)</td>
</tr>
<tr>
<td>NA</td>
<td>640 (26,1)</td>
<td>100 (19,7)</td>
<td>512 (27,3)</td>
</tr>
<tr>
<td>Estrogen receptor status (TMA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>positive</td>
<td>1,189 (48,6)</td>
<td>238 (46,9)</td>
<td>924 (49,3)</td>
</tr>
<tr>
<td>negative</td>
<td>521 (21,3)</td>
<td>83 (16,4)</td>
<td>426 (22,7)</td>
</tr>
<tr>
<td>NA</td>
<td>739 (30,2)</td>
<td>196 (36,7)</td>
<td>526 (28,0)</td>
</tr>
<tr>
<td>Progesterone receptor status (TMA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>positive</td>
<td>941 (38,4)</td>
<td>186 (36,7)</td>
<td>742 (39,6)</td>
</tr>
<tr>
<td>negative</td>
<td>761 (31,1)</td>
<td>129 (25,4)</td>
<td>631 (32,7)</td>
</tr>
<tr>
<td>NA</td>
<td>747 (30,5)</td>
<td>192 (37,9)</td>
<td>521 (27,8)</td>
</tr>
<tr>
<td>HER2 amplification (CISH, TMA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>positive</td>
<td>291 (11,9)</td>
<td>48 (9,5)</td>
<td>236 (12,6)</td>
</tr>
<tr>
<td>negative</td>
<td>1,387 (56,6)</td>
<td>289 (55,2)</td>
<td>1,083 (57,6)</td>
</tr>
<tr>
<td>NA</td>
<td>771 (31,5)</td>
<td>179 (35,3)</td>
<td>590 (29,9)</td>
</tr>
<tr>
<td>p53 expression (TMA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>269 (11,0)</td>
<td>44 (8,7)</td>
<td>225 (11,7)</td>
</tr>
<tr>
<td>negative-low</td>
<td>1,266 (51,8)</td>
<td>228 (45,0)</td>
<td>1,035 (54,0)</td>
</tr>
<tr>
<td>NA</td>
<td>912 (37,2)</td>
<td>235 (46,4)</td>
<td>643 (34,3)</td>
</tr>
<tr>
<td>Ki-67 expression (TMA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low-moderate</td>
<td>996 (40,6)</td>
<td>197 (38,9)</td>
<td>778 (41,5)</td>
</tr>
<tr>
<td>high</td>
<td>566 (23,2)</td>
<td>93 (18,3)</td>
<td>462 (24,6)</td>
</tr>
<tr>
<td>NA</td>
<td>886 (36,2)</td>
<td>217 (42,8)</td>
<td>636 (33,9)</td>
</tr>
<tr>
<td>Adjuvant systemic therapy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>not given</td>
<td>1,477 (60,3)</td>
<td>390 (78,9)</td>
<td>1,087 (56,7)</td>
</tr>
<tr>
<td>given</td>
<td>894 (36,5)</td>
<td>109 (21,3)</td>
<td>761 (40,6)</td>
</tr>
<tr>
<td>NA</td>
<td>78 (3,2)</td>
<td>9 (1,8)</td>
<td>51 (2,7)</td>
</tr>
</tbody>
</table>

(The percentages may not equal to 100% due to rounding)
NA=not available
**=those with wrong diagnosis (n=46) or zero follow-up (n=14) were excluded
**Age at diagnosis: mean 60.5 years, range 23.3-96.2 years
***Tumor size: mean 23.5mm, range 1-180mm

Table 8. Clinicopathological characteristics of 2,883 women in FinProg database.
9.2 Molecular subtypes in screen-detected breast cancer (Study I)

The purpose of the Study I was to explore differences in the molecular profiles in screen-detected cancers compared to non-screen-detected cancer. Luminal A (ER+ and/or PR+, HER2-) subtype was more frequent among screen-detected tumors than in non-screen-detected tumors (73% vs. 64%) in the age group of 50-69 years (p =0.02). In line with this observation, HER2+/ER- tumors were not as frequent among screen-detected cancers compared to non-screened tumors (5% vs. 13%, p =0.049). The distribution of the molecular subtypes of cancers found in mammography screening in the age group of 50 to 69 was similar to that of non-screen-detected cancers among patients aged ≥70 years (p=0.88).

9.3 Prognosis of patients with screen-detected breast cancer (Studies I, II, III)

In Studies I and II, DDFS was used as the outcome measure. The median follow-up time was 9.4 years in Study I and 9.5 years in Study II. BCSS was the outcome measure in Study III and the median follow-up time was 15.4 years.

Ten-year DDFS was more favorable for patients with screen-detected cancer than those with breast cancer detected by other methods (p <0.0001) (Studies I-II). The survival advantage associated with detection in screening persisted also during a median of 15 years of follow-up (Study III). The 15-year breast cancer-specific survival of patients with screen-detected breast cancer was 86%, whereas it was only 66% among patients with non-screen-detected breast cancer (p <0.0001, RR =2.91).

Women with screen-detected breast cancer had more favorable survival in each tumor size category investigated (0-10mm, 11-20mm, 21-30mm and >30mm in Study II, 0-10mm, 11-20mm, 21-50mm and >50mm in Study III). However, such survival comparisons could not be made for patients with a very large tumor (> 5 cm in diameter), because only few such tumors were detected in screening (Study II-III).

In node-negative screen-detected breast cancer the survival advantage associated with screen-detection was most obvious among patients with a small tumor (≤ 2cm)(Study II-III). The 15-year breast cancer specific survival for T1N0 cancer in the age group of 50 to 69 years was 91.1% in the screen-detected group as compared to 82.4% in the non-screened group (p =0.0003, RR =2.05; Study III).

The association between screen-detection and favorable outcome was observed also in patients with node-positive disease. Seventy-one percent of the patients with screen-detected cancer with
nodal involvement survived for 15-years follow-up compared to 51% of the patients with node-positive breast cancer detected outside of the screening program (p =0.0011, RR =1.95) (Study III).

Of the different molecular subtypes, luminal A tumors were associated with the best prognosis, whereas patients with HER2+/ER- cancer had the worst outcome. Patients with screen-detected breast cancer of the luminal A subtype had significantly better outcome than patients with non-screened breast cancer of the same subtype (10y DDFS; 90% vs. 75%, respectively, p<0.0001). A similar difference in patient outcome was observed also in the HER2+/ER- subtype (79% vs. 46%, p =0.04) (Study I).

Because there were several differences between screen-detected and non-screen-detected breast cancers in univariate survival analyses, we performed a Cox multivariate survival analysis to find out whether screen-detection has an independent prognostic value in risk estimations. We included all variables that were significantly associated as single factors with prognosis into the model. In Study I, the molecular subtype, tumor size, the number of positive nodes, histological grade and age at diagnosis were included in the multivariate analyses besides the method of detection. Cancer detection in mammography screening remained as an independent prognostic factor in a multivariate analysis (HR 1.79, p =0.011) along with tumor size, the nodal status and grade. In Study II, tumor size, the number of positive nodes, tumor PR expression, histological grade, presence of HER2 gene amplification and age at diagnosis were included in the multivariate model. The detection of cancer at screening was an independent prognostic factor with a relative hazard of 1.90 (p=0.01) for cancer detection outside of screening. In Study III, the same covariables were included in the model. The hazard ratio for cancer detection outside of screening was 1.69 (p =0.028), and the method of detection remained as a favorable independent prognostic factor also during a median of 15 years of follow-up. The risk for contralateral breast cancer was not associated with the method of detection.

9.4 Comparison of the FinProg and SEER cohorts (Study IV)

In Study IV, overall survival, breast cancer-specific survival and the relative survival were used as outcome measures. The median follow-up time was 9.5 years for the FinProg and 9.7 years for the SEER series.

Age at diagnosis differed significantly between the two cohorts. In the US cohort, breast cancer was diagnosed at the age of 65 or over in 43% of the cases, whereas in the Finnish cohort only 36% of the patients were in this age group (p <0.0001). Correspondingly, a larger proportion of the patients were aged 50-64 years in the FinProg dataset (36%) as compared with the SEER data (30%). These differences may be in part due to differences in the screening policies. Screening was offered more often for older women in the US as compared to Finland, where screening was offered mainly for 50-59 year old women in 1991-92.
The proportions of node-negative and node-positive breast cancer patients was similar in the two datasets (p =0.70). Hormone receptor positivity was more frequent in the SEER dataset than in the FinProg database (p <0.0001).

Forty-six percent of the breast cancers in the SEER database were poorly differentiated (grade 3 or 4) as compared with 27% (grade 3) in the FinProg database. There were only 12 % grade 1 cancers in the SEER database as compared with 26% in the FinProg series (p<0.0001). The distribution of the clinical variables in the FinProg and SEER series are shown in Table 1 in the Study IV article.

Despite these differences in variables between two cohorts, the survival curves examined closely overlapped and were of similar shape in both series. Yet, survival was slightly better in the SEER series in some subgroups as compared to the FinProg series. The survival difference between the two datasets was largest in women aged 65 or over, and was significant also in the subset of patients with a NPI over 5.40, grade 2 or 3 and for patients with the St Gallen criteria classified as intermediate to high.

Tumor size, the number of examined axillary lymph nodes, the number of positive axillary lymph nodes, histologic grade, tumor ER and PR status, age at diagnosis and the series (the SEER or FinProg) were included in a multivariate model. All of the listed variables were independent prognostic factors except for the series suggesting that the risk estimations based on commonly used prognostic factors are transportable across these populations and that the small differences between the SEER and the FinProg series in outcome can be explained by different distributions of common prognostic factors.
10 DISCUSSION

10.1 General remarks

Breast cancer is the leading cause of cancer deaths in women worldwide accounting for 14% of all cancer deaths in 2008 (Jemal et al. 2011). Mammography screening programs have been introduced to detect breast cancers at an early stage, before cancer dissemination and to reduce breast cancer mortality. The screening programs have been successful in finding more early stage tumors, but at the same time the incidence has increased. The breast cancer mortality reduction related to mammography screening in postmenopausal women has been estimated to be 20-50% (Kerlikowske et al. 1995)(van Schoor et al. 2011)(Tabar et al. 2011)(Kalager et al. 2010).

On average, screen-detected breast cancers have a more favorable tumor biological profile than non-screen-detected breast cancers. It has, however, been unclear whether the survival differences can be explained by other factors related to the extent of disease and prognosis. Furthermore, little has been known about the long-term survival of breast cancer patients detected at mammography screening as compared to those detected outside of screening. According to the current results, screen-detection is a favorable prognostic factor, independent of several established prognostic factors, and the survival advantage related to breast cancer detection at screening persists over a long follow-up time.

There was no significant difference in the risk for contralateral breast cancer according to the method of detection. The risk for contralateral breast cancer is higher in patients with the first breast carcinoma of the lobular type. Lobular carcinomas may be difficult to detect by mammography, (Ma et al. 1992)(Li et al. 2003) which may at least partly explain the result. The proportion of lobular carcinoma in patients with screen-detected cancer was equal to the patients with non-screen-detected cancer in the current study. Further studies are needed to investigate other possible explanations for the similar risk for contralateral breast cancer in both screen-detected and non-screen-detected breast cancers.

10.2 Strengths and limitations of the study

The Finnish Cancer Registry has collected information about virtually all cancer cases diagnosed in Finland since 1953. In this study, outcome information of patients was up-dated from the files of the Finnish Cancer Registry as well as from the hospital case records, and data about the
Various definitions of screen-detection have been used in studies that compared survival of patients with screen-detected cancer to that of women with non-screen-detected breast cancer. A Finnish study (Sarkeala, Heinavaara & Anttila 2008b) investigated mortality reduction from breast cancer both among the invitees and true attendees. Mortality reduction among the invitees was only 22% as compared to 28% reduction among the true attendees. Hence, by comparing screening invitees to non-invitees, the potential survival advantage of screen-detection may be underestimated. Survival of patients with interval cancer is reported to be similar to that of non-screen-detected population, both groups having less favorable survival than patients with screen-detected cancers (Holmberg et al. 1986)(Burrell et al. 1996). In some previous studies, interval cancers have been included in the screen-detected group, since these patients participated in the screening program (Immonen-Raiha et al. 2005). Although including interval cancers into the screen-detected group gives information about the efficacy of screening program on the whole, survival outcomes associated with screen-detected cancers will be confounded. Our study compared the survival of screening attendees with non-attendees giving more accurate estimations of survival associated with patient’s whose cancer were screen-detected. On the other hand, few women in the non-screened group may have undergone mammography without symptoms or signs of breast cancers, but the great majority had either detected a breast lump or had breast symptoms, so this confounding effect (opportunistic screening) was small in the present series.

Screening rounds were not considered in this study. The association between the screening round and tumor aggressiveness remains unclear. Cancers detected at the first round of screening have been reported to have a low malignant potential, the tumors generally being small and often node-negative (Hakama et al. 1995). Cancers detected at subsequent rounds were of a similar size and had similar nodal status as the interval cancers. According to another study (Klemi et al. 2003) significant differences in the prognostic variables of cancers detected were not found between the screening rounds but size of the series was limited. In our study women with breast cancer detected at the first screening round had similar outcomes to women with breast cancer detected at the subsequent rounds, although the data on screening rounds was incomplete.

The actual benefit of mammography screening and the balance between adverse effects and true benefits are under continuous debate, most of the criticism stem from the biases associated with screening.

In order to reduce the selection bias, we chose to investigate the nationwide FinProg database which represents 53% of women with breast cancer patients diagnosed in Finland during 1991-92, and the great majority of all breast cancers diagnosed within the selected regions. The attendance rate to mammography screening is high in Finland, 89% at the time of this study (Dean, Pamilo 1999)(Sarkeala et al. 2004) which is high as compared with many other countries (Saalasti-
Koskinen et al. 2009)(Giordano et al. 2012). These factors, in part, reduce the risk of a major selection bias as compared with surveys from countries with a lower attendance rate or with a less comprehensive study cohort.

In 1991-1992 most breast cancer patients were treated with relatively similar adjuvant therapies. The therapies administered may cause some bias, but at the time of the study only a minority of the patients received systemic adjuvant therapies, and they were likely administered similarly regardless of the mode of detection of cancer.

Breast cancers found in mammography screening are detected earlier in their natural course than non-screen-detected cancers (the lead time bias). The adjustments for known prognostic factors and long term follow-up reduce the effect of lead time bias. The available models for lead time correction are based on approximation of the rate of transition of an asymptomatic but mammography detectable to a symptomatic cancer. In all analyses with this study, in situ carcinomas were excluded and adjustments were made according to cancer size, the nodal status and stage in order to reduce the effect of the lead time bias and the stage shift. The long-term follow-up time also partly reduces the effects of the lead time.

Mammography screening reveals more in situ carcinomas than would be diagnosed otherwise (Malmgren, Atwood & Kaplan 2008). Even though untreated ductal carcinoma in situ (DCIS) may develop into invasive carcinoma, DCIS is associated with excellent prognosis. The incidence of breast cancer in women who are at screening age has increased, although early detection of many in situ cancers at screening is expected to lead to a decrease in the incidence of invasive cancers. It could be assumed that with the introduction of a screening program breast cancer incidence among elderly women should also decrease. However, the incidence of breast cancer among the elderly has not decreased as predicted in the screened populations (Zahl, Maehlen & Welch 2008). This raises a question about overdiagnosis caused by mammography screening programs. Part of overdiagnosis could be attributed to the detection of in situ cancers (Puliti et al. 2012). To reduce the potential bias related to overdiagnosis in the current study we included only patients with invasive cancer, and we also excluded cancers detected at autopsy (Moody-Ayers, Wells & Feinstein 2000). Whether there has been a change in the incidence of cancers detected at autopsy after the introduction of screening programs remains to be studied.

Although a wide selection of prognostic factors was observed, the information about e.g. lymphovascular or perineural invasion was not available in the current study. Lymphovascular and perineural invasion, however, could be associated with dissemination potential and earlier symptoms. Further studies on factors that would be capable of distinguishing indolent tumors from potentially invasive tumors are also needed.
10.3 Tumor biological profile

In general, screen-detected breast cancers are smaller, less often node-positive, more often of a special histological type, and more often well differentiated as compared to cancers found outside of screening (Palka et al. 2008)(Olsson et al. 2012)(Mook et al. 2011a)(Nagtegaal et al. 2011). In concordance with these previous studies, patients with screen-detected breast cancer had a small tumor size and had seldom nodal involvement in the current study.

Previous studies have reported that screen-detected tumors are less often ductal compared to non-screen-detected tumors (Klemi et al. 1992)(Chuwa et al. 2009), although in some studies the association was weak (Dawson et al. 2009)(Klemi et al. 2003). We found only a weak association between the method of detection and cancer histological type in the current series.

Small cancers are more often well differentiated, and a few previous studies have found a significant association between the method of detection and histological grade (Burrell et al. 1996)(Dong et al. 2008)(Crosier et al. 1999). In the current series the proportion of well differentiated tumors was significantly higher in the subset of small (≤10mm) screen-detected tumors than in non-screen-detected cancers of similar size, but this association weakened with increasing tumor size.

Screen-detected tumors have been reported to have a low mitotic activity, a small S-phase fraction and a low Ki-67 expression level (Porter et al. 1999)(Crosier et al. 1999)(Dawson et al. 2009)(Ernst et al. 2002)(Groenendijk et al. 2003). A favorable proliferation marker profile was also observed in the current series in screen-detected cancer.

The HER2 gene amplification and HER2 protein expression did not differ significantly between the two diagnostic groups, which is concordant with previous studies (Palka et al. 2008)(Dawson et al. 2009). Several studies report a significant association between the method of detection and the hormone receptor status (Dong et al. 2008)(Dong et al. 2008)(Dawson et al. 2009)(Nagtegaal et al. 2011) (Gill et al. 2004) although a lack of a difference has also been reported (Palka et al. 2008). No statistically significant association between the hormone receptor status (ER and PR) and screen-detection was observed in the current series. The most recent guidelines changed the cut-off point for positivity from ≥10% to ≥1% and thus increased the proportion of tumors to be classified as positive. We assessed the relationship between the method of detection and hormone receptor expression using both cut-offs, but found no significant associations.

A comparison of the molecular subtypes between screen-detected and non-screen-detected breast cancers showed that screen-detected tumors were more likely to be of the luminal A subtype, a finding which is supported by others (Dawson et al. 2009).
10.4 Prognosis

In the current study the survival advantage of patients with screen-detected tumor could not be explained by the patient characteristics or the tumor biological profile. This is in concordance with other studies that found favorable survival to be associated with screen-detected cancer even after adjustment for a series of known prognostic variables (Mook et al. 2011b)(Olsson et al. 2012).

In most previous studies, survival analyses were restricted to after 10 years from diagnosis or a shorter follow-up period. Breast cancer recurrences usually occur within the 10 years from the diagnosis, but can be observed even up to 15-25 years after diagnosis (Joensuu, Pylkkänen & Toikkanen 1999)(Joensuu, Toikkanen 1995). The lead time (time from screen-detection to the symptomatic stage) in breast cancer screening is estimated to be 3-4 years. Furthermore, the effect of some prognostic factors (ER, tumor size, the lymph node status and grade) are time-dependent and their prognostic efficacy may decrease with time (Warwick et al. 2004)(Anderson et al. 2006). Whether the favorable survival associated with screen-detection decreases with time has not been well known.

In Study II we estimated the 10-year DDFS in patients with screen-detected or non-screen-detected cancer. The survival benefit of screen-detection remained even after adjustments for several commonly used prognostic factors, and screen-detection emerged as an independent prognostic factor. In order to study the influence of screening on long-term survival we updated the dataset and analyzed 15-year breast cancer specific survival (Study III). Screening remained as an independent prognostic factor, which is in concordance with the only previous study with a long-term follow-up (Shen et al. 2005). This previous study, however, was based on relatively old data (the HIP study from the 1960’s with a median of 16 years follow-up, and the CNBSS1-2 study from the 1980’s with a median of 12 years follow-up) and did not have information available about biomarkers such as ER, HER2, p53 and Ki67.

Molecular subtypes are associated with the outcome of breast cancer patients. Women with luminal tumor have the best prognosis while the HER2+/ER- and triple negative breast cancers are associated with less favorable prognosis (Carey et al. 2006). This was verified in the current series, but the molecular subtypes could not explain the differences in survival associated with the method of tumor detection. Screen-detected luminal A and HER2+/ER- tumors were associated with better survival as compared with non-screen-detected tumors, which is in concordance with prior results (Dawson et al. 2009), and along with the multivariate analyses carried out suggests that the survival benefit associated with cancer screen-detection is not fully explained by the evaluated tumor biological factors.

Most of the biological factors examined had independent prognostic value, but the differences in these tumor-related factors could not explain the effect of the method of cancer detection. Screen-detection remained as a favorable independent prognostic factor in multivariate analyses. These results suggest that a significant length bias is associated with cancer detection at screening.
even when the effect of several common prognostic factors is accounted for in a multivariate analysis.

In the FinProg database patients with screen-detected breast cancer were given adjuvant systemic therapy less frequently than those with non-screen-detected cancer suggesting that the differences in treatment could not explain the survival advantage of screening as a method of detection.

10.5 Generalizability

To examine the representativeness of some of the current findings, we compared the clinicopathological and survival data between the FinProg series and another large dataset, the SEER series from the US. Survival was slightly better for the US breast cancer patients although the survival curves were of similar shape and generally closely overlapping. The results from a previous study comparing survival of breast cancer patients in the SEER series with a European breast cancer dataset (EUROCARE) (Sant et al. 2004) are in line with our results. Survival of the US patients was better also as compared to the EUROCare series with patients from six European countries (Sant et al. 2003).

When we compared the clinicopathological variables in the FinProg and the SEER series, there were differences which might have explained the observed survival differences. The age distribution was different between the series the proportion of patients aged 65 or older at the time of diagnosis being significantly higher in the SEER dataset than in the FinProg dataset. The screening policies are different in Europe, including Finland, and the US. In 1991-92, when women in the FinProg series were diagnosed, mammography screening was offered biennially mainly for women aged 50-59. Today, screening is offered for women aged 50-69 which is in line with the new recommendations by the USPSTF (US Preventive Services Task Force). In the US, elderly women are screened with mammography more frequently than European women, which might explain the greater proportion of early stage tumors in elderly women in the US. Further studies are needed to explore whether survival in Finland and the US might be even more convergent with this screening policy concerning a wider age range. Survival of elderly patients with early stage tumor is favorable and these differences in age and screening policies may explain in part of the survival differences. The differences in the adjuvant systemic treatments such as administration of tamoxifen more frequently to the U.S. patients with small, node-negative breast cancer might also influence survival. The effect of differences in the use of adjuvant therapies between countries remains to be studied. Interestingly, the survival advantage of ER-positivity weakened with time, which was also the case in the SEER series. The number of examined axillary lymph nodes was higher in the US in 1991-2 raising the question of more thorough search for occult metastases and a more precise definition of the stage at the time of the primary diagnosis. In general, a precise
definition of the stage of the disease is necessary for the therapy decision-making, but some recent data suggest that axillary dissection may offer no benefits over a sentinel-node biopsy followed by breast irradiation and systemic adjuvant treatment, and is associated with more harms (Giuliano et al. 2011). In situations where the investigations are less thorough, survival is lower indicating that a proportion of cancers that are misclassified as early stage tumors are actually more advanced tumors with lower survival rates. However, more intensive diagnostic activity in the US can also lead to earlier diagnosis without any true advantage for the patients. Women will be aware of their diseases earlier and will live longer with the diagnosis.

The SEER is a public-use database which collects clinicopathological information from patients with various types of cancer in the U.S. However, it collects cancer incidence and survival rates covering only approximately 28% of the U.S. population (www.seer.cancer.gov). The medical centers participating in the SEER Program may be self-selected group of institutions with more advanced diagnostic and treatment facilities than institutions not participating in the program.

The survival differences between the two series were smaller than differences associated with frequently use prognostic factors within each series indicating the generalizability and transportability between large cohorts of breast cancer patients.

10.6 Survival advantage associated with cancer detection at screening

The survival benefit of screen-detection was evident and could not be explained by differences in tumor stage and biology. Even when we adjusted for a series of known prognostic factors in multivariate analyses in order to minimize the length bias, part of the effect on survival was still attributable to unidentifiable factors associated with screening. We also stratified T1 tumors into smaller subcategories according to tumor size, but the differences in survival persisted in favor of the screen-detected groups. However, screen-detected breast cancers were still likely smaller than the cancers found by other means within each size category.

Norwegian researchers compared the breast cancer incidence in women before and after initiating mammography screening (Zahl, Maehlen & Welch 2008). After an observation period, all patients in the non-screened group were screened. After screening of the controls, breast cancer incidence still remained higher in the screening group than in the controls suggesting that sometimes breast cancer spontaneous regression may occur. To explore spontaneous regression as an explanation for the survival benefit of screen-detection, further studies are needed.

Even though it was evident in the current studies that screen-detected breast cancers are associated with more favorable outcomes than cancers detected outside of screening, this was not fully explained by the biological factors examined. Novel factors, such as factors associated with cancer invasiveness, metastasis, apoptosis and senescence need to be examined, as well as factors
linked with the host immune defense. The current results are compatible with a substantial length bias associated with breast cancer detection in mammography screening that needs to be taken out when considering the risk of breast cancer recurrence and when selecting adjuvant systemic treatments.
11 CONCLUSIONS

In the current thesis the influence of mammography screening on breast cancer prognosis was investigated. The associations between both clinicopathological and molecular factors and the method of cancer detection were studied. The generalizability and transportability of the results were also investigated by comparing two large breast cancer databases, one from Finland and another one from the USA. Based on the results presented in this thesis, the main conclusions are:

1. Breast cancers detected by mammography screening programs have a less aggressive biological profile as compared with cancers of roughly similar size found outside of screening.

2. Women with breast cancer diagnosed within a mammography screening program have better prognosis than women whose cancer has been detected by other means even after adjustments are made for several commonly used prognostic factors. This survival advantage is evident also during a long follow-up time.

3. Data of the method of breast cancer detection should be collected to improve individual risk estimations and prognostication, and accounted for in decision making process to avoid overtreatment of patients with screen-detected breast cancer.

4. Survival estimates based on frequently used prognostic factors are generalizable and transportable between large population-based series.
12 ACKNOWLEDGEMENTS

This thesis was carried out at the Department of Oncology, Helsinki University Central Hospital and at the Institute for Molecular Medicine Finland (FIMM) during the years 2003-2013. I wish to express my sincere gratitude for the Director of FIMM, Professor Olli-Pekka Kallioniemi for the opportunity to use excellent conditions for scientific work at FIMM. I also want to thank all the staff at FIMM for creating an exciting place to work.

I express my deepest and warmest appreciation to my supervisor Docent Johan Lundin for his excellent and patient and enduring guidance. Johan has given me valuable and constructive support during these years. Without his enthusiastic encouragement and support, I wouldn’t been able to complete this thesis.

I would like to express my very great thankfulness to my other supervisor, Academic Professor Heikki Joensuu. He has provided me valuable criticism and support as a supervisor and also as a director of the Department of Oncology at the Helsinki University Central Hospital.

I am deeply grateful for the valuable comments from Docent Arja Jukkola-Vuorinen and Professor Nea Malila, who as reviewers improved this thesis in many important respects. I also like to express my very great appreciation to Docent Minna Tanner who accepted the invitation to be the opponent. Special thanks should be given to Professor Jorma Isola, who introduced me the breast histopathology and tumor microarray technology. I would like to offer special recognition to Harri Sihto who taught me the principles of immunohistochemistry and professionally performed the stainings of FinProg TMA slides I had collected.

Without the innovative information technology advice from Mikael Lundin, we would not have the case-match prognostic tool which helps in comparing the survival estimations of breast cancer patients with different marker profiles for both clinicians and researcher. Micke has always been ready to help in any IT-related problems.

I wish to acknowledge the help provided by our group member, Nina Linder, who has helped me updating the database, scoring the markers and also given me valuable tips in several practical things. I had also many lovely scientific and non-scientific discussion with our group member, Juho Konsti, thanks Juho. Hélène Rotkirch has also provided help in tremendous work of database updating.

I express my very great appreciation to the co-authors Professor Kaija Holli, Professor Taina Turpeenniemi-Hujanen, Professor Vesa Kataja, Docent Ahti Anttila, Docent Ralf Bützow, Docent Päivi Heikkilä, MD Liisa Sillas and MD Maarit Sarlomo-Rikala.
I would like to extend my thanks to the laboratory staff and other collaborators in each University Hospitals in Finland, in Kuopio, Oulu, Tampere, Turku and Helsinki. I would also thank the assistance provided by the staff at the Finnish Cancer Registry and Statistics Finland.

I would like to express my great appreciation from the financial support which was provided by Biomedicum Helsinki Foundation, the Sigrid Juselius Foundation, the Academy of Finland, the Finnish Cancer Society, the Research Funds of Helsinki University Central Hospital, a Special Governmental Research Grant, the Yamanouchi European Foundation, Finska Läkaresällskapet, the Instrumentarium foundation for Science and Medicinska Understödfsöreningen Liv och Hälsa.

I also appreciate the help which I gave from my uncle Jyrki and my lovely cousin and goddaughter Nea. I also want to thank Marjatta Koskinen for the help with the cover of this thesis.

I also want to thank my parents-in-low Riitta and Mikko, who have adopted me as a full member to the Lehtimäki family.

I want also thank to my late father, Jaakko, who taught me to never give up. I also want to thank my sister Leena and her family for standing by at the good and bad days.

I am extremely thankful to my mother Riitta and stepfather Juha-Matti who have always supported me whatever I have been doing. You have always believed in me, encouraged me, loved me - and also taken care of our children during the most hectic moments.

Finally, my deepest gratitude goes to my loving husband Markus and our lovely (and lively) children, Eero, Laura, Lotta and Oskar – you are the sunshine of my life! I love you with all my heart.

Helsinki, November 2013

Tiina Lehtimäki
13 REFERENCES


Boyle, P. 2003, "Mammographic breast cancer screening: after the dust has settled", Breast (Edinburgh, Scotland), vol. 12, no. 6, pp. 351-356.


Ederer F, H.H. 1959, "Instructions to IBM 650 programmers in processing survival computations. Methodological note No. 10", *Bethesda, MD: End Results Evaluation Section, National Cancer Institute;1959*.


O'Malley, F.P., Pinder, S.E. 2006, Breast Pathology, First edn, Elsevier Inc.


Seal, M.D. & Chia, S.K. 2010, "What is the difference between triple-negative and basal breast cancers?", *Cancer journal (Sudbury, Mass.)*, vol. 16, no. 1, pp. 12-16.


www.cancerregistry.fi.


