



LUND UNIVERSITY

Prognostic factors in breast cancer with focus on proliferation and non-linear effects

Forsare, Carina

2016

[Link to publication](#)

Citation for published version (APA):

Forsare, C. (2016). *Prognostic factors in breast cancer with focus on proliferation and non-linear effects*. [Doctoral Thesis (compilation), Department of Clinical Sciences, Lund]. Lund University: Faculty of Medicine.

Total number of authors:

1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Prognostic factors in breast cancer with focus on proliferation and non-linear effects

Carina Forsare



LUND
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at the lecture hall in the Radiotherapy building, Skåne Oncology
Clinic, Lund, Sweden, Friday the 11th of November 2016, at 9:00 am

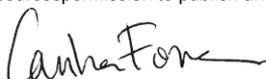
Faculty opponent

Professor Björn Naume

Department of Cancer Medicine, Oslo University Hospital, Oslo, Norway

Organization LUND UNIVERSITY	Document name DOCTORAL DISSERTATION	
Department of Clinical Sciences, Lund Oncology and Pathology	Date of issue November 11 th 2016	
Author(s) Carina Forsare	Sponsoring organization	
Title and subtitle Prognostic factors in breast cancer with focus on proliferation and non-linear effects		
<p>Abstract</p> <p>Breast cancer is the most common cancer today among women in the Western world. In Sweden in 2015, more than 8 000 women were diagnosed with breast cancer. For a large number of patients, the prognosis is good. The majority of the patients are cured by primary surgical treatment, and many of the recurring patients survive for long periods of time. Additional (adjuvant) treatment is administered to prevent recurrences, although many patients are already cured by the primary surgery, thereby leading to over-treatment. Despite adjuvant treatment, more than 1 400 patients die from breast cancer every year, demonstrating the need for tools to better tailor treatment. The most important and routinely used prognostic and treatment predictive factors are age at diagnosis, tumor size, spread to the lymph nodes, histological grade, hormone receptors, factors for proliferation and expression of a growth factor receptor (HER2).</p> <p>The overall aim of this thesis was to investigate prognostic and predictive markers in breast cancer and to find ways to improve the use of established factors, with a focus on proliferation and non-linear effects. Factors for proliferation were studied in different breast cancer patient cohorts and with different methods.</p> <p>In Study I, we investigated whether an index, CAGE, based on cyclin A, histological grade, and estrogen receptor, could provide valuable prognostic information. CAGE was evaluated on tissue microarray slides, with samples from 219 premenopausal node-negative breast cancer patients. CAGE together with HER2 identified 53% of the patients as low risk with a 5-year distant disease-free survival of 95%. In Study II, we investigated whether the results of Study I could be confirmed in independent and larger series, replacing cyclin A with the worldwide used Ki67, creating KiGE. In chemo-naïve N0/N1 patients, KiGE alone identified 57% of the patients as low risk, with a 5-year event-free survival of 92%.</p> <p>In Study III, we studied five factors for proliferation (mitotic activity index (MAI), phosphorylated histone H3 (PPH3), cyclin B1, cyclin A, and Ki67), separately and in combinations. Since all of the prognostic factors have pros and cons, our hypothesis was that combining factors would circumvent these issues. We demonstrated that a combination of MAI and cyclin A improved prognostication compared to use of the factors individually.</p> <p>Although it is a well-known fact that information is lost when predictors are dichotomized (divided into two groups) or categorized into three groups, for clinical decision-making, this is often performed. In Study IV, we investigated whether the prognostication could be improved by postponing dichotomization/categorization until the last step in the process. We concluded that dichotomization definitely leads to information loss and should be avoided. Categorization improved prognostication, whereas using non-linear transformations only moderately improved the predictions.</p> <p>In conclusion, the studies included in this thesis demonstrated that several markers for proliferation could be used and that combinations seems to improve prognostication, compared to examining factors individually. With these combinations, it was possible to identify patients with a low risk of developing recurrence so that adjuvant chemotherapy might be avoided. Dividing prognostic factors in more groups than two gives better predictive performance, but keeping them continuous was only moderately better. However, by postponing categorization until the very last step of the prognostic modeling strategy, more possibilities for individual predictions were enabled.</p>		
Key words breast cancer, proliferation, prognosis, non-linear, validation		
Classification system and/or index terms (if any)		
Supplementary bibliographical information	Language English	
ISSN and key title 1652-8220	ISBN 978-91-7619-348-8	
Recipient's notes	Number of pages 98	Price
	Security classification	

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature: 

Date: 2016-09-27

Prognostic factors in breast cancer with focus on proliferation and non-linear effects

Carina Forsare



LUND
UNIVERSITY

Lund University
Faculty of Medicine, Department of Clinical Sciences, Lund
Oncology and Pathology

© Eva Carina Forsare

Coverphoto by Friedensreich Hundertwasser (1928-2000).

Hundertwasser

715 Hommage to Schröder-Sonnenstern, silk screen in 8 colours, 1972

© 2016 Hundertwasser Archive, Vienna

Lund University, Faculty of Medicine, Department of Clinical Sciences Lund,
Oncology and Pathology

ISBN 978-91-7619-348-8

ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University, Lund 2016



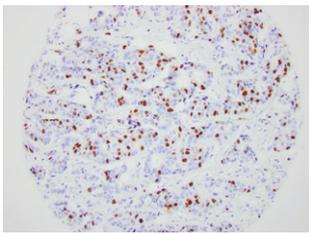
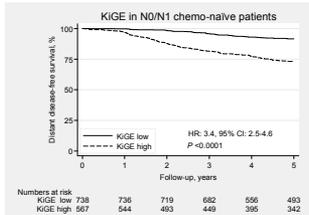
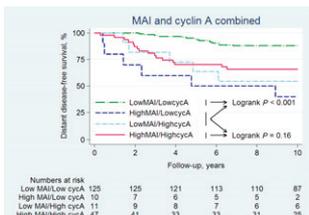
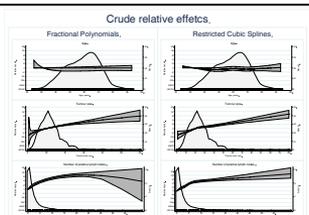
Climb mountains and reach new heights

Content

Thesis at a glance	9
Included studies	11
Abbreviations	13
Abstract	15
Introduction	17
Breast cancer	21
General background	21
Diagnostics	22
Prognostic and treatment predictive factors in breast cancer	22
Proliferation	23
Markers of proliferation	23
Other prognostic and treatment predictive factors in breast cancer	27
TNM classification	27
Age at diagnosis	28
Histological grade	28
Estrogen and progesterone receptors: ER and PgR	28
Human epidermal growth factor receptor 2	29
Molecular profiling	29
Other promising prognostic factors	30
Treatment	31
Surgery	31
Radiotherapy	31
Endocrine therapy	32
Chemotherapy	32
Anti-HER2 therapy	33
Positive and negative effects of adjuvant treatment	33
Background to the thesis	34
Aims of the thesis	35
Overall aim	35
Study I	35
Study II	35
Study III	36
Study IV	36

Patients	37
Study I and III	37
Study II	38
Study IV	39
Methods	41
Tissue microarray	41
Immunohistochemistry	42
Antibodies used in this thesis	43
Statistical methods and concepts	45
Statistics used but not extensively described	45
The Kaplan-Meier method	45
The log-rank test	46
Cox regression	46
Schoenfeld's test	46
Stratification in Cox regression models	46
<i>P</i> -value	47
<i>C</i> -index	47
Fractional polynomial	47
Multivariable FP	48
Restricted cubic splines	48
Endpoints	48
Results	51
Study I	51
Study II	51
Study III	52
Study IV	53
Discussion	55
Conclusions	65
Future perspectives	67
Populärvetenskaplig sammanfattning	69
Acknowledgments	73
References	75

Thesis at a glance

Study	Question	Patients and Methods	Figure	Results
I	Could an index, CAGE, based on cyclin A, histological grade, and estrogen receptor (ER), provide valuable prognostic information?	CAGE was evaluated on tissue microarray (TMA)-slides, with samples from 219 premenopausal, node-negative breast cancer patients. Five-year distant disease-free survival (DDFS) was used as the endpoint.		CAGE together with HER2 identified 53% of the patients as low risk with a 5-year DDFS of 95% (95% CI: 89–98%).
II	In independent and larger patient series, could the results of Study I be confirmed by replacing cyclin A with Ki67 and creating KiGE?	KiGE was evaluated on TMA slides with samples from 1 305 chemo-naïve N0/N1 breast cancer patients. Five-year event-free survival was used as the endpoint.		The results from Study I were confirmed. KiGE identified 57% of the patients as low risk with a 5-year event-free survival of 92% (95% CI: 89–93%).
III	Would the use of a combination of factors for proliferation, instead of using them individually, improve prognostication?	Mitotic activity index (MAI), phosphorylated histone H3 (PPH3), cyclin B, cyclin A, and Ki67 were evaluated in premenopausal, node-negative breast cancer patients. Ten-year DDFS was used as the endpoint.		MAI together with cyclin A showed stronger prognostic value when combined than when each proliferation factor was used individually.
IV	Would prognostication be improved by keeping predictors continuous as long as possible?	Age at diagnosis, tumor size, and number of positive lymph nodes were evaluated in models based on dichotomized, categorized, and continuous predictors.		Categorization improved prognostication, whereas using fractional polynomial and restricted cubic spline transformations only moderately improved the predictions.

Included studies

Study I. Combination of the proliferation marker cyclin A, histological grade, and estrogen receptor status in a new variable with high prognostic impact in breast cancer

Carina Strand, Cecila Ahlin, Pär-Ola Bendahl, Marie-Louise Fjällskog, Ingrid Hedenfalk, Per Malmström, Mårten Fernö, *Breast Cancer Res Treat*, 131(1):33–40 2012

Study II. The combination of Ki67, histological grade and estrogen receptor status identifies a low risk group among 1,853 chemo-naïve women with N0/N1 primary breast cancer

Carina Strand, Martin Bak, Signe Borgquist, Gunilla Chebil, Anna-Karin Falck, Marie-Louise Fjällskog, Dorthe Grabau, Ingrid Hedenfalk, Karin Jirström, Marie Klintman, Per Malmström, Hans Olsson, Lisa Rydén, Olle Stål, Pär-Ola Bendahl, Mårten Fernö, *SpringerPlus* 2(1):111 2013

Study III. The prognostic value of mitotic activity index (MAI), phosphohistone H3 (PPH3), cyclin B1, cyclin A, and Ki67, alone and in combinations, in node-negative premenopausal breast cancer.

Marie Klintman M, **Carina Strand**, Cecila Ahlin, Sanda Beglerbegovic, Marie-Louise Fjällskog, Dorthe Grabau, Einar Gudlaugsson, Emiel Janssen, Kristina Lövgren, Ivar Skaland, Pär-Ola Bendahl, Per Malmström, Jan Baak, Mårten Fernö, *PloS one* 8(12):e81902 2013

Study IV. Will using age, tumor size, and lymph node status as continuous predictors improve prognostication in breast cancer?

Carina Forsare, Martin Bak, Anna-Karin Falck, Dorthe Grabau†, Fredrika Killander, Per Malmström, Lisa Rydén, Olle Stål, Marie Sundqvist, Pär-Ola Bendahl*, Mårten Fernö* (*shared senior authorship), Manuscript

† Deceased December 2015

Papers reprinted with permission by the publishers.

Studies not included but of relevance for the thesis:

Training artificial neural networks directly on the concordance index for censored data using genetic algorithms

Jonas Kalderstam, Patrik Eden, Pär-Ola Bendahl, **Carina Strand**, Mårten Fernö, Mattias Ohlsson, *Artificial intelligence in medicine* 58(2):125-132 2013

Histological grade provides significant prognostic information in addition to breast cancer subtypes defined according to St Gallen 2013

Anna Ehinger, Per Malmström, Pär-Ola Bendahl, Christopher Elston, Anna-Karin Falck, **Carina Forsare**, Dorthe Grabau, Lisa Rydén, Olle Stål, Mårten Fernö, *Acta Oncologica*, In press

Abbreviations

AI	aromatase inhibitor
ANN	artificial neural network
ASCO	American Association of Clinical Oncology
BCS	breast conserving surgery
BCSS	breast cancer-specific survival
CAGE-index	cyclin A, grade and ER
CDK	cyclin dependent kinase
CI	confidence interval
CISH	chromogenic <i>in situ</i> hybridization
CTCs	circulating tumor cells
ctDNA	circulating tumor DNA
DAB	diaminobenzidine
DIA	digital image analysis
DDFS	distant disease-free survival
DMFS	distant metastasis-free survival
ER	estrogen receptor
FAC	5-fluorouracil, doxorubicin (Adriamycin), and cyclophosphamide
FEC	5-fluorouracil, epirubicin, and cyclophosphamide
FISH	fluorescence <i>in situ</i> hybridization
FOVs	fields of vision
FP	fractional polynomial
HER2	human epidermal growth factor receptor 2
HR	hazard ratio

IHC	immunohistochemistry
KiGE-index	Ki67, grade and ER
MAI	mitotic activity index
MFP	multivariable fractional polynomial
N0	lymph node negative
N1	1–3 positive lymph nodes
PAI-1	plasminogen activator inhibitor type 1
PgR	progesterone receptor
PI	prognostic index
PPH3	phosphorylated histone H3
RCS	restricted cubic spline
SERM	selective estrogen receptor modulator
SISH	silver enhanced <i>in situ</i> hybridization
SPF	S-phase fraction
TAM	tamoxifen
TMA	tissue microarray
TNM	tumor node metastasis
uPA	urokinase plasminogen activator
WHO	World Health Organization

Abstract

Breast cancer is the most common cancer today among women in the Western world. In Sweden in 2015, more than 8 000 women were diagnosed with breast cancer. For a large number of patients, the prognosis is good. The majority of the patients are cured by primary surgical treatment, and many of the recurring patients survive for long periods of time. Additional (adjuvant) treatment is administered to prevent recurrences, although many patients are already cured by the primary surgery, thereby leading to over-treatment. Despite adjuvant treatment, more than 1 400 patients die from breast cancer every year, demonstrating the need for tools to better tailor treatment. The most important and routinely used prognostic and treatment predictive factors are age at diagnosis, tumor size, spread to the lymph nodes, histological grade, hormone receptors, factors for proliferation and expression of a growth factor receptor (HER2).

The overall aim of this thesis was to investigate prognostic and predictive markers in breast cancer and to find ways to improve the use of established factors, with a focus on proliferation and non-linear effects. Factors for proliferation were studied in different breast cancer patient cohorts and with different methods.

In *Study I*, we investigated whether an index, CAGE, based on cyclin A, histological grade, and estrogen receptor, could provide valuable prognostic information. CAGE was evaluated on tissue microarray slides, with samples from 219 premenopausal node-negative breast cancer patients. CAGE together with HER2 identified 53% of the patients as low risk with a 5-year distant disease-free survival of 95%. In *Study II*, we investigated whether the results of *Study I* could be confirmed in independent and larger series, replacing cyclin A with the worldwide used Ki67, creating KiGE. In chemo-naïve N0/N1 patients, KiGE alone identified 57% of the patients as low risk, with a 5-year event-free survival of 92%.

In *Study III*, we studied five factors for proliferation (mitotic activity index (MAI), phosphorylated histone H3 (PPH3), cyclin B1, cyclin A, and Ki67), separately and in combinations. Since all of the prognostic factors have pros and cons, our hypothesis was that combining factors would circumvent these issues. We demonstrated that a combination of MAI and cyclin A improved prognostication compared to use of the factors individually.

Although it is a well-known fact that information is lost when predictors are dichotomized (divided into two groups) or categorized into three groups, for clinical decision-making, this is often performed. In *Study IV*, we investigated whether the prognostication could be improved by postponing dichotomization/categorization until the last step in the process. We concluded that dichotomization definitely leads to information loss and should be avoided. Categorization improved prognostication, whereas using non-linear transformations only moderately improved the predictions.

In conclusion, the studies included in this thesis demonstrated that several markers for proliferation could be used and that combinations seems to improve prognostication, compared to examining factors individually. With these combinations, it was possible to identify patients with a low risk of developing recurrence so that adjuvant chemotherapy might be avoided. Dividing prognostic factors in more groups than two gives better predictive performance, but keeping them continuous was only moderately better. However, by postponing categorization until the very last step of the prognostic modeling strategy, more possibilities for individual predictions were enabled.

Introduction

The processes by which normal cells evolve into tumor cells follow a complicated route. As described by Hanahan and Weinberg in 2000, cells must to acquire certain abilities, i.e., hallmark capabilities, to become tumorigenic and ultimately malignant [1]. These hallmark capabilities include the following: (1) sustaining proliferation, (2) evading growth suppressors, (3) resisting cell death, (4) enabling replicative immortality, (5) inducing angiogenesis, and (6) activating invasion and metastasis. In 2011, these hallmarks were revised, and two more hallmark capabilities and two enabling characteristics were added. The two added hallmarks are deregulating cellular energetics and avoiding immune destruction, and the two enabling characteristics are tumor-promoting inflammation and genomic instability and mutation [2] (Figure 1).

It is believed that most cancers must acquire the same hallmark capabilities and that the two enabling characteristics contribute to driving them forward. First and foremost is the development of genomic instability, which generates random mutations (including chromosomal rearrangements). The second characteristic involves tumor-promoting inflammation of (pre)-malignant lesions, which in turn promotes tumor progression [2].

As previously mentioned, one of the hallmarks of cancer is sustained proliferation, which is considered to be one of the most fundamental traits of cancer cells [2]. Cyclins regulates the cell cycle and hence the growth. Too many or too few cyclins at the wrong time point are important reasons for a cancer's uncontrolled growth [3]. The levels of cyclins vary during the cell cycle. They are synthesized and degraded as needed to drive the different stages of the cell cycle forward (Figure 2). They bind to and regulate the activity of the cyclin-dependent protein kinases (CDKs) [4]. CDKs never act on their own: they depend on the cyclins for proper functioning [5, 6]. The activities of the various cyclin/CDK-complexes must be modulated to impose control on specific steps in the cell cycle (Figure 3.). This modulation is achieved by changing the levels and availability of cyclins during various phases of the cell cycle. Tumorigenic cells have inactivated or lost the late G₁-check point (R-point) [5].

The cyclin/CDK-complexes have two main functions, as follows: (1) to activate the enzymes that drive the cell cycle forward and into respective cell phases; and (2) in a negative feedback loop, to induce degradation of the cyclins from previous

phases, thereby driving the cell into the next phase. Since the cyclins of respective phases are degraded continuously, the phase transitions are irreversible.

The main phases of the cell cycle are as follows. G_0 (G stands for gap phase) is a non-growing state during which the cells wait for external signals before deciding to enter the active cell cycle. G_1 , also known as the growth phase, is the phase during which the cells' supply of proteins is increased, the number of organelles, such as mitochondria and ribosomes are increased, and the cell grows in size. G_1 is also the critical phase during which cells wait to enter into next phase. Damaged cells should be repaired before entering the S-phase. The G_1 checkpoint ensures that everything is ready for DNA synthesis. The S-phase (DNA synthesis) is the phase during which DNA is replicated. In the G_2 -phase, the cells prepare themselves for entering into M-phase. The G_2 checkpoint ensures that everything is ready for mitosis, and finally the M-phase (mitosis) is the phase during which cells are divided [3, 5, 6] (Figure 3).

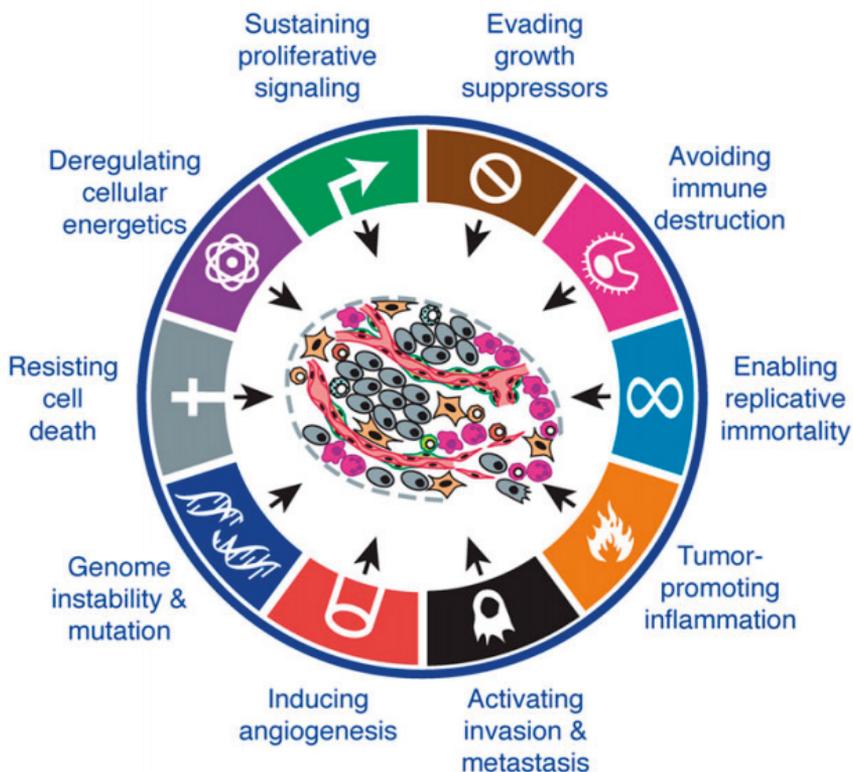


Figure 1. Hallmarks of cancer. Reprinted with permission from Elsevier [2]

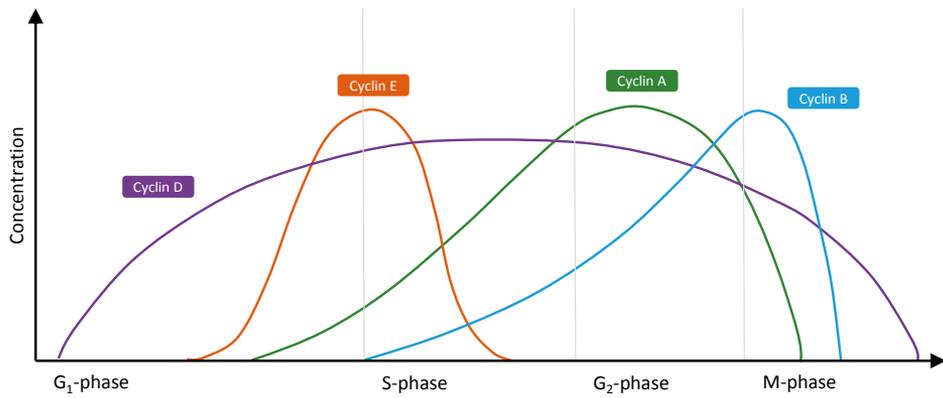


Figure 2. Levels of cyclins during the cell cycle

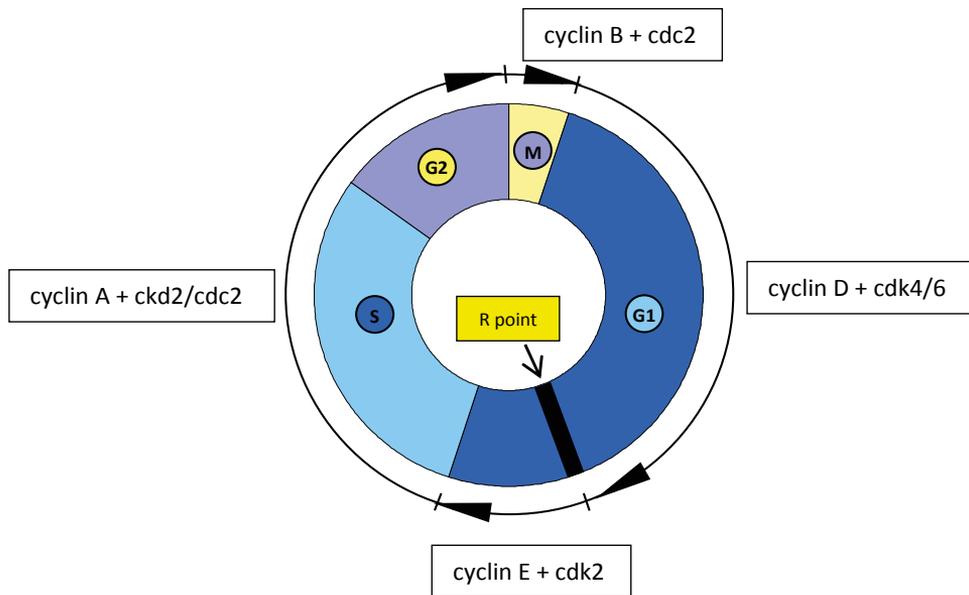


Figure 3. The cell cycle

Cancer is a disease of uncontrolled cell proliferation. One might presume that the control system within cancer cells is organized in a different manner than that of normal cells. In truth, they use almost identical control systems. Cancer cells make only minor modifications to the control steps. Many types of cancer cells have inactivated one or more of the checkpoint controls, which helps them to accumulate mutant genes and altered karyotypes, driving their neoplastic growth [5-7].

Breast cancer

General background

Breast cancer is the most common cancer today among women in the Western world. In Sweden in 2015, 8 022 women and 41 men were diagnosed with breast cancer [8]. For a large number of patients, the prognosis is good. The majority of the patients are cured by primary surgical treatment, and many of the recurring patients survive for long periods of time.

In Sweden, the breast cancer incidence has increased by 1.6% annually over the last 20 years and by 2.6% over the last 10 years [8], but at the same time, the mortality has decreased. Due to mammographic screening (introduced in Sweden in the 1980s), breast cancers are detected at earlier stages [9]. Early detection, together with improved local and systemic adjuvant treatment, is considered to be important reasons for the improved cure rate of breast cancer patients [10-12]. However, unlike other cancers, breast cancer has no ‘cure point’, indicating that, regardless of how long after the breast cancer diagnosis, the risk of death among breast cancer patients will always be higher than for women of the same age without breast cancer [13, 14]. The survival is dependent on adequate surgery and the administration of adjuvant therapy to patients at risk of developing recurrence. Sweden has one of the best 5-year survival rates for breast cancer worldwide [15]. The 5-year survival rate is currently almost 90%, and the 10-year survival rate is greater than 80%. In 2014, 1 405 women died of breast cancer in Sweden [8]. The leading cause of breast cancer related death is distant metastasis. Metastasis is the process by which malignant cells spread from the primary tumor site to other distant vital organs. Breast cancer predominantly spreads to the lymph nodes, skeleton, lungs, liver, and/or brain.

The majority of breast cancers are sporadic, but 5-10% of patients have a family history of the disease that suggests an inherited predisposition [16]. The most obvious susceptibility genes with high penetrance mutations are BRCA1 or BRCA2 [17, 18]. However, a large proportion of patients with a family history of breast cancer do not carry germline BRCA gene mutations, and their increased risk might be caused by defects in moderate or low penetrance genes, by combinations of genes or by still unknown factors [19].

In addition to genetic inheritance and being a woman, known risk factors for developing breast cancer are increasing age, early menarche and late menopause, nulliparity or older age at first full-term pregnancy [20-22], and hormone replacement therapy after menopause (use for >10 years) [23, 24]. Current use of oral contraceptives or use during the previous ten years moderately increases the risk of developing breast cancer [25, 26]. Lifestyle factors such as high intake of saturated fat, being overweight or obese after menopause, or excessive alcohol consumption, are also risk factors for breast cancer [17, 27].

Diagnostics

In Sweden, the majority of breast cancers are detected with mammography and the remaining cancers by the patient herself or her physician. Once a cancer is suspected, a diagnosis is reached by the triple diagnostic approach, which includes clinical examination (palpation), mammography, and investigation by fine or core needle aspiration [28, 29]. Further investigation by surgery is undertaken if at least one of these diagnostic tools provides a reason to suspect malignancy. When diagnosed, it is recommended that the patient be discussed at a multidisciplinary conference before surgery [30, 31]. Breast density, which has begun to be recognized as a risk factor for breast cancer, can decrease the diagnostic sensitivity and hence mask malignancies: therefore, ultrasound in combination with mammography might be recommended for young women (<30 years old) [32, 33].

Prognostic and treatment predictive factors in breast cancer

Prognostic and treatment predictive factors in breast cancer can help to predict the risk of breast cancer recurrence and to facilitate the choice of optimal treatment. Although these factors have improved treatment decisions, not every patient receives optimal therapy: hence, some patients are over-treated and others are under-treated. To assess the risk of recurrence and to select patients who will benefit from adjuvant therapy, prognostic factors are used. A prognostic factor is correlated with recurrence in untreated patients. A predictive factor, in contrast, is correlated with the likelihood of a response to a certain treatment [34, 35]. The generally accepted prognostic factors are lymph nodes status, tumor size, age at diagnosis, histologic grade, hormone receptors (estrogen (ER) and progesterone (PgR)), human epidermal growth factor receptor 2 (HER2), and proliferation. ER

and HER2 are also used as predictive factors [36, 37]. Ki67, frequently used to determine the rate of proliferation, is only recommended by the St Gallen guidelines [36]. The American Society of Clinical Oncology (ASCO) guidelines, however, do not recommend the use of Ki67 for guidance in the choice of adjuvant chemotherapy [37]. The molecular profiles, *Oncotype DX*[®] and *MammaPrint*[®], have been approved by the St Gallen and ASCO guidelines [36, 37]. Recently finished or ongoing trials evaluate the prognostic and predictive value of these profiles [38-40]. Web-based tools such as Adjuvant! Online or PREDICT can be used to estimate the risk of recurrence (within a certain time frame, e.g., 5 years) for single patients [41-43]. However, such predictions should be interpreted at the group level and not as predictions for individual patients [44].

Proliferation

As previously mentioned, one of the hallmarks of cancer is sustained proliferation, which is considered to be one of the most fundamental traits of cancer cells [2]. Therefore, several methods have been developed to estimate the proliferation rate of tumor cells, and numerous studies have shown that high proliferation rates, irrespective of evaluation method, are associated with worse prognosis in breast cancer [4, 45-56]. Proliferation rates are, in general, higher in triple-negative (ER/PgR/HER2-negative), HER2-positive and receptor-negative tumors, compared to ER-positive tumors [57-60]. Furthermore, it seems as if the added prognostic value of proliferation is most pronounced in ER-positive/HER2-negative tumors [61, 62]. Measurement of proliferation rates can be performed either by the assessment of single factors, e.g., using immunohistochemistry (IHC), or as the main common denominator of different genetic profiles [63, 64].

A high proliferation rate has also been shown to be associated with a more favorable response to chemotherapy [65, 66]. Hence, a validated proliferation marker might be useful not only for prognostic purposes but also, more importantly, for treatment predictive purposes.

Markers of proliferation

Mitotic activity

Mitosis count is one of the three components of histological grade (together with tubular formation and nuclear atypia) [67]. According to the Elston and Ellis grading system, mitoses should be assessed in ten fields of vision (FOVs), starting at the periphery of the tumor [67]. According to the World Health Organization

(WHO), the number of mitoses is categorized into three groups as a part of histological grading: group 1 (0–5 mitoses), group 2 (6–10 mitoses), and group 3 (>10 mitoses) [68]. The WHO has published recommendations for a more strict mitosis counting procedure: mitotic activity index (MAI) [69]. Instead of ten random FOVs, ten FOVs in ten neighboring consecutive fields should be assessed. Furthermore, the manner in which field diameter is defined is stricter. When the WHO protocol has been strictly followed, MAI has been highly reproducible [47].

Some studies have shown that the prognostic value of histological grade is mainly due to mitosis [55, 70-72]. Other studies have shown that nuclear atypia is the most important part of histological grading [73]. However, problems with reproducibility have been reported [74-77].

Thymidine labeling index

Another of the first markers for the measurement of proliferation rate was thymidine-labeling index (TLI), which measures the proportion of cells synthesizing DNA at a given time. Fresh tumor material is required. The tissue is incubated in medium containing thymidine labeled with the nucleotide tritium (^3H). The level of incorporated ^3H -thymidine into the tissue is measured, and a ratio of labeled to non-labeled DNA is calculated [78, 79]. Several studies have shown the prognostic and predictive importance of TLI [66, 78, 80-82].

S-phase fraction

Using a flow cytometer, the S-phase fraction (SPF) can be analyzed either in paraffin-embedded or fresh frozen material. However, the majority of publications showing correlations with prognosis in breast cancer were performed on fresh frozen material. The first publication showing a correlation between SPF and prognosis in breast cancer was published in 1984 [83]. Before the cells are processed through the flow cytometer, they are mechanically disrupted and stained with propidium iodide. DNA histograms are obtained, which show the distribution of cells corresponding to the phases of the cell cycle, in which the major peak usually represents the G_0/G_1 -phase. Some of the disadvantages are that, for best performance, fresh frozen tissue is required, a larger tumor volume than, for instance, for IHC is needed, and a mixture of normal cells and cancer cells is analyzed. Several studies have shown a correlation between a high SPF and poor prognosis in breast cancer [84-86]. The measurement of SPF by flow cytometry, combined with PgR and tumor size, was used as a prognostic index in parts of Sweden and Finland in the 1990s [86, 87].

Ki67

Proliferation can also be evaluated by measuring Ki67, which is a cell cycle antigen. It was originally described in Kiel (1983): hence, Ki stands for Kiel, and

67 is the number of the original clone in the 96-well plate from which it was derived [88]. Ki67 is expressed by all proliferating cells in the body, and it is present in all active phases of the cell cycle: G₁, S, G₂ and M, peaking during the mitotic phase, (hence not in G₀) [89, 90]. The exact function of Ki67 is unknown, but a role has been suggested in ribosomal RNA-synthesis [91, 92].

An association between the overexpression of Ki67 and worse outcomes in early breast cancer has been observed in several studies [47, 53, 93, 94]. Ki67 might also be an indicator of better response to chemotherapy [95]. The independent prognostic utility of Ki67 has been shown, especially among patients with ER-positive tumors [61]. Expression of Ki67 can separate ER-positive tumors into two subclasses, a low proliferative group ('luminal A-like'), associated with a favorable prognosis: and a high proliferative group ('luminal B-like'), associated with a less favorable prognosis [60].

Because Ki67 carries both prognostic and predictive information, and it is affordable and easily measured on paraffin embedded tumor material, it is widely used. However, there is not yet an international consensus on the assessment of Ki67, although recommendations have been proposed [96, 97]. The ASCO guidelines from 2016 still do not recommend the use of Ki67 to guide the choice of adjuvant chemotherapy [37], because of a lack of consensus concerning methodology, especially cut-off values [98]. Different antibodies and scoring systems have been used [99]. However, in the 2011 St Gallen guidelines, Ki67 was recommended to be used to distinguish between slowly proliferating 'luminal A-like' and rapidly proliferating 'luminal B-like' tumors [100]. Ki67 can furthermore be used to subdivide histological grade 2 into groups with different prognosis [61, 62]. Studies have shown poor reproducibility, and efforts have been undertaken to standardize assessment, choice of cut-off, antibody, etc., but issues with poor reproducibility persist [96, 98, 101-103].

Cyclins

Cyclins can be studied on either the DNA or the protein level, using, e.g., fluorescence *in situ* hybridization (FISH) or IHC. In this thesis the focus is on the protein level as measured by IHC.

Cyclin D

Cyclins D1, D2 and D3 exist: however, in breast cancer, the most studied cyclin is cyclin D1, which is associated with CDK4/6 to ensure the advancement of the G₁-phase. Overexpression of cyclin D1, often due to amplification of the coding gene, is found in a large proportion of breast cancers but only rarely in benign and premalignant lesions. Although it is involved in the activation of proliferation, studies have shown links to a less malignant phenotype, with high levels of cyclin D1 correlated with better prognosis [104-108]. In contrast, an association with

favorable prognosis has been shown on the protein level, while amplification of the gene has shown an association with poor prognosis [109, 110].

Cyclin E

Cyclins E1 and E2 localize to different sites in breast cancer cells, but they also co-localize within the cell [111]. Although genes on different chromosomes encode cyclin E1 and cyclin E2, they both activate CDK2 [111]. Together with CDK2, cyclin E2 forms a catalytically active complex, and it has been suggested that cyclin E1 and cyclin E2 phosphorylates different sets of substrates and thereby regulate the progression of cells through the G₁-phase and into the S-phase [112]. The critical determinant for entering into S-phase is cyclin E in cooperation with CDK2. Several studies have identified cyclin E as important in breast cancer prognosis [113-118]. Recently, a study by Karakas et al., showed that cytoplasmic cyclin E, together with phosphorylated CDK2, could be used as a biomarker for aggressive breast cancer because it predicted shorter recurrence-free and overall survival [119].

Cyclin A

Cyclin A1 is expressed during meiosis and embryogenesis, while cyclin A2 is expressed in dividing somatic cells [120]. Cyclin A2 is responsible for the initiation of DNA replication and it is also involved in G₂ to M-phase transition. It can activate two different CDKs: CDC2 (also called CDK1) and CDK2. The level of cyclin A2 rises in the early S-phase and falls in the mid M-phase. As cells enter the S-phase, cyclin A2 replaces cyclin E as the partner of CDK2, thereby enabling the S-phase to progress. As the cell moves further and enters the G₂-phase, cyclin A2 switches partner and binds to CDC2. Like cyclin E, cyclin A2 has been associated with worse outcomes in breast cancer patients in some studies, but other studies have been unable to confirm this finding (perhaps due to a lack of consensus concerning cut-off values, analogous to Ki67) [4, 45, 49, 121-126]. In this thesis, cyclin A2 is hereafter referred to as cyclin A.

Cyclin B

It has been suggested that cyclin B1 is involved in chromatin condensation and that cyclin B2 is involved in disassembling the Golgi apparatus [127]. The final step in the cell cycle is controlled by the cyclin B1–CDK1/CDC2 complex, also known as the mitosis-promoting factor. To date, only a few publications have found that high expression of cyclin B1 was correlated with poor survival [128-130].

Cyclin H

Recently, another cyclin, cyclin H, has been shown to be correlated with outcomes in breast cancer, and once again, it was most pronounced for ER-positive breast

cancer. Patel et al., found that the majority of tumors with high expression of cyclin H were ER-positive with a better prognosis [131].

Phosphorylated histone H3, PPH3

Phosphorylation of histone H3, called PPH3 or PHH3, occurs late in the G₂ and M-phases, and it is involved in initiating chromatin condensation. It is needed for the cell to progress to mitosis. PPH3 is assessed with IHC and is evaluated similarly to MAI. Positive cells are counted in ten consecutive FOVs, starting at the periphery in the area with the highest proliferation. Difficulties in counting mitotic figures can be associated with condensed mitoses due to hypoxia or suboptimal fixation: using PPH3 circumvents this issue. Studies have shown a correlation between high expression of PPH3 and poor prognosis [55, 132-136].

Other prognostic and treatment predictive factors in breast cancer

TNM classification

The most important prognostic factor is the TNM-classification, in which T stands for tumor size, N for spread to the regional lymph nodes, and M for the presence of distant metastasis [137]. TNM classifies the tumor and provides a strong indication of prognosis. Patients are classified into five prognostic stages (0–IV), and higher stages have worse prognosis.

Patients with larger tumors have worse prognosis than patients with smaller tumors [138]. Accordingly, tumor size is divided into four groups with increasing risk for recurrence: T1: 1–20 mm; T2: 21–50 mm; T3: >50 mm; and T4: a tumor of any size, with skin or chest wall involvement.

The more lymph nodes that are involved, the higher the risk is for developing distant metastases [35]. Lymph node involvement is divided into: N0: no involvement; N1: 1–3 positive lymph nodes involved; N2: 4–9 lymph nodes involved; and N3: ≥10 lymph nodes involved.

If the patient has a tumor that is less or equal to 2 cm in size, with no spread or dissemination to the lymph nodes in the axilla or to distant sites, the tumor is classified as stage I, with better prognosis than for higher stages [35]. Stage IV tumors are considered incurable [139].

Since the introduction of mammographic screening, the fraction of small tumors with no spread to the lymph nodes has increased: hence tumors are diagnosed and surgically removed at lower stages today [9].

Age at diagnosis

Approximately 5–10% of women diagnosed with breast cancer are younger than 40 years old [8]. Studies have showed that breast cancer in young women is associated with worse prognosis [140, 141]. Fredholm et al., showed that women <35 years old had worse breast cancer-specific survival (BCSS) (69% 10-year breast cancer-specific survival) compared to 76% for 35–39 years old, and 84% for 40–49 years old [142]. Their study reported that the risk for relapse in young women was most pronounced in stage I–II and luminal B tumors. Kroman et al., found a similar age trend, with younger women having a higher risk of dying: however this trend vanished in patients who received chemotherapy. They suggested that young age alone should be considered a high risk factor [143]. Elderly patients in general have better outcomes than patients younger than 60 years old [144]. The majority of newly diagnosed breast cancer patients are 50 years old or older [8]. In Sweden in 2015, the median age at diagnosis was 64 years old [8].

Histological grade

Histological grading, analyzed on hematoxylin/eosin-stained tissue slides using a light microscope, is the combined evaluation of the presence of tubules, how pleomorphic the nuclei are, and mitotic activity. Each category is scored 1–3, and these scores are added into a histological grade score (3–9 points). Grade 1 is defined as 3–5 points, grade 2 as 6–7 points, and grade 3 as 8–9 points. Grading was first described by Bloom and Richardson [145] and was later revised by Elston and Ellis in 1991 [67]. Several studies have shown correlations with prognosis [75, 146, 147], other studies have shown problems with reproducibility [67, 75, 147, 148].

Estrogen and progesterone receptors: ER and PgR

Estrogens play an important role in the development of breast cancer. The majority of all breast cancers are ER-positive (85% in Sweden in 2015) [149]. The incidence of ER-positive cases increases with age. There are two types of ER, ER α and ER β . ER α has proved to be correlated with sensitivity to endocrine treatment.

ER β has not been as well studied. The term ER in this thesis refers to ER α . By also evaluating PgR, it is possible to improve the prediction of outcome, and in the St Gallen guidelines 2013, PgR was included to guide the distinction between ‘luminal A-like’ and ‘luminal B-like’ breast cancer [36]. However, according to the latest meta-analysis from EBCTCG, PgR was not predictive of response to endocrine treatment [150].

Both ER and PgR are present in the cell nucleus and can be routinely detected by IHC in the clinic.

Human epidermal growth factor receptor 2

The cell surface tyrosine kinase receptor human epidermal growth factor receptor 2 (HER2) is a member of the epidermal growth factor receptor family. HER2 is both prognostic and predictive [151, 152]. Anti-HER2-targeted therapy has been administered since the late nineties for metastatic breast cancer and since 2005 as an adjuvant treatment. In Sweden, 12-15% of breast cancers overexpressed HER2 and/or were amplified [149, 153].

The HER2 protein is assessed by IHC and gene amplification by FISH, chromogenic *in situ* hybridization (CISH), or silver enhanced *in situ* hybridization (SISH). Since 2011, using the Inform HER2 Dual ISH, chromosome 17 and the HER2 gene can be chromogenically stained at the same time.

Molecular profiling

Historically, prognostic factors have been identified and used to stratify patients into subgroups with distinct biology, prognosis and response to treatment. For the majority of these factors, evaluation was/is performed on a single protein level. Since the introduction of gene expression profiling techniques, it is possible to perform this analysis on a multi-gene level [154]. With distinct combinations of genes, Perou and Sorlie pioneered the use of these new techniques to identify molecular classes [57, 63, 155]. Others soon followed [58, 156-158]. Important results from these studies are that ER-positive breast cancer is a distinctively different form of breast cancer than ER-negative [159]. Furthermore, molecular subtypes identified two luminal subtypes (luminal A and B), which are ER-positive with different expression of PgR and Ki67, HER2-enriched and basal-like (ER/HER2/PgR-negative) tumors. Luminal A tumors, which are low proliferating and HER2-negative, generally have better outcomes than luminal B tumors [160]. Since the introduction of these techniques, a plethora of prognostic breast cancer profiles have been reported [156, 158, 161-164]. Many studies have shown the

importance of proliferation [60, 158, 165, 166]. To date, only a few profiles are in clinical use and are recommended in international guidelines [36]. Recommended by the ASCO guidelines are *Oncotype DX*[®] recurrence score and PAM50-Risk of Recurrence Score (ROR) [37, 156, 162]. Furthermore, the clinical utility and validity are being tested in recently closed or ongoing trials, such as MINDACT (testing the utility of MammaPrint[®]), TAILORx and RxPONDER (testing the usefulness of *Oncotype DX*[®]) [38-40, 167-169].

Transcriptomic-based assays are being increasingly used in clinical practice, particularly in the US. However, due to complicated technologies, high required expertise and expense, it is not available for the majority of patients worldwide. As an alternative, the St Gallen International Expert consensus 2013 panel has approved surrogate criteria for determining the molecular subtypes [36]. This surrogate classification considers different combinations of the four biomarkers, ER, PgR, HER2, and Ki67, and can be used when it is not possible to perform genomic tests, due to cost or other reasons. The markers are combined as follows:

‘Luminal A-like’ (ER-positive, PgR high, HER2-negative, Ki67 low);

‘Luminal B-like (HER2-negative)’ (ER-positive, HER2-negative, and at least one of Ki67 high or PgR low);

‘Luminal B-like (HER2-positive)’ (ER-positive, HER2-positive, any Ki67 and any PgR);

‘HER2 positive’ (ER-negative, PgR low, HER2-positive); and

‘Triple negative’ (ER-negative, PgR low, HER2-negative).

Other promising prognostic factors

Examples of other promising prognostic factors are circulating tumor cells (CTCs), tumor-infiltrating lymphocytes, and circulating tumor DNA: however these factors do not have the evidence required to be included in the current guidelines [37, 97]. According to the ASCO guidelines, urokinase plasminogen activator (uPA) and plasminogen activator inhibitor type 1 (PAI-1) can be used for ER/PgR-positive, HER2-negative breast cancer: however the level of recommendation is weak [37].

Treatment

After surgery, a combination of different treatments is administered to a large proportion of patients to prevent disseminated cancer cells from creating metastases at distant sites. Treatments can be administered as local therapy (radiotherapy) or systemic therapy (endocrine, chemotherapy, and/or anti-HER2 targeted treatment). Systemic therapy can also be administered before surgery (neoadjuvant). For patients with metastatic breast cancer, systemic palliative treatment is given. If the primary tumor is inoperable due to the size of the tumor, neoadjuvant therapy can be administered to shrink the tumor and render it surgically removable. In Sweden, postmenopausal, node-positive patients can be offered bisphosphonate treatment as an additional adjuvant treatment [31].

If not elsewhere stated, the recommendations are according to the Swedish guidelines and the presented figures represents Swedish breast cancer patients [31].

Surgery

The primary therapy is removal of the tumor, either by performing breast-conserving surgery (BCS) or by removing the whole breast, i.e., mastectomy [170, 171]. Cancers that are detected by mammography are statistically more likely to be treated with BCS [172]. In Sweden, three quarters of breast cancer patients are operated on with BCS (if the tumors are <30mm) [149]. The sentinel lymph node procedure (introduced in 1990) is used to decide whether more lymph nodes must be removed or not [173, 174]. Injection of radioactively labeled albumin and a blue dye identifies the sentinel nodes [175]. The sentinel nodes are surgically removed and are sent to the pathology department, where they are examined perioperatively on frozen sections. If the lymph nodes contain metastases ≥ 0.2 mm in diameter, further axillary dissection is performed [31]. Ongoing trials are studying whether axillary dissection can also be omitted when macrometastases (>2mm) are detected [176, 177].

Radiotherapy

Postoperative radiotherapy is administered as a precaution in order to reduce the risk of local and regional recurrences. Patients who have undergone BCS are offered radiotherapy. Furthermore, radiotherapy is offered to patients who have undergone mastectomy if the tumor is >50 mm, or at least one positive lymph node is detected in the axilla [31]. Several studies comparing mastectomy with

BCS followed by radiotherapy have shown no difference in survival between the two methods [178-181].

Endocrine therapy

If a tumor is considered ER-positive, which in Sweden indicates that the tumor expresses >10% ER (the St Gallen [97] and ASCO guidelines [37] recommend a cut-off of >1% for positivity), endocrine treatment is administered to the great majority of patients [11, 31].

Endocrine treatment alternatives consist of either tamoxifen (TAM), which is an anti-estrogen and binds to ER in an antagonistic manner, aromatase inhibitors (AIs), which target the enzyme aromatase (aromatase is involved in the process of converting androgens into estrogens), or drugs suppressing the ovarian production of estrogens. Selective ER modulators (SERMs), such as TAM, antagonize ER function. TAM belongs to the SERM group of drugs that mimic the effects of estrogen in some tissues, while opposing them in others [182]. For instance, TAM inhibits ER activation in the breast while stimulating it in the endometrium [183].

Five years of adjuvant TAM (vs. no treatment) reduces the relative risk of recurrence by approximately 50%, and mortality by 30% [184]. Five years of TAM is recommended as adjuvant therapy, but a benefit of prolonged treatment has been reported [185, 186]. TAM has effects in both pre- and menopausal patients, whereas AIs are effective only in postmenopausal women because AIs do not suppress estrogen production in the ovaries [187, 188]. Compared to five years of TAM given to postmenopausal breast cancer patients, AIs reduce 10-year relative breast cancer mortality by 15% [184, 189]. Selected patient groups, mainly node-positive, are offered prolonged endocrine treatment, but it is today not recommended to administer AIs for more than five years. However, studies of more than five years of AI are ongoing [190].

Chemotherapy

Adjuvant chemotherapy is recommended for patients at high-risk for recurrence, i.e., either young age, high proliferation rate, high histological grade, node-positivity, large tumor size, and HER2-positive, ER-negative and/or PgR-negative tumors. The most common chemotherapy regimen consists of an anthracycline-based combination, i.e., FEC (5-fluorouracil, epirubicin, and cyclophosphamide) or FAC (5-fluorouracil, doxorubicin (Adriamycin), and cyclophosphamide), followed by a taxane, either paclitaxel or docetaxel [191].

Anti-HER2 therapy

If a patient has an HER2-positive tumor, indicating an IHC-score of 3+ and/or with the HER2 gene amplified, the monoclonal antibody trastuzumab, targeted at HER2, is recommended for use in the adjuvant setting. In 2015, HER2-positive cancers represented 14% of cases in Sweden [149]. The monoclonal antibody pertuzumab and TMD-1 are other anti-HER2 directed drugs that are being studied in the adjuvant setting [192].

Positive and negative effects of adjuvant treatment

Large meta-analyses have shown that breast cancer mortality is reduced, and overall survival is improved by adjuvant treatment [11, 193]. However, all adjuvant treatment regimens also induce side effects. After radiotherapy, the typical local acute side effects are erythema and pneumonitis [194]. Late side effects of the cardiovascular system with increased cardiac mortality have been documented in several studies with long-term follow-up, among others, in the SB II:1 trial from the South Swedish Breast Cancer Group [195].

For those patients treated with adjuvant chemotherapy, the majority experience mild or moderate nausea and vomiting [196]. The most dangerous side effect of chemotherapy is bone marrow suppression and risk of neutropenic fever.

The most common side effects with TAM are menopausal symptoms and increased risk of thrombosis and pulmonary embolism [197]. A common side effect of AIs is arthralgia and AIs can also cause osteoporosis [184, 198]. TAM is generally better tolerated, at least in postmenopausal patients, and it has been reported to delay myocardial arteriosclerosis [199].

In addition to the side effects, many treatments are expensive. Nevertheless, the benefit from adjuvant therapy outperforms the side effects, and strategies for personalizing treatment are being investigated. Although a large proportion of breast cancer patients are cured by surgery alone, the majority receives systemic treatment despite the risk of recurrence being low. Number one of the top ten research questions in the web-based consultation performed by Dowsett et al., was to identify tools to better guide the selection of patients for whom adjuvant chemotherapy could be avoided [200]. This thesis focuses on improved prognostication and is thus one of many approaches for attaining this goal.

Background to the thesis

When this thesis was initiated, other groups had shown that, using gene expression profiles, patients with ER-positive breast cancer and also breast cancer classified as histological grade 2, could be subdivided into two groups with distinctively different prognosis [59, 63, 201]. In the ER-positive subgroup, genes associated with proliferation seemed to be the most important, whereas genes associated with immune response were more important among ER-negative breast cancers [59]. Since the most important genes in the ER-positive subgroup were associated with proliferation, we wanted to investigate whether the use of a single factor (Ki67), could provide similar results. The answer was yes [61]. At the same time, Ahlin et al., showed that cyclin A could be used as a marker of proliferation [126]. Because cyclin A seemed easier to evaluate (contrast-rich and crisp staining), we wanted to investigate this further.

Although it is a well-known fact that information is lost when predictors are dichotomized or categorized, for clinical decision-making, this is often performed. One possible strategy could be to postpone dichotomization/categorization until the last step in the modeling process, thereby exploiting the information included among continuous predictors. Surprisingly few publications have addressed this issue, and we addressed it in this thesis.

Aims of the thesis

Overall aim

The overall aim was to investigate prognostic and predictive markers in breast cancer and to find methods of improving the use of established factors, with a focus on proliferation and non-linear effects. These results could be used in clinical decision-making to determine which patients need or do not need adjuvant systemic treatment, especially chemotherapy, after primary surgery.

Study I

Our aim was to evaluate the performance of the new variable CAGE: a combination of cyclin A, histological grade, and ER, in premenopausal node-negative breast cancer patients.

Our hypothesis was that an index based on cyclin A, ER, and histological grade would provide valuable prognostic information. Furthermore, we wanted to investigate the prognostic importance of cyclin A alone and in subgroups with different histological grade, and in relation to genetic grade.

Study II

The aim was, in a larger dataset, to confirm the previously defined prognostic index (CAGE) and, instead of cyclin A, to combine the worldwide used proliferation marker Ki67 with histological grade and ER (KiGE).

Our hypothesis was that KiGE was equally valid to CAGE and that it could confirm the prognostic value of the combination of a proliferation factor with histological grade and ER.

Study III

The aim was to evaluate the prognostic value of MAI, PPH3, cyclin B1, cyclin A, and Ki67, alone and in combinations, in node-negative breast cancer.

Our hypothesis was that a combination of proliferation factors would improve prognostication.

Study IV

The primary aim was proof of principle, i.e., to evaluate whether accounting for non-linear effects of the three factors age at diagnosis, tumor size, and number of positive lymph nodes improved prognostication.

Our hypothesis was that, by keeping the predictors continuous for as long as possible in the modeling process, prognostication would be improved.

Patients

Studies in thesis

Original cohorts/studies

Study I and III, (N0)

SB 91:B

Study II, (chemo naïve N0/N1)

SB II:2 pre and post
Malmö cohort
Bone marrow metastases cohort
Odense cohort
Uppsala study

Study IV

SB II:1 pre and post
SB II:2 pre and post
Bone marrow metastases cohort
Odense cohort
Kalmar cohort

Study I and III

The patients came from the SB 91:B study population ($N = 237$, 59 with distant metastases) [202], which between 1991 and 1994 included premenopausal lymph node-negative women. Adjuvant treatment was administered in 29 (13%) of the patients. It was a trial with the purpose of studying the prognostic importance of prospectively analyzed SPF. This trial was also a part of the node-negative breast cancer (NNBC)-trial [86]. Fourteen patients were excluded due to lack of paraffin blocks. Cases were also excluded if fewer than 200 cells were counted for cyclin A, leaving 219 patients included in *Study I*.

In **Study III**, samples from $N = 221$ patients were evaluated for PPH3, $N = 195$ for MAI, $N = 199$ for Ki67, and $N = 217$ for cyclin B. At the time of **Study III**, one additional case of cyclin A was excluded due to an error in the database, leaving $N = 218$ evaluated for cyclin A in **Study III**.

Study II

In total, 1 854 women with primary breast cancer originating from two randomized clinical studies, three cohorts and one case-control study were included. Only patients who were node-negative or with 1–3 positive lymph nodes were included. Patients were excluded if they had received chemotherapy or if information on adjuvant therapy, Ki67, histological grade or ER was missing. The patients came from the following patient materials:

SB II:2-pre ($N = 221$, 68 with distant recurrences): Between 1986 and 1991, women were included in this randomized trial of premenopausal stage II breast cancer, with the aim of comparing the effects of two years of TAM vs. no adjuvant systemic treatment [203, 204].

SB II:2-post ($N = 166$, 22 with distant recurrences): Between 1983 and 1991, women were included in this randomized trial studying two vs. five years of TAM in postmenopausal women [205].

The Malmö cohort ($N = 217$, 32 with recurrences): This consecutive series in Malmö enrolled breast cancer patients between 1988 and 1992 with the purpose of evaluating biomarkers on tissue microarray (TMA) slides [206].

The Bone marrow metastases cohort ($N = 379$, 27 with distant recurrences): This consecutive series in the South Swedish Health Care Region enrolled breast cancer patients between 1999 and 2003 with the purpose of studying cytokeratin positive cells in bone marrow aspirates from the sternum [207].

The Odense cohort ($N = 539$, 86 with distant recurrences): This consecutive series in Odense enrolled patients between 1980 and 1990 with the purpose of evaluating the prognostic value of Chalkley counting in a population-based cohort [208].

The Uppsala study ($N = 166$ cases and 166 controls) was used as an additional dataset. Patients were enrolled in the study between 1993 and 2004, and within this cohort, cases were defined as women who died from breast cancer. The purpose of the study was to evaluate the prognostic value of cyclin A in node-negative patients, and furthermore to validate previously suggested cut-off values [126].

Study IV

A derivation set was compiled of 4 477 women with primary breast cancer originating from four multicenter randomized clinical studies and three cohorts. The patients came from the following patient materials:

SB II:1-pre ($N = 362$, 125 with distant metastases): Between 1978 and 1983, women were included in this randomized trial with the purpose of evaluating the effects of chemotherapy and radiotherapy, alone and in combination [195].

SB II:1-post ($N = 624$, 237 with distant metastases): Between 1978 and 1985, women were included in this randomized trial studying the effects of TAM for one year and radiotherapy, alone and in combination, in postmenopausal women [195, 209].

SB II:2-pre ($N = 561$, 233 with distant metastases) [203], SBII:2-post ($N = 1,553$, 382 with distant metastases) [205], the Odense cohort ($N = 841$, 251 with distant metastases) [208], and the Bone marrow metastases cohort ($N = 536$, 87 with distant metastases) [207]. For more information on the last four studies, see above beneath the subheading **Study II**.

A limited number of patients ($N = 537$) were excluded due to missing information on follow-up, number of positive lymph nodes, and/or tumor size.

The Kalmar cohort ($N = 1 132$, 311 with distant metastases), a consecutive series of primary breast cancers, was used as a validation dataset.

Methods

Tissue microarray

Tissue microarray (TMA), as a high-throughput technique for paraffin-embedded material, was introduced in 1998 [210]. One of the advantages of TMA is that it cost-effectively enables simultaneously staining of hundreds of samples on only one slide meaning that, using only a limited amount of tissue, fewer amounts of antibody are necessary, with minimal differences in staining conditions. The disadvantages include representability and heterogeneity, i.e., the risk of not capturing the region of interest in the tumor tissue due to the small tumor area in the sample.

In this thesis, TMA-blocks were prepared from paraffin-embedded blocks using a manual arrayer (Beecher Instruments Inc., Sun Prairie, WI, USA). Two 0.6-mm cores (*Studies I and II*) and cores of 1.0 mm (*Study III*) were obtained from representative areas (i.e., with tumor cells) of each primary tumor block and were transferred to a recipient block. Then, 3–4 μm sections were cut from each block and transferred to glass slides (Menzel Superfrost[®] Plus), dried at room temperature and then incubated in a heat chamber for 2 hours at 60°C.

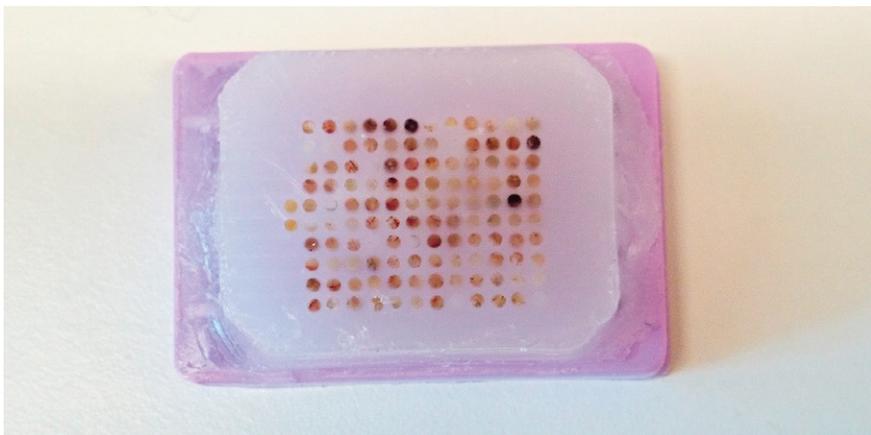


Figure 4. Breast cancer TMA paraffin block.

Immunohistochemistry

In tumor diagnosis, immunohistochemistry (IHC) is commonly used to detect antigens and to assess the distribution and localization of proteins in the tissue. Its advantages include ease of methodology, low costs, and high efficiency. Furthermore, distinct cancer cells can be evaluated, thereby avoiding evaluation of non-malignant cells in the tumor. Its disadvantages include choice of antibody, and bad staining quality depending on poor tissue fixation and poor reproducibility of evaluators. These disadvantages underpin the need for defined and validated guidelines to render the choice of cut-offs easier and the use of established antibodies, and to guide the methods of evaluation.

Immunohistochemistry was used in *Study I* (cyclin A), *Study II* (Ki67), and *Study III* (PPH3, cyclin B1, cyclin A and Ki67). To visualize the proteins of interest, primary antibodies directed against the antigen were used. A secondary antibody, usually conjugated with biotin directed to the primary antibody, was added and visualized with a chromogen, e.g., diaminobenzidine (DAB). The protein of interest was colored brown, and the surrounding tissue was colored blue with hematoxylin.

The TMA-slides were deparaffinized in xylene and rehydrated in decreasing concentrations of ethanol solutions (from absolute ethanol to distilled water). Immune staining was performed in an automated Autostainer *Plus* (Dako Denmark A/S), according to the manufacturers procedures’.

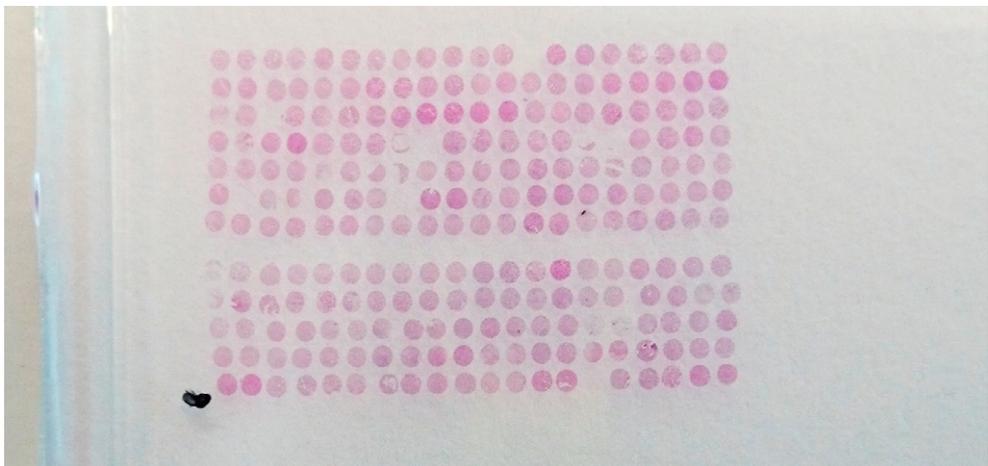


Figure 5. Breast cancer TMA-slide. Stained with hematoxylin/eosin. White areas in the rows represent missing TMA-cores.

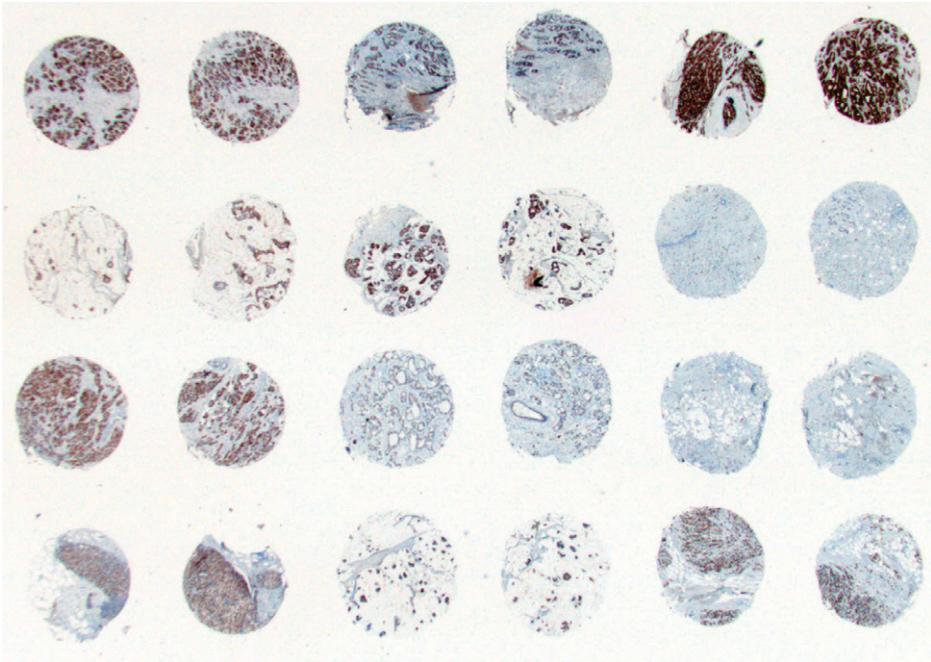


Figure 6. Breast cancer TMA-slide. TMA-cores stained for Ki67 and counterstained with Mayer's hematoxylin (blue). Brown color represents cells positive for Ki67.

Antibodies used in this thesis

Cyclin A

The cyclin A2 (NCL-Cyclin A, 1:100, Novocastra Laboratories) antibody was applied at room temperature. The staining was evaluated by one more experienced (Cecilia Ahlin, CA) and one less experienced (Carina Forsare, CF) investigator. Numbers of positive cells were counted in the TMA-cores with the most positivity, and a minimum of 200 cells were counted. If there were not sufficient cells in the first core, additional cells in the second core were counted until 200 cells in total were counted. High cyclin A was predefined at the 7th decile of the cyclin A distribution, which corresponded to $\geq 15\%$ (CA) and $\geq 17\%$ (CF) positively stained tumor cells, respectively. The agreement between the investigators evaluations of cyclin A expression was good (as measured by kappa statistics, value of 0.71). In **Studies I** and **III**, we chose to focus on the results from the more experienced investigator (CA).

Ki67

The analysis of Ki67 has been described elsewhere [61, 211]. In short, the MIB-1 antibody (DAKO, K5001, 1:500, Copenhagen, Denmark) was used. As predefined, cases above the 7th decile of the empirical Ki67 distribution (which corresponded to 20% positive cells) was used as cut-off.

PPH3/PHH3

The anti-phosphohistone H3 antibody (ser 10 Upstate #06-570, Lake Placid, NY, USA) at a 1:1500 dilution was used. Previous assessments of PPH3 were performed on whole sections, starting at the periphery of the tumor in ten consecutive FOVs with a total area of 1.59 mm². To mimic whole sections, the number of positively stained nuclei in one TMA-core was multiplied by 1.59 and divided by the total area of the TMA, 1.13 mm². The cut-off value was predefined at the 7th decile, which corresponded to ≥ 7 positive cells.

MAI

MAI was assessed by one experienced pathologist (JB) on hematoxylin-eosin whole sections according to the MMMCP 1987 protocol [212]. The area with the highest proliferation, in the periphery of the tumor, was chosen. In ten consecutive FOVs, structures that were mitotic figures were identified. The MAI was defined as the total number of mitoses in an area of 1.59 mm². The same cut-off as previously described was chosen, with ≥ 10 mitoses defined as high risk [55].

Cyclin B1

Slides were stained with the cyclin B1 antibody (cyclin B1 Y106, 1:200, Epitomics Inc., Burlingame, CA, USA). A minimum of 200 cells was counted. Cyclin B1 was assessed as previously described [129]. The cut-off value was predefined at the 7th decile of the cyclin B1 distribution, which corresponded to $\geq 12.5\%$ positive cells.

ER, PgR, HER2, and histological grade

ER, PgR, HER2, and histological grade were analyzed and evaluated as described elsewhere and previously defined cut-offs were used [126, 203, 206, 207, 213].

Statistical methods and concepts

The statistical analysis software Stata (Stata Corp, College Station, TX, USA) was used for data analyses in all of the studies included in this thesis. In *Study I*, Stata 11.1 2010 was used and in *Studies II* and *III*, Stata 12.1 2012 was used. For analyses of the case-control study in *Study II*, SAS was used (SAS Institute, Inc.). In *Study IV*, Stata 14.1 2016 was used.

Statistics used but not extensively described

Stepwise Cox analysis with backward elimination was used to fit the best model when highly correlated factors were included in the same multivariable model. An interaction term was used to test the null hypothesis of the effect of a factor in different subgroups, e.g., the effect of cyclin A in ER-positive vs. ER-negative breast cancer. Kappa statistics were used to evaluate the agreement between evaluators. Pearson's correlation coefficient, Pearson's χ^2 test, and Pearson's χ^2 test for trend were used for the analyses of associations between factors. Forest plots were used to visualize hazard ratios (HR:s) and 95% confidence intervals (CI:s) and the overall measurement of effect.

The Kaplan-Meier method

In survival analysis, each studied subject has a study time. It could have begun when the subject entered the study or when a treatment was started. In our case, it was the date of diagnosis of breast cancer. The time ends when the endpoint is reached, i.e., when the event of interest occurs (for example death, recurrence, or distant metastasis): the study time also ends if the subject leaves the study due to a cause other than the event studied. If the time of the event is not reached (patients move away, refuse to continue in the study, or die of other causes), the study time is said to be censored. For patients who do not experience the event studied, the total survival time cannot be calculated. We do not know whether a censored patient would have experienced an event at a later time-point or not. The Kaplan-Meier method is a statistical procedure commonly used to calculate unbiased estimates of survival probabilities. The plot of these estimates, as a step-function of follow-up time, is known as a Kaplan-Meier graph. Such graphs were frequently used in the studies included in this thesis to illustrate differences in survival based on patterns of prognostic factors [214-216].

The log-rank test

The log-rank test is used to compare survival in groups of patients. The null hypothesis to be tested is that there is no difference in survival between groups, and the alternative is either that the survival curves differ (the default test) or that they differ and are ordered. For two groups, these alternatives are equivalent, but for >2 ordered groups, the log-rank test for trend has greater power and should therefore be used. Briefly, the observed and expected numbers of events in each group at each event time are compared and combined into test statistics, which are compared to a chi-square distribution with appropriate numbers of degrees of freedom. It can only be used for one variable at a time [215, 216].

Cox regression

The Cox proportional hazards regression model is commonly used to quantify relative effects, so-called hazard ratios (HR) of prognostic factors. A Cox regression model can be used to predict the hazard of the outcomes during the follow-up time. The relative effects of the factors in the model are assumed to be constant, i.e., independent of time, assumptions which must be checked. For instance, for a dichotomized prognostic factor with crossing survival curves, this model assumption is not fulfilled [215, 217].

Some of the advantages with Cox regression modeling, compared to the log-rank test, are that it provides effect measurements and not only *P*-values. Furthermore, variables measured on a continuous scale can be evaluated without categorization and adjustment for confounders is straightforward.

Schoenfeld's test

In Cox proportional hazards regression models, the effects are assumed to be proportional over time, and with Schoenfeld's test, this non-proportional hazards assumption can be formally tested [218].

Stratification in Cox regression models

Stratification allows for different mortality in different patient strata, e.g., studies, but equal effects of the factor relatively. It is used to ascertain that the estimated relative effect is a true effect and not an artifact due to the combination of patient series with different characteristics.

***P*-value**

A *P*-value is the probability of obtaining the outcome actually observed or an outcome that is even more extreme, assuming the null hypothesis to be true. If the *P*-value is less than a chosen critical value, the null hypothesis can be rejected, i.e., there is significant difference in predicted risk between the subgroups studied [215]. In this thesis, the typical applications of significance testing were to test whether survival curves were equal (log-rank) and to test factors effect on survival (HR = 1.00 means no effect). In this thesis, all of the *P*-values were two-sided and values less than 0.05 were considered significant.

C-index

Different measurements of predictive performance and model fit have been suggested in the literature [219]. We chose Harrell's concordance index (*C*). Harrell's *C* is a generalization of the area under the ROC curve for survival data: therefore, it measures how well a model discriminates between different responses, i.e., is the predicted response low for patients not experiencing the event and high for patients experiencing the event? A Cox-model with good discrimination has low predicted hazards for patients with long follow-up not experiencing an event and high hazards for patients experiencing an event, especially if the event occurs early. The index is defined as the proportion of all evaluable pairs of patients, for which the patient with the highest predicted hazard also has the longest survival time. A pair of patients both of whom have censored survival times is not evaluable, whereas a pair is evaluable in whom one of the patients has a low predicted hazard and a survival time that is censored at a later time-point than the time for the event of the other patient. A model leading to perfect prediction yields $C = 1.00$ whereas $C = 0.50$ is not better than guessing [220].

Fractional polynomial

Non-linear effects were modeled using fractional polynomial (FP) transformations [221, 222] in Cox regression models. Linear combinations of a series of 8 pre-defined power transformations (x^p) of the predictors of interest as covariates were compared, and the best fitting transformation of each predictor were chosen. Following the standard definition of FPs, the powers p were chosen from the set $(-2, -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, 2, 3)$, where x^0 is defined as $\ln(x)$. A first degree FP is hence defined as $\beta_0 + \beta_1 \cdot x^p$ and a second degree FP as a linear combination of two of the power transformations, e.g., $\beta_0 + \beta_1 \cdot x^{\frac{1}{2}} + \beta_2 \cdot x^2$. First degree FPs will guarantee

monotonicity, whereas second degree FPs are more flexible, allowing for either a maximum or a minimum. Higher order FPs will be even more flexible, but the number of evaluated transformations grows exponentially, leading to severe multiple testing problems. To avoid over-fitting, a function selection procedure based on a closed test procedure was used, which added flexibility, i.e., extra polynomial terms, only if the model fit improved significantly at the chosen overall significance level after adjustment for multiple testing [222].

Multivariable FP

The multivariable FP procedure (MFP), which is an extension of the function selection procedure based on FPs, was used to derive an FP-based prognostic index based on the originally continuous or integer valued predictors [222]. The MFP procedure applies a combination of backward and forward selection, starting from the full complex model, to find a parsimonious model, which is neither too simple nor too complex. It examines whether the effect of a continuous variable is better modeled by a non-linear member of the class of FP functions or by a linear function.

Restricted cubic splines

Non-linear covariate effects were also modeled using restricted cubic splines (RCS). For each covariate, k , so-called knots, was chosen, which uniquely defines $k-1$ polynomial transformations of the covariate. The definition of these transformations guaranteed that any linear combination of the $k-1$ spline variable would be linear before the first knot, a piecewise cubic polynomial between the adjacent knots, and linear again after the last knot. To avoid over-fitting, we decided to use five knots located at the 5th, 27.5th, 50th, 72.5th, and 95th percentiles, respectively, as recommended by Harrell [223]. This definition worked for age and tumor size: however, for the number of positive lymph nodes (a variable with almost 40% zeros), we chose to place the five knots at 1, 2, 3, 4, and 10 positive nodes, respectively.

Endpoints

In this thesis, distant disease-free survival (DDFS) was used as the endpoint in *Studies I* and *III*. In *Study IV*, we have called this endpoint DMFS (distant metastasis-free survival). Distant metastasis, but not locoregional recurrences was considered an event, and so was death by/of or with breast cancer. If a patient dies

due to breast cancer, most likely distant metastases existed: hence this is included in the endpoint. Contralateral breast cancer is not included as an event since we cannot know then whether a distant metastasis spread from the primary tumor or from the contralateral breast cancer and could thereby not be used to calculate time to endpoint. Death from other causes was not considered as an event.

For **Study II**, DDFS (as described above) was available for 1 305 of the patients. For 217 of these patients (one study cohort), we only had the definition breast cancer event: hence in **Study II**, all of these endpoints were compiled together and referred to as event-free survival. For the 332 patients in the case-control study included in **Study II**, the endpoint was breast cancer death.

Results

Study I

Cyclin A divided histological grade 2 tumors into two groups with significantly different DDFS (HR: 15, 95% CI: 4.3-52, $P < 0.001$). Furthermore, it was only in the ER-positive subgroup that cyclin A was a prognostic factor (HR: 5.8, 95% CI: 2.2–16), $P < 0.001$ vs. HR: 1.5, 95% CI: 0.55–3.9, $P = 0.44$ for the ER-negative subgroup). A strong, but not significant, interaction of cyclin A between ER-positive and ER-negative cases was identified (HR: 3.9, 95% CI: 0.98–16, $P = 0.054$).

Next we evaluated the combination variable CAGE, in which low CAGE consisted of all histological grade 1 cases and histological 2/ER-positive/low cyclin A cases. The high CAGE group was defined as histological grade 3, histological grade 2/ER-negative, or histological grade 2/ER-positive/high cyclin A. We found that CAGE was an independent prognostic factor for DDFS in multivariable analysis (HR: 4.1, 95% CI: 1.6-10, $P = 0.002$), adjusted for HER2 and age. Low CAGE and HER2-normal identified 53% as low-risk patients with a 5-year DDFS of 95% (95% CI: 89–98%). Furthermore, a strong correlation between cyclin A and genetic grade, as defined by Sotiriou [158], was identified ($P < 0.001$, χ^2).

Study II

Since Ki67 is the generally accepted proliferation marker in clinical practice, Ki67 was used to validate the index defined in *Study I*. Comparably, the low KiGE-group was defined as histological grade 1 tumors and grade 2/ER-positive/low Ki67. High KiGE consisted of all of the other cases. Irrespective of menopausal status, lymph node status or endocrine therapy or not, KiGE was found to subdivide patients into groups with significantly different predicted risks of recurrence. Furthermore, we confirmed that the prognostic effect of proliferation was most pronounced in ER-positive and histological grade 2 tumors.

To evaluate the importance of KiGE in patients with the same characteristics as in **Study I**, subgroups were created, with *Set 1* corresponding to **Study I**. The other subsets (*Sets 2–5*) were created to evaluate whether KiGE was prognostic in a wide range of subsets:

Set 1: node-negative (N0), no adjuvant therapy, ≤ 50 years old at diagnosis ($N = 169$, 20 with events), HR: 4.4 (95% CI: 1.4–13), $P = 0.010$;

Set 2: N0, no adjuvant therapy, >50 years old at diagnosis ($N = 488$, 55 with events), HR: 4.0 (95% CI: 2.2–7.2), $P < 0.001$;

Set 3: 1–3 positive lymph nodes (N1), no adjuvant therapy ($N = 167$, 39 with events), HR: 4.1 (95% CI: 1.9–8.6), $P < 0.001$;

Set 4: N0, adjuvant endocrine therapy ($N = 291$, 39 with events), HR: 3.0 (95% CI: 1.5–6.0), $P = 0.002$; and

Set 5: N1, adjuvant endocrine therapy ($N = 407$, 82 with events), HR: 4.1 (95% CI: 2.5–6.8), $P < 0.001$.

A significant association between event-free survival and high vs. low KiGE was found for all of the subsets and furthermore when including all 1 522 patients (HR: 3.9, 95% CI: 2.9–5.2, $P < 0.001$). A case-control study ($N = 166$ cases and 166 controls) was also included to corroborate the results further. For that study, breast cancer death was significantly associated with high vs. low KiGE (odds ratio: 2.7, 95% CI: 1.7–4.3, $P < 0.0001$). The results were similar in multivariable analysis adjusted for age at diagnosis, tumor size, and adjuvant endocrine treatment, and as well as when adjusting for HER2.

Among the subgroup of patients in whom DDFS was available ($N = 1\ 305$), KiGE identified a low risk group (constituting 53% of the patients) with a 5-year DDFS of 92% (95% CI: 89–93%).

Study III

The purpose was to evaluate the prognostic value of MAI, PPH3, cyclin B1, cyclin A, and Ki67, alone and in combinations. We confirmed the strong prognostic value of all of the studied proliferation factors and that combining two factors improved the effects. The importance of proliferation, especially in ER-positive breast cancer, was thus further verified.

In univariable analysis, high vs. low MAI was the strongest prognostic proliferation factor for 10-year DDFS (HR 3.3, 95% CI: 1.8–6.1, $P < 0.001$), followed by PPH3, cyclin A, Ki67, and cyclin B1. In the subgroup of ER-positive

tumors, a very strong prognostic effect of high MAI was found (HR: 7.0, 95% CI: 3.1–16, $P < 0.001$). No prognostic effect was found in ER-negative patients. MAI added prognostic value in histological grade 2 (HR: 7.2, 95% CI: 3.1–22, $P = 0.001$) and 1 (HR: 11, 95% CI: 2.3–55, $P = 0.003$) but not in histological grade 3. Furthermore, because the strongest prognostic value was found for MAI alone, one factor at a time was combined with MAI. When combining MAI and/or high cyclin A, a stronger prognostic relationship was found (HR 4.2, 95% CI: 2.2–7.0, $P < 0.001$) than for MAI in univariable analysis. Combinations of two proliferation factors were added to multivariable models, including age, ER-status, HER2-status, and adjuvant medical treatment. A high risk of developing distant recurrence was defined as having at least one of the two proliferation factors high. The combination of MAI and cyclin A showed a stronger prognostic effect on DDFS than when analyzing each factor separately (HR: 3.8, 95% CI: 1.6–8.7, $P = 0.002$).

Study IV

In the derivation set, risk stratification, based on models allowing for non-linear effects of continuous and integer-valued predictors, using FP models or RCS, was compared to stratification based on models for categorized predictors. Using FP transformations, univariable non-linear effects were detected for tumor size and the number of positive lymph nodes. For age, non-linear transformations did not improve the model fit significantly compared to the linear identity transformation chosen by the MFP procedure (C -index 0.513). The best fit for tumor size was a square root transformation (C -index 0.594), whereas for the number of positive lymph nodes, a combination of two polynomial terms provided the best fit (C -index 0.665). In multivariable analyses, categorization of each factor using two or three cut-points was found to improve prognostication compared to dichotomization (C -index 0.628 vs. 0.674). Modeling non-linear effects using MFP and RCS transformations modestly improved the C -index (0.695 vs. 0.696).

The model with categorized predictors generated 31 groups with different predicted relative hazards of distant metastases. In exploratory analyses, we investigated whether prognostication, within three homogenous risk groups according to the model with categorized predictors, could be improved using continuous predictors modeled with MFP transformations. As proof of principle, three groups with >100 patients and different levels of risk were chosen. A low-risk group (>50 years old, T1, and node-negative), an intermediate-risk group (>50 years old, T2, and 1–3 positive lymph nodes), and a high-risk group (>50 years old, T2, and at least ten positive lymph nodes) were identified. The estimated 10-

year DMFS was 92% for the patients within the low-risk group with the lowest predicted relative hazard, compared to 79% for the patients with the highest predicted relative hazard within this group. The corresponding 10-year DMFS based on categorized predictors was 88%.

The value of a prediction model is determined not by its performance in the dataset from which it was derived (the derivation set) but by its ability to perform well on independent validation data (the validation set). The Kalmar cohort ($N = 1\ 132$) was therefore used to validate risk stratification based on FP and RCS transformed predictors from the derivation set. Four multivariable prediction models fitted in the derivation set were evaluated: dichotomized, categorized, MFP, and RCS models. For each model, the same predictor transformations as in the derivation set was applied and weights estimated in the derivation set were used to calculate the values for the prognostic indices (PIs) for the patients in the validation set. The validation *C*-index for the model with dichotomized predictors was 0.675, 0.700 for the model with categorized predictors, and 0.705 respective 0.701 for models with FP or RCS transformed predictors.

The prognostic discrimination was further analyzed by calculation of HR:s for four groups based on the 16th, 50th, and 84th percentile following the recommendation by Royston and Altman [219]. The HR:s after MFP transformation of continuous predictors for the derivation and validation set showed similar appearance. The calibration of the high-risk group was good between validation and derivation, but poorer for the lower risk groups.

Discussion

More than 80% of breast cancer patients receive adjuvant treatment, although the majority is cured by surgery alone, leading to over-treatment [197]. Established clinical factors, such as hormone receptors, age at diagnosis, number of affected lymph nodes, tumor size, histological grade and HER2 are used in the clinic to sub-classify patients and to assign adequate treatment, but they could be improved for individual clinical decision-making. The advantage with new techniques, such as gene expression analysis, is the possibility of analyzing many genes at the same time and to finding gene patterns, instead of single gene expressions [63, 64, 158, 201, 224]. This ability might enable more individual-based treatments and large groups of patients could be divided into smaller and better-characterized groups that could benefit from individualized therapy. To identify tools to better guide the selection of patients for whom adjuvant chemotherapy might be avoided was number one of the top ten research questions in the web-based consultation performed by Dowsett et al., [200].

Before this thesis was initiated, it had been shown that gene expression profiles could be used to subdivide ER-positive breast cancer into groups with distinctively different prognosis and that the large proportion of tumors that are considered histological grade 2 (with intermediate risk for developing recurrence and hence difficulty in clinical decision making) using the Genomic Grade Index could be divided into one group with prognosis more similar to grade 1 and one group with prognosis more similar to grade 3 [158]. Furthermore, it was genes associated with proliferation that predominantly subdivided histological grade 2 [59, 63, 64, 158, 201]. It had also been shown that this subdivision of histological grade 2 could be performed using a single biomarker (Ki67) analyzed with IHC [61, 62].

We addressed these questions further in *Study I*. By combining cyclin A, histological grade, and ER into CAGE, and by creating an index based on CAGE and HER2, a subgroup of node-negative patients with low risk for recurrence (5-year DDFS of 95%, constituting 53% of the patients) for whom adjuvant chemotherapy might be avoided, was identified. The majority (95%) of the patients in the low-risk group were not administered adjuvant endocrine therapy, which was in accordance with standard procedure at that time of diagnosis (1991–1994). Today, the ER-positive patients in this low-risk group would have been recommended adjuvant endocrine therapy, which most likely would have

improved their 5-year DDFS even further. This low-risk group was characterized by low proliferation (low cyclin A), low histological grade (grade 1 and 2), and ER positivity, resembling luminal A breast cancer. The low recurrence rate for these patients suggests that adjuvant chemotherapy will have limited efficacy. In line with this suggestion, a recent study by Nielsen and co-workers concluded that patients with ‘luminal A-like’ tumors did not benefit from adjuvant cyclophosphamide-based chemotherapy [225]. With **Study I**, we corroborated the results that genes associated with proliferation could be used to subdivide ER-positive and histological grade 2 tumors into groups with different prognosis. Hence, our hypothesis that cyclin A could be used to subdivide histological grade 2 was fulfilled.

Although the staining of cyclin A was crisp and contrast-rich, we could not corroborate that evaluation of cyclin A was easier to perform than evaluation of Ki67 (kappa-value 0.71 for cyclin A compared to kappa-value >0.80 for Ki67) [61]. One explanation might be that the evaluation of cyclin A compared to Ki67 was performed by different persons with different levels of experience.

Based on the results of **Study I**, we wanted to confirm the importance of CAGE in a larger cohort, replacing cyclin A with the worldwide used factor Ki67 creating KiGE. In **Study II**, we evaluated this new index, KiGE, and showed that KiGE and CAGE performed equally well. Among node-negative patients, a low risk group constituting 57% of the patients had a 5-year DDFS of 92%, similar to the more expensive gene profile MammaPrint[®], in which the primary objective of the study was to achieve a 95% CI with a lower boundary of at least 92% [58]. For patients with 1–3 positive lymph nodes and with highly ER-positive tumors, chemotherapy can be recommended if there are other risk factors (i.e., Ki67 or cyclin A): otherwise, endocrine therapy might be sufficient [226]. Similar to the results in **Study II**, one of the aims of the MINDACT trial and the primary aim of RxPONDER trial are that within the group of patients with 1–3 positive lymph nodes, there exist patients with low risk of developing recurrence, a group of patients who might be spared adjuvant chemotherapy [39, 40].

It should be mentioned that all patients in the MINDACT study were given endocrine treatment, compared to only 8 (4%) patients in **Study I** and 824 (54%) in **Study II** received endocrine therapy.

The strength of **Study I** was that the material was homogenous with regard to age at diagnosis (all premenopausal), and all were node-negative. Moreover, only 21 patients received adjuvant chemotherapy. At the time of **Study I**, it was one of few studies including untreated patients. One weakness was that the number of patients included was small. We circumvented this weakness in **Study II** with a larger cohort. A downside with **Study II** was the mixture of endpoints, and that many small retrospective studies were included. One large prospective study perhaps

would have been better. In contrast, the mixture of studies could also be considered as strength of the study, and hence, the robustness of the KiGE-index was strengthened by the factors being evaluated in different health care regions and by different persons using different cut-offs. And furthermore, the prognostic value of the KiGE-index could be demonstrated in several subsets. We used TMA-cores in our analyses, and it can be argued that TMA only captures a small fraction of the tumor. However, several studies have shown similar associations with survival for TMA-cores as for whole sections (including Ki67) [95, 123, 227-230].

The Cox proportional hazards model assumes proportional hazards over time. To avoid problems with non-proportional hazards, the time to follow-up can be restricted. In the first studies, we chose five years of follow-up since the risk of recurrence is greatest during the first five years (and MINDACT used five years). However, since many breast cancer patients experience recurrence beyond five years, we extended the follow-up to ten years in the other studies. The main prognostic value lay in predicting early recurrence. So far, there are no known single factors to predict long-term recurrence [231]. The EndoPredict score is a tool that can provide additional information in identifying late distant metastases in the ER-positive/HER2-negative subgroup and might thereby select patients who need prolonged endocrine therapy [163].

At the St Gallen conference 2013, two years after the publication of *Study I*, an index including ER, PgR, Ki67, and HER2 was presented [36]. Had we included HER2 in our CAGE-index, the indices would have been very similar. A recent study showed that both genomic grade and centrally reviewed Ki67 could be used to improve clinicopathological models [232].

In the present thesis, prognostic factors with a focus on proliferation were studied. There are many different techniques and factors that can be used to study proliferation. Some demand large amounts of fresh frozen tissue (SPF and older versions of gene expression analyses): others have poor reproducibility (Ki67). The advantage with IHC is that, unlike genetic profiling, it allows for the selection of malignant cells only, and the analysis can be performed in areas of the section with the highest proliferation. Sampling of tissue for genetic profiling inevitably includes normal cells.

Mitotic counting is a part of histological grading, which is performed in clinical practice. However, apoptotic cells can be mistaken for mitotic cells. Using PPH3 circumvents this problem because it does not stain apoptotic cells and can, in that respect, support MAI assessments [136]. There is an inconsistency in defining cut-offs and choice of antibody. Different numbers of positive cells have been counted. Should cells in hot spot areas be counted or should an average be calculated? One explanation for the poor reproducibility of Ki67 evaluations could be intratumoral heterogeneity, with cells in the periphery having higher expression

of Ki67. This uneven distribution within the section has an impact on the evaluation method to choose [233]. Aleskandarany et al., showed that the highest expression hot spots best reflect the outcomes with lymph node metastasis as the endpoint [234]. Others do not recommend hot spots [235]. There is however no standard method on evaluating Ki67. A recent study evaluated nine different ways of assessing Ki67 and reported that Ki67 assessed on hotspots and in the periphery of the tumor resulted in higher Ki67 median values compared to measuring average Ki67 (27.45% vs. 16.96%) [236].

Another study suggested that different cut-offs for Ki67 should be used for TMA vs. whole sections [230]. However, Ki67 is prognostic over a broad spectrum of cut-offs [237]. A common cut-off is $\geq 20\%$ cells positive for Ki67; however, in Sweden, lab-specific cut-off values are used. The St Gallen 2015 conference recommends Ki67, with values 20–29% as cut-off, for discrimination between ‘luminal A-like’ and ‘luminal B-like’ [97]. Furthermore, recent results from our group showed that the value of Ki67 and PgR, for discrimination between luminal A- and B-like tumors seemed to be restricted to histological grade 2 [103].

To date, the evaluation of Ki67 has not been sufficiently stable to be included in the ASCO guidelines [37]. The utility of Ki67 will hopefully be improved by the introduction of standardized assessments and guidelines [96, 237, 238]. This is important since Ki67 is valuable, not only as a prognostic marker, but also for monitoring the response to neoadjuvant treatment [239].

The issues with reproducibility remain, although when applying national Swedish guidelines and recommendations for the evaluation of biomarkers, especially Ki67, good reproducibility can possibly be obtained [240]. To reduce interlaboratory and interobserver variations, repeated reproducibility studies are recommended [235, 238].

As mentioned above, prognostic factors have their pros and cons: nevertheless, studies have shown the prognostic value of proliferation regardless of the technique used. However, there is awareness that traditional biomarkers are insufficient in stratifying patients into appropriate risk groups for some of the reasons mentioned above. Using molecular profiles, a number of promising predictive models have been discovered: MammaPrint[®], Oncotype DX[®], Genomic Grade Index, and PAM50-ROR. These signatures can predict outcomes although they are based on different genes [58, 156, 158, 162]. This fact indicates that there exists an overlap of biological processes across gene patterns. Combining information from more than one gene signature would potentially increase the predictive power. Studies have shown that, by combining different genetic profiles, the prognostic potential can be increased [241, 242]. Other studies have shown the prognostic value of combining biomarkers [60, 243-247].

Using flow cytometry for analysis of the fraction of cells present in S-phase and, furthermore, combining the results with PgR-status and tumor size, members of our group showed, as early as 1990, a strong prognostic relationship by combining factors [87]. Furthermore, in 1999, Edén et al., showed that clinical markers performed as well as microarray gene expression profilers, and they suggested that a combination would improve prognosis [248]. The same year as the publication of **Study I**, Cuzick et al., published a combination of ER, PgR, Ki67 and HER2 into an IHC-score, which they called IHC4 [249]. According to them, IHC4 performed as well as the 21-gene recurrence score (i.e., *Oncotype DX*[®]) in predicting time to distant recurrence. The value in combining prognostic factors was also shown by Synnestvedt et al., who combined vascular invasion, histological grade, HER2, and Ki67. Similar to our results they found a large (61.4%) low-risk group with no unfavorable factors with only 3.6% distant relapses (median follow-up time was 86 months) who would not benefit from systemic adjuvant therapy [250].

Another combination-index is NPI+, which is a biomarker-based Nottingham prognostic index, as described by Rakha et al., and refined by Green et al., [251-253]. Regular NPI does not consider the biological heterogeneity. NPI+, which considers the IHC evaluation of ten biomarkers, classifies tumors into seven biological classes, similar to the intrinsic subtypes. These seven classes are then further classified into prognostic subgroups by adding information about prognostic clinicopathological variables (i.e., ER, PgR, cytokeratins, epidermal growth factor receptors, p53, and mucin), based on beta values generated in Cox regression analysis.

We hypothesized that, by combining two or more proliferation factors, the shortcomings previously mentioned could be substantially reduced. In **Study III**, we studied five factors for proliferation, separately and in combinations. We found that the strongest combination was MAI together with cyclin A. One possible explanation could be that they cover different phases of the cell cycle. MAI covers the M-phase, and cyclin A covers the S-phase. The results of **Study III** were based on a relatively small number of patients (~200 patients) and must be validated in a larger cohort: however, the results provide an indication that a combination of proliferation factors better predict survival. Due to a lack of power, we were not able to investigate whether a combination of more than two factors would have performed even better. In contrast, a combination of two factors for proliferation is very likely more practical in clinical routine than using three or more.

The results in **Study III** showed that MAI was a stronger prognostic factor than Ki67, especially when combined with cyclin A. Baak et al., reported similar results that MAI was the strongest predictor in node-negative breast cancer [55]. With the same focus on MAI and with standardized assessments and guidelines,

MAI might replace Ki67 as the prognostic marker for proliferation. Since MAI is already a part of histological grading, performed routinely on hematoxylin/eosin-slides, issues with choice of antibody, cut-off, hotspot or not, would be circumvented.

One way of improving the evaluation of Ki67 might be to use digital image analysis (DIA). The results have shown that DIA outperforms manual assessment [254]. In the luminal B subtype the results from DIA assessments of Ki67 were compared to manual assessments, and the DIA assessments performed the best. Others studies have shown similar result, with DIA outperforming manual scoring [255, 256].

To date, only a few molecular profiles/signatures have come into clinical use [257], and they have shed light on the inter-heterogeneity of breast cancer. The intrinsic subtypes have been refined [162, 258] and different combinations of profiles have been suggested [64, 259-262].

The first results from the MINDACT trial reported that 23.2% of patients (who did not receive chemotherapy) were assigned low genomic risk and high clinical risk (according to Adjuvant! Online) and had a 5-year DMFS of 94.7% (95% CI: 92.5 to 96.2). Thus, their primary objective of the study, to achieve a 95% CI with a lower boundary of at least 92%, was met [263]. Based on these results from the MINDACT study, patients classified as low genomic risk according to the 70-gene signature (i.e., MammaPrint[®]) could be recommended to forego chemotherapy. They had similar risk as those classified as low clinical and high genomic risk, hence MammaPrint[®] only added value to those classified as high clinical risk.

The TAILORx trial showed that the 21-gene recurrence score could select 16% of patients with low risk for metastasis (99% 5-year survival without distant metastasis) [40]. The results from chemotherapy vs. no chemotherapy in the intermediate group, as identified by the 21-gene assay, are still pending.

Nevertheless, these results are promising: however in countries with sparse economic resources, molecular profiling is still not an option and simpler and less expensive techniques, such as IHC, are especially important. Therefore, the St Gallen guidelines recommend ER, PgR, HER2, and Ki67 as IHC surrogate markers for intrinsic subtypes [36].

The use of prognostic factors in breast cancer has a long history dating back to the invention of the TNM classification system. Categorization of prognostic factors is intuitively appealing since the clinically relevant question is often to select between two (or a few) treatment options, but categorization of individual factors is not necessary for the construction of good prediction models [264]. Categorization is problematic since this will in general lead to information loss and hence lower power to detect true associations. Moreover, how should one

define the cut-offs? Should it be the median or some other percentile? Is it more correct to select the upper third of the patients as a cut-off, based on the argument that, in the long run, approximately one third of unselected breast cancer patients will develop recurrence? In addition, when applying a cut-off, patients on either side of the cut-off will be considered to have different predicted risks, rather than similar. Our aim was to show that prognostication could be improved by avoiding the categorization of risk factors and by allowing for non-linear effects using FP and RCS transformations. Sauerbei et al., evaluated the use of FP transformations in Cox modeling, comparing categorized with continuous prognostic factors, such as age, menopausal status, number of positive lymph nodes, ER, PgR, tumor size and grade. They concluded that important information could be extracted by allowing for FP transformations instead of traditional modeling [265]. Risk groups based on predictions from more flexible models of this type could be used to identify patients who could be spared adjuvant therapy and/or patients in need of additional treatment [265, 266].

In *Study IV*, more than 5 500 patients were compiled together and we observed that categorization into 3–4 groups, e.g., number of positive lymph nodes into N0, N1–3, N4–9, and N10+ instead of positive vs. negative, improved the prediction, as seen by increase in C-index. A multivariable model with age and tumor size in three categories and number of lymph nodes in four improved the C-index from 0.628, with the variables dichotomized in the model, to 0.674. Contradictory to our expectations, predictive performance was only moderately improved by using the predictors continuously. By increasing the number of cut-points, the possibly non-linear effects of the predictors seemed to be well captured. These findings strengthen the way age, tumor size, and number of lymph nodes are used in the clinic today.

Categorization is in general a more robust approach compared to FP or RCS transformations. However, using continuous values can have the advantage of making it possible to form risk groups of any size. Few publications so far have addressed this topic. Recently, Ejlersen et al., used FP transformations in developing a model for the prediction of excess mortality after adjuvant endocrine therapy [266]. Their model included age, tumor size, number of positive lymph nodes, and ER positivity in percentages. When using FP transformations instead of categorized variables, better identification of patients without excess mortality was found [266]. In contrast to our results, they showed that FP transformation performed better than models based on categorized variables. One explanation might be that their model included ER positivity in percentage and that their cohort was more homogenous (postmenopausal high-risk patients who all received five years of adjuvant endocrine therapy). In contrast to other studies, we did not see a non-linear effect for age [140-143]. This could perhaps be explained by, in our study, more than one third (23/69) of the patients younger than 35 years old

being treated with adjuvant chemotherapy, compared to 12% (511/4406) of the patients older than 35 years old. Furthermore, Fredholm et al., showed that differences in survival connected to age were primarily present in stage I and II breast cancers [140]. In *Study IV*, the majority of the patients were in stage II. Kroman et al., observed a similar age trend, with younger women having a higher risk of dying: however, this trend vanished in patients who received chemotherapy [143].

We modeled potentially non-linear relationships using both MFP and RCS transformations, and the results were strikingly similar. Both splines and FP transformation have their advantages and disadvantages [222]. FPs are more sensitive to outliers, which can be controlled for by restricting the degrees of freedom for each factor. On the other hand, RCS can lead to over-fitting of the data especially if many knots are used [222]. One advantage with the MFP analysis is that it stops itself from over-fitting the model. By incorporating prior knowledge into the statistical modeling strategy, it might be possible to avoid relationships that are not biologically true.

An alternative modeling strategy is artificial neural networks (ANNs). In a dataset, which to a large extent overlapped with the derivation set in *Study IV*, ANNs was applied. The results showed nearly identical performance between the ANN and Cox models [267]. The Cox and ANN models had almost the same *C*-indices, 0.71 vs. 0.70, which were comparable to the results in *Study IV*, with a *C*-index of 0.695 for the MFP-model.

As for *Study II*, a limitation with our study was that many different cohorts, instead of one large population-based cohort, were included. On the other hand, the results in the derivation set could be validated in another independent more population-based dataset. In the validation dataset, similar improvements in performance were obtained when more cut-points were used. The discrimination and calibration was found to be better for high-risk patients than for patients whose prognostic factors indicated lower risk. The relative effect estimates were found to be smaller when the models fitted in the derivation set were applied to the validation set. This could be explained by minor over-fitting to the derivation set.

In *Study IV*, only the three factors age at diagnosis, tumor size, and number of positive lymph nodes were included. A clinically applicable model should include all prognostic factors in use, i.e., according to current guidelines, also ER, PgR, HER2, Ki67, and histological grade [37, 97]. We did not have complete information for these additional factors: otherwise it might have been possible to create a nomogram that could be used for clinical decisions similar to the nomogram for DCIS or PREDICT, into which continuous values could be incorporated [42, 43, 268].

With the results from *Study IV*, the use of dichotomized factors in *Study I, II*, and *III* could be considered suboptimal. On the other hand, for *Study I* and *III* we did not have continuous values for all of the factors studied. Since *Study II* was a validation of the results in *Study I*, the factors were used accordingly.

Conclusions

In *Studies I* and *II*, a combination of a marker for proliferation (cyclin A or Ki67), histological grade, and estrogen receptor could identify over 50% of the patients as low risk, with more than 90% of the patients having 5-year event-free survival.

In *Study III*, a combination of MAI and cyclin A improved prognostication compared to using the factors individually.

In *Study IV*, dichotomization of prognostic factors measured on a continuous or integer-valued scale led to considerable information loss and should thus be avoided. Models allowing for non-linear effects did not outperform models with predictors categorized into three or four groups.

Future perspectives

Despite the use of single markers as surrogate makers for the intrinsic subtypes, the trend is toward the increased use of molecular profiles. My belief is that age at diagnosis, tumor size, and number of lymph nodes (the factors studied in *Study IV*) will be measured similarly in the near future, even after implementation of genetic profiles. The clinicopathological markers, however, might be replaced or combined with other techniques. One recent study showed that, when combining one gene predictor, i.e., PAM50-ROR, with established prognostic markers (tumor size and proliferation), the performance compared to pure molecular scores was improved [269]. Hence, a hybrid score, i.e., integrating a molecular signature such as PAM50-ROR (ProSigna) with clinicopathological factors, might increase the prognostic ability.

In Sweden, the South Sweden Cancerome Analysis Network – Breast (SCAN-B) consortium was launched in 2010 to address this challenge [270]. The aim is to establish infrastructure towards the implementation of genetic analyses in routine clinical practice. Tumors and blood (to date not paraffin-embedded material) are consecutively collected and molecularly classified. These classifications might in the future complement the clinicopathological factors and guide treatment decisions. A similar project is the Clinical Sequencing of Cancer in Sweden (Clinseq), which started in 2013 in Stockholm. The aim is to establish an infrastructure for genomic profiling and to perform classification of cancer. For both SCAN-B and Clinseq, the goal is to develop diagnostic tests based on molecular profiling, with the possibility of tailored treatment and hopefully thereby improving survival for all breast cancer patients. Furthermore, samples are collected in tumor banks for future projects.

Many of my projects would not have been possible if not for the large collection of tumor material and databases with patient characteristics and information about survival. For future research projects, collection of high quality tumor banks (eg., tissue, blood, paraffin blocks) and databases is greatly important. Moreover, to design studies in which the benefit of different new treatments for breast cancer can be evaluated is also important. Of additional importance is to design prospective studies in which a group of people is followed prior to developing breast cancer. Karma, the Karolinska Mammography Project for Risk Prediction of Breast Cancer, is one such study. Karma is a prevention study initiated in 2011

by researchers at the Karolinska Institute. The focus is primarily to identify high-risk women, based on lifestyle factors, genetic changes and mammographic morphology. More than 70 000 women have been included so far. If women at high risk for developing breast cancer could be identified in early stages, preventive treatments could be offered, and thereby the incidence and mortality in breast cancer could be reduced.

To identify markers that predict metastatic potential at an early stage is a challenge. Circulating tumor cells (CTCs) and circulating tumor DNA (ctDNA) in the blood are promising new markers in breast cancer. To analyze CTCs or ctDNA and relate this to molecular subtype could potentially provide information, which could guide individual treatment further.

Populärvetenskaplig sammanfattning

Bröstcancer är en av de vanligaste cancersjukdomarna bland kvinnor. I Sverige diagnostiseras mer än 8000 kvinnor varje år. Det betyder att varje timme drabbas någon av bröstcancer. Tack vare hälsoundersökningar som mammografi är majoriteten av tumörerna idag små när de upptäcks och oftast utan spridning till lymfkörtlarna i armhålan. Därmed är chansen till bot större och idag lever nästan 90% av patienterna fem år efter diagnosen. Den mest effektiva behandlingsformen är kirurgi, vilken botar majoriteten av patienterna. För att minska risken för återfall ger man tilläggsbehandling, så kallad adjuvant behandling, men eftersom en stor grupp redan är botade innebär detta ibland att för mycket behandling ges. Hur kan man då identifiera de patienter som inte behöver extra behandling? Och trots att det ges tilläggsbehandling är det fortfarande drygt 1400 kvinnor som varje år dör av bröstcancer. Därför är det även viktigt att hitta den grupp av patienter som underbehandlas d.v.s. de som får återfall och som möjligen hade behövt ytterligare eller annan typ av behandling. Mycket forskning ägnas åt att identifiera faktorer och åt försök att kombinera befintliga faktorer, för att använda dem som hjälp för att ge rätt behandling. Målet är att på så sätt bättre definiera och behandla de patienter som har hög risk för återfall.

För att bestämma vilken ytterligare behandling som ska ges använder man prognostiska och behandlingsprediktiva faktorer. Dessa faktorer korrelerar till sjukdomens förlopp, aggressivitet och känslighet för behandling. De viktigaste som används rutinmässigt är ålder vid diagnos, tumörstorlek, spridning till lymfkörtlar i armhålan, histologisk grad, uttryck av olika biomarkörer som östrogenreceptorn (ER), progesteronreceptorn (PgR), faktorer för tillväxthastighet, så kallad proliferation (Ki67), och uttryck av tillväxtfaktor receptorn (HER2). Histologisk grad och proliferation avspeglar cellens förmåga till okontrollerad celledelning och tillväxt. Histologisk grad delas in i tre grupper där grad 3 har sämst och grad 1 bäst prognos. För grad 2, som är den stora mellangruppen, behöver man andra verktyg för att kunna välja rätt behandling.

Nyligen har profiler som beskriver genuttryck börjat användas som ett försök att bättre gruppera patienter i olika riskgrupper baserat på beräknad risk. När projekten i denna avhandling initierades hade bröstcancer med hjälp av genuttryck grupperats i ER-positiva och ER-negativa tumörer. Dessa grupper har dessutom olika prognos. ER-positiva, histologisk grad 2 tumörer kunde delas i en grupp med

prognos mer lik grad 1 och en grupp med prognos mer lik grad 3. Gener involverade i proliferation var särskilt viktiga för denna uppdelning. Användandet av genuttrycks-profiler ökar, framför allt i USA. De är dock tekniskt krävande och kostsamma och inte tillgängliga för majoriteten av patienter i andra delar av världen. Andra studier har visat att en liknande uppdelning gick att göra med enklare tekniker och vi ville studera detta vidare.

I **Study I**, fann vi att kombinationen av en markör för proliferation (cyclin A), histologisk grad och östrogenreceptorn (ER), i ett index vi kallade CAGE tillsammans med HER2 identifierade en relativt stor lågrisk grupp (53%), med liten risk för fjärrmetastasering (5-års fjärrmetastas-fri överlevnad 95%). Studien omfattade bara drygt 200 patienter. Nästa steg blev därför att validera indexet i ett oberoende och större material. Detta gjordes i **Study II**, i vilket Ki67, som är den proliferationsmarkör som används mest i kliniken, användes i stället för cyclin A. Vi kombinerade faktorerna på samma sätt som i **Study I** och skapade indexet KiGE. I **Study II** såg vi att KiGE, precis som CAGE i **Study I**, identifierade en stor grupp av patienter med liten risk för återfall, som troligen inte behöver erbjudas cellgifter. I en webb-baserad konsultation var den viktigaste frågan just att hitta de patienter som inte behöver ges cellgifter.

Alla tekniker och markörer har sina för- och nackdelar. I **Study III** undersökte vi om en kombination av flera markörer för proliferation kunde ge ett säkrare mått för tumörens tillväxt och därmed också säkrare bestämma prognosen. Vi fann bl.a. att kombinationen av cyklin A och MAI (som är ett mått på själva celledelningen) hade större prognostisk effekt än om man tittade på faktorerna var för sig.

Det finns idag ett begränsat antal behandlingsalternativ och för att bestämma om behfaktorer i två eller flera grupper. I statistiska modeller är detta inte nödvändigt. Problem med att dela in i grupper, kategorisera, är bland annat hur man då väljer gränsvärden. Värden på var sida om gränsen blir automatiskt olika (hög respektive låg risk) i stället för snarlika.

I **Study IV** undersökte vi om det gav mer prognostisk information att låta värdena vara kontinuerliga så länge som möjligt i modellen. Vi såg att det definitivt var bättre att dela faktorerna i fler kategorier än två. Däremot blev det bara marginell skillnad om man behöll dem kontinuerliga i analysen. För den enskilde patienten kan det dock fortfarande göra skillnad och målet i dagens sjukvård är att hitta en så väl individanpassad behandling som möjligt.

Sammantaget visar studierna i denna avhandling att flera olika markörer för proliferation kan användas för att förutsäga sjukdomsförlopp och att en kombination av dem förstärker det prognostiska värdet. Med dessa kombinationer är det möjligt att identifiera patienter med så låg risk för återfall att nyttan av cellgifter är tveksam. Vi såg också att det är bättre att dela prognostiska faktorer i

fler grupper än två, men bara marginellt bättre att behålla dem kontinuerliga. Däremot, om man behåller faktorerna kontinuerliga så länge som möjligt i analysen, ökar möjligheten att skapa subgrupper och då också möjligheten att göra individuella beräkningar om risk för återfall.

Acknowledgments

Without support this work would not have been possible. I would like to express my special thanks to:

Mårten Fernö, my main supervisor and mentor, for encouraging me to perform this achievement. This could not have been done without you. Thank you for professional guidance throughout the work. For your never-ending enthusiasm and support. We both like the taste of blood in our mouths, but prefer the taste of wine.

Pär-Ola Bendahl, my co-supervisor, for introducing me into the field of statistics. I definitely know more now than when I started this journey, but I am far from your knowledge on the topic. Thank you for educating me with enthusiasm.

Ingrid Hedenfalk, my co-supervisor, aka lab-supervisor. Somehow the lab-project was replaced by statistics, but I needed you by my side nevertheless. I want to thank you for sharing your knowledge in molecular research, and your knowledge of the latest publications really impresses me.

Per Malmström, my co-supervisor on the clinical side. Thank you for sharing your never-ending knowledge in breast cancer. Willingly answering my questions and always with a smile.

Bo Baldetorp, Håkan Olsson, Mef Nilbert, Signe Borgquist, Lars Ekblad, and **Ingrid Wilson**, previous and present Heads of the Department of Oncology and Pathology. Thank you for creating an inspiring and creative research environment.

All my co-authors I thank for important contributions and collections of valuable databases. A special thanks to **Cecilia Ahlin, Dorthe Grabau** (who sadly passed away in December 2015), **Marie Klintman, Lisa Rydén, Olle Stål,** and **Marie Sundquist** for fruitful collaborations.

The participating departments of the **South** and **South-East Swedish Breast Cancer Groups**, of the **Uppsala-Örebro region**, and of the **Odense region** I thank for providing samples and clinical follow-up.

To all former and present PhD students connected to the Department. Thank you for interesting discussions at journal clubs and educational sessions. A special thanks to **Looket Dighe** and **Sara Jansson**, my present roommates. For different

reasons so conveniently leaving the room almost for myself, with lots of space to write this thesis. Hope to see you soon!

Susanne André and **Emelie Jalgén**, thank you for arranging all the practical aspects of performing a dissertation.

To all staff at the Department. I cannot mention you all, but you make this workplace enjoyable.

Kristina Lövgren for excellent skills in histopathological techniques. I miss sharing office with you. **Sara Baker** for taking me out of my thesis-bubble. Talking about both important and less important things. **Therese Sandberg**, thank you for cheering me on. I will keep on hitting you with deadlines! **Eva Rambech** for providing me with delicious tea.

Cecilia Hegardt, my lunchdate. Thank you for listening and being a good friend.

My parents, **Kerstin** and **Bertil** and my brother **Mattias**. Thank you for bringing me up in a stable and supporting home. Without you I would not be me. You are loving, caring, and always there for me.

Vilhelm and **Nina**, the bonus of my life. Thank you for all fun and exciting things we do together. Soon it's high time for "fredags-låten".

Oskar, the joy of my life. You brighten my day and take the edge off things and remind me of what really matters in life. "Åh, vi fiter i det då!"

And last, but definitely not the least! **Anders**, the love of my life. Life is so much better together with you. Thank you for believing in me, encouraging and supporting me, always. I love you endlessly.

All breast cancer patients included in the studies, thank your for participating. It would not have been possible without you.

This thesis was supported by funds from the Swedish Cancer Society, the Swedish Research Council, the Gunnar Nilsson Cancer Foundation, the Mrs. Berta Kamprad Foundation, the Anna and Edwin Bergers Foundation, the Swedish Breast Cancer Association (BRO), the Swedish Cancer and Allergy Foundation, the Skåne University Hospital Research Foundation, Skåne County Council's Research and Development Foundation, and Governmental Funding of Clinical Research within the National Health Service.

References

1. Hanahan D, Weinberg RA: **The hallmarks of cancer.** *Cell* 2000, **100**(1):57-70.
2. Hanahan D, Weinberg RA: **Hallmarks of cancer: the next generation.** *Cell* 2011, **144**(5):646-674.
3. Sherr CJ: **The Pezcoller lecture: cancer cell cycles revisited.** *Cancer Res* 2000, **60**(14):3689-3695.
4. Konigsberg R, Rogelsperger O, Jager W, Thalhammer T, Klimpfinger M, De Santis M, Hudc M, Dittrich C: **Cell cycle dysregulation influences survival in high risk breast cancer patients.** *Cancer investigation* 2008, **26**(7):734-740.
5. Norbury C, Nurse P: **Animal cell cycles and their control.** *Annual review of biochemistry* 1992, **61**:441-470.
6. Weinberg R: **The Biology of Cancer:** Garland Science, Taylor & Francis Group; 2007.
7. Hartwell LH, Weinert TA: **Checkpoints: controls that ensure the order of cell cycle events.** *Science* 1989, **246**(4930):629-634.
8. Socialstyrelsen: **Cancerincidens i Sverige.** 2014.
9. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, Mandelblatt JS, Yakovlev AY, Habbema JD, Feuer EJ, Cancer I, Surveillance Modeling Network C: **Effect of screening and adjuvant therapy on mortality from breast cancer.** *N Engl J Med* 2005, **353**(17):1784-1792.
10. Youlden DR, Cramb SM, Dunn NA, Muller JM, Pyke CM, Baade PD: **The descriptive epidemiology of female breast cancer: an international comparison of screening, incidence, survival and mortality.** *Cancer Epidemiol* 2012, **36**(3):237-248.
11. Early Breast Cancer Trialists' Collaborative G: **Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials.** *Lancet* 2005, **365**(9472):1687-1717.
12. Scharl A, Kuhn T, Papatthemelis T, Salterberg A: **The Right Treatment for the Right Patient - Personalised Treatment of Breast Cancer.** *Geburtshilfe Frauenheilkd* 2015, **75**(7):683-691.
13. Woods LM, Rachet B, Lambert PC, Coleman MP: **'Cure' from breast cancer among two populations of women followed for 23 years after diagnosis.** *Ann Oncol* 2009, **20**(8):1331-1336.
14. Baade PD, Youlden DR, Chambers SK: **When do I know I am cured? Using conditional estimates to provide better information about cancer survival prospects.** *The Medical journal of Australia* 2011, **194**(2):73-77.

15. Coleman MP, Forman D, Bryant H, Butler J, Rachet B, Maringe C, Nur U, Tracey E, Coory M, Hatcher J, McGahan CE, Turner D, Marrett L, Gjerstorff ML, Johannesen TB, Adolfsson J, Lambe M, Lawrence G, Meechan D, Morris EJ, Middleton R, Steward J, Richards MA, Group IMW: **Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995-2007 (the International Cancer Benchmarking Partnership): an analysis of population-based cancer registry data.** *Lancet* 2011, **377**(9760):127-138.
16. Foulkes WD: **Inherited susceptibility to common cancers.** *N Engl J Med* 2008, **359**(20):2143-2153.
17. McPherson K, Steel CM, Dixon JM: **ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics.** *BMJ* 2000, **321**(7261):624-628.
18. Loman N, Johannsson O, Kristoffersson U, Olsson H, Borg A: **Family history of breast and ovarian cancers and BRCA1 and BRCA2 mutations in a population-based series of early-onset breast cancer.** *J Natl Cancer Inst* 2001, **93**(16):1215-1223.
19. Stratton MR, Rahman N: **The emerging landscape of breast cancer susceptibility.** *Nat Genet* 2008, **40**(1):17-22.
20. Layde PM, Webster LA, Baughman AL, Wingo PA, Rubin GL, Ory HW: **The independent associations of parity, age at first full term pregnancy, and duration of breastfeeding with the risk of breast cancer. Cancer and Steroid Hormone Study Group.** *Journal of clinical epidemiology* 1989, **42**(10):963-973.
21. Collaborative Group on Hormonal Factors in Breast C: **Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease.** *Lancet* 2002, **360**(9328):187-195.
22. Ewertz M, Duffy SW, Adami HO, Kvale G, Lund E, Meirik O, Møller A, Soini I, Tulinius H: **Age at first birth, parity and risk of breast cancer: a meta-analysis of 8 studies from the Nordic countries.** *Int J Cancer* 1990, **46**(4):597-603.
23. Jernstrom H, Bendahl PO, Lidfeldt J, Nerbrand C, Agardh CD, Samsioe G: **A prospective study of different types of hormone replacement therapy use and the risk of subsequent breast cancer: the women's health in the Lund area (WHILA) study (Sweden).** *Cancer causes & control : CCC* 2003, **14**(7):673-680.
24. Fournier A, Berrino F, Riboli E, Avenel V, Clavel-Chapelon F: **Breast cancer risk in relation to different types of hormone replacement therapy in the E3N-EPIC cohort.** *Int J Cancer* 2005, **114**(3):448-454.
25. **Breast cancer and hormonal contraceptives: further results. Collaborative Group on Hormonal Factors in Breast Cancer.** *Contraception* 1996, **54**(3 Suppl):1S-106S.
26. Marchbanks PA, McDonald JA, Wilson HG, Folger SG, Mandel MG, Daling JR, Bernstein L, Malone KE, Ursin G, Strom BL, Norman SA, Wingo PA, Burkman RT, Berlin JA, Simon MS, Spirtas R, Weiss LK: **Oral contraceptives and the risk of breast cancer.** *N Engl J Med* 2002, **346**(26):2025-2032.
27. Nelson NJ: **Studies on how lifestyle factors may affect breast cancer risk and recurrence.** *J Natl Cancer Inst* 2012, **104**(8):574-576.

28. Hermansen C, Skovgaard Poulsen H, Jensen J, Langfeldt B, Steenskov V, Frederiksen P, Jensen OM: **Diagnostic reliability of combined physical examination, mammography, and fine-needle puncture ("triple-test") in breast tumors. A prospective study.** *Cancer* 1987, **60**(8):1866-1871.
29. Kaufman Z, Shpitz B, Shapiro M, Rona R, Lew S, Dinbar A: **Triple approach in the diagnosis of dominant breast masses: combined physical examination, mammography, and fine-needle aspiration.** *J Surg Oncol* 1994, **56**(4):254-257.
30. Chang JH, Vines E, Bertsch H, Fraker DL, Czerniecki BJ, Rosato EF, Lawton T, Conant EF, Orel SG, Schuchter L, Fox KR, Zieber N, Glick JH, Solin LJ: **The impact of a multidisciplinary breast cancer center on recommendations for patient management: the University of Pennsylvania experience.** *Cancer* 2001, **91**(7):1231-1237.
31. **Nationellt Vårdprogram Bröstcancer 2014**
[<http://www.swebcg.se/wp-content/uploads/2016/09/NVP-Bröstcancer-2014.pdf>]
32. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, Jong RA, Hislop G, Chiarelli A, Minkin S, Yaffe MJ: **Mammographic density and the risk and detection of breast cancer.** *N Engl J Med* 2007, **356**(3):227-236.
33. Checka CM, Chun JE, Schnabel FR, Lee J, Toth H: **The relationship of mammographic density and age: implications for breast cancer screening.** *AJR American journal of roentgenology* 2012, **198**(3):W292-295.
34. Hayes DF, Isaacs C, Stearns V: **Prognostic factors in breast cancer: current and new predictors of metastasis.** *Journal of mammary gland biology and neoplasia* 2001, **6**(4):375-392.
35. Cianfrocca M, Goldstein LJ: **Prognostic and predictive factors in early-stage breast cancer.** *Oncologist* 2004, **9**(6):606-616.
36. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thurlimann B, Senn HJ, Panel m: **Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013.** *Ann Oncol* 2013, **24**(9):2206-2223.
37. Harris LN, Ismaila N, McShane LM, Andre F, Collyar DE, Gonzalez-Angulo AM, Hammond EH, Kuderer NM, Liu MC, Mennel RG, Van Poznak C, Bast RC, Hayes DF, American Society of Clinical O: **Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline.** *J Clin Oncol* 2016, **34**(10):1134-1150.
38. Cardoso F, Piccart-Gebhart M, Van't Veer L, Rutgers E, Consortium T: **The MINDACT trial: the first prospective clinical validation of a genomic tool.** *Mol Oncol* 2007, **1**(3):246-251.
39. Sparano JA: **TAILORx: trial assigning individualized options for treatment (Rx).** *Clin Breast Cancer* 2006, **7**(4):347-350.
40. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, Geyer CE, Jr., Dees EC, Perez EA, Olson JA, Jr., Zujewski J, Lively T, Badve SS, Saphner TJ, Wagner LI, Whelan TJ, Ellis MJ, Paik S, Wood WC, Ravdin P, Keane MM, Gomez

- Moreno HL, Reddy PS, Goggins TF, Mayer IA, Brufsky AM, Toppmeyer DL, Kaklamani VG, Atkins JN, Berenberg JL, Sledge GW: **Prospective Validation of a 21-Gene Expression Assay in Breast Cancer.** *N Engl J Med* 2015, **373**(21):2005-2014.
41. Ravdin PM, Siminoff LA, Davis GJ, Mercer MB, Hewlett J, Gerson N, Parker HL: **Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer.** *J Clin Oncol* 2001, **19**(4):980-991.
 42. Wishart GC, Azzato EM, Greenberg DC, Rashbass J, Kearins O, Lawrence G, Caldas C, Pharoah PD: **PREDICT: a new UK prognostic model that predicts survival following surgery for invasive breast cancer.** *Breast Cancer Res* 2010, **12**(1):R1.
 43. Wishart GC, Bajdik CD, Dicks E, Provenzano E, Schmidt MK, Sherman M, Greenberg DC, Green AR, Gelmon KA, Kosma VM, Olson JE, Beckmann MW, Winqvist R, Cross SS, Severi G, Huntsman D, Pylkas K, Ellis I, Nielsen TO, Giles G, Blomqvist C, Fasching PA, Couch FJ, Rakha E, Foulkes WD, Blows FM, Begin LR, van't Veer LJ, Southey M, Nevanlinna H, Mannermaa A, Cox A, Cheang M, Baglietto L, Caldas C, Garcia-Closas M, Pharoah PD: **PREDICT Plus: development and validation of a prognostic model for early breast cancer that includes HER2.** *Br J Cancer* 2012, **107**(5):800-807.
 44. Henderson R, Keiding N: **Individual survival time prediction using statistical models.** *Journal of medical ethics* 2005, **31**(12):703-706.
 45. Bukholm IR, Bukholm G, Nesland JM: **Over-expression of cyclin A is highly associated with early relapse and reduced survival in patients with primary breast carcinomas.** *Int J Cancer* 2001, **93**(2):283-287.
 46. Kuhling H, Alm P, Olsson H, Ferno M, Baldetorp B, Parwaresch R, Rudolph P: **Expression of cyclins E, A, and B, and prognosis in lymph node-negative breast cancer.** *J Pathol* 2003, **199**(4):424-431.
 47. van Diest PJ, van der Wall E, Baak JP: **Prognostic value of proliferation in invasive breast cancer: a review.** *J Clin Pathol* 2004, **57**(7):675-681.
 48. Colozza M, Azambuja E, Cardoso F, Sotiriou C, Larsimont D, Piccart MJ: **Proliferative markers as prognostic and predictive tools in early breast cancer: where are we now?** *Ann Oncol* 2005, **16**(11):1723-1739.
 49. Baldini E, Camerini A, Sgambato A, Prochilo T, Capodanno A, Pasqualetti F, Orlandini C, Resta L, Bevilacqua G, Collecchi P: **Cyclin A and E2F1 overexpression correlate with reduced disease-free survival in node-negative breast cancer patients.** *Anticancer Res* 2006, **26**(6B):4415-4421.
 50. Beresford MJ, Wilson GD, Makris A: **Measuring proliferation in breast cancer: practicalities and applications.** *Breast Cancer Res* 2006, **8**(6):216.
 51. Whitfield ML, George LK, Grant GD, Perou CM: **Common markers of proliferation.** *Nature reviews Cancer* 2006, **6**(2):99-106.
 52. Desmedt C, Sotiriou C: **Proliferation: the most prominent predictor of clinical outcome in breast cancer.** *Cell Cycle* 2006, **5**(19):2198-2202.
 53. de Azambuja E, Cardoso F, de Castro G, Jr., Colozza M, Mano MS, Durbecq V, Sotiriou C, Larsimont D, Piccart-Gebhart MJ, Paesmans M: **Ki-67 as prognostic**

- marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients.** *Br J Cancer* 2007, **96**(10):1504-1513.
54. Stuart-Harris R, Caldas C, Pinder SE, Pharoah P: **Proliferation markers and survival in early breast cancer: a systematic review and meta-analysis of 85 studies in 32,825 patients.** *Breast* 2008, **17**(4):323-334.
 55. Baak JP, Gudlaugsson E, Skaland I, Guo LH, Klos J, Lende TH, Soiland H, Janssen EA, Zur Hausen A: **Proliferation is the strongest prognosticator in node-negative breast cancer: significance, error sources, alternatives and comparison with molecular prognostic markers.** *Breast Cancer Res Treat* 2009, **115**(2):241-254.
 56. Patani N, Martin LA, Dowsett M: **Biomarkers for the clinical management of breast cancer: international perspective.** *Int J Cancer* 2013, **133**(1):1-13.
 57. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslén LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D: **Molecular portraits of human breast tumours.** *Nature* 2000, **406**(6797):747-752.
 58. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH: **Gene expression profiling predicts clinical outcome of breast cancer.** *Nature* 2002, **415**(6871):530-536.
 59. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, Martiat P, Fox SB, Harris AL, Liu ET: **Breast cancer classification and prognosis based on gene expression profiles from a population-based study.** *Proc Natl Acad Sci U S A* 2003, **100**(18):10393-10398.
 60. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, Watson M, Davies S, Bernard PS, Parker JS, Perou CM, Ellis MJ, Nielsen TO: **Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer.** *J Natl Cancer Inst* 2009, **101**(10):736-750.
 61. Klintman M, Bendahl PO, Grabau D, Lovgren K, Malmstrom P, Ferno M: **The prognostic value of Ki67 is dependent on estrogen receptor status and histological grade in premenopausal patients with node-negative breast cancer.** *Mod Pathol* 2010, **23**(2):251-259.
 62. Aleskandarany MA, Rakha EA, Macmillan RD, Powe DG, Ellis IO, Green AR: **MIB1/Ki-67 labelling index can classify grade 2 breast cancer into two clinically distinct subgroups.** *Breast cancer research and treatment* 2011, **127**(3):591-599.
 63. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P, Borresen-Dale AL: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proc Natl Acad Sci U S A* 2001, **98**(19):10869-10874.
 64. Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, Desmedt C, Ignatiadis M, Sengstag T, Schütz F, Goldstein DR, Piccart M, Delorenzi M: **Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures.** *Breast Cancer Res* 2008, **10**(4):R65.

65. MacGrogan G, Mauriac L, Durand M, Bonichon F, Trojani M, de Mascarel I, Coindre JM: **Primary chemotherapy in breast invasive carcinoma: predictive value of the immunohistochemical detection of hormonal receptors, p53, c-erbB-2, MiB1, pS2 and GST pi.** *Br J Cancer* 1996, **74**(9):1458-1465.
66. Silvestrini R, Luisi A, Zambetti M, Cipriani S, Valagussa P, Bonadonna G, Daidone MG: **Cell proliferation and outcome following doxorubicin plus CMF regimens in node-positive breast cancer.** *Int J Cancer* 2000, **87**(3):405-411.
67. Elston CW, Ellis IO: **Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up.** *Histopathology* 1991, **19**(5):403-410.
68. **International Agency for Research on Cancer: Pathology and Genetics of Tumours of the Breast and Female Genital Organs**
[<http://www.iarc.fr/en/publications/pdfs-online/pat-gen/bb4/>]
69. WHO: **Pathology and genetics of tumours of the breast and female genital organs.** *IARC Press, Lyon* 2003.
70. Baak JP, van Diest PJ, Voorhorst FJ, van der Wall E, Beex LV, Vermorken JB, Janssen EA: **Prospective multicenter validation of the independent prognostic value of the mitotic activity index in lymph node-negative breast cancer patients younger than 55 years.** *J Clin Oncol* 2005, **23**(25):5993-6001.
71. Gudlaugsson E, Skaland I, Janssen EA, van Diest PJ, Voorhorst FJ, Kjellevoid K, zur Hausen A, Baak JP: **Prospective multicenter comparison of proliferation and other prognostic factors in lymph node negative lobular invasive breast cancer.** *Breast Cancer Res Treat* 2010, **121**(1):35-40.
72. Lende TH, Janssen EA, Gudlaugsson E, Voorhorst F, Smaaland R, van Diest P, Soiland H, Baak JP: **In patients younger than age 55 years with lymph node-negative breast cancer, proliferation by mitotic activity index is prognostically superior to adjuvant!** *J Clin Oncol* 2011, **29**(7):852-858.
73. Rank F, Dombernowsky P, Jespersen NC, Pedersen BV, Keiding N: **Histologic malignancy grading of invasive ductal breast carcinoma. A regression analysis of prognostic factors in low-risk carcinomas from a multicenter trial.** *Cancer* 1987, **60**(6):1299-1305.
74. Harvey JM, de Klerk NH, Sterrett GF: **Histological grading in breast cancer: interobserver agreement, and relation to other prognostic factors including ploidy.** *Pathology* 1992, **24**(2):63-68.
75. Boiesen P, Bendahl PO, Anagnostaki L, Domanski H, Holm E, Idvall I, Johansson S, Ljungberg O, Ringberg A, Ostberg G, Ferno M: **Histologic grading in breast cancer--reproducibility between seven pathologic departments. South Sweden Breast Cancer Group.** *Acta Oncol* 2000, **39**(1):41-45.
76. Medri L, Volpi A, Nanni O, Vecchi AM, Mangia A, Schittulli F, Padovani F, Giunchi DC, Zito A, Amadori D, Paradiso A, Silvestrini R: **Prognostic relevance of mitotic activity in patients with node-negative breast cancer.** *Mod Pathol* 2003, **16**(11):1067-1075.

77. Al-Janabi S, van Slooten HJ, Visser M, van der Ploeg T, van Diest PJ, Jiwa M: **Evaluation of mitotic activity index in breast cancer using whole slide digital images.** *PLoS One* 2013, **8**(12):e82576.
78. Meyer JS, Friedman E, McCrate MM, Bauer WC: **Prediction of early course of breast carcinoma by thymidine labeling.** *Cancer* 1983, **51**(10):1879-1886.
79. Silvestrini R, Daidone MG, Valagussa P, Di Fronzo G, Mezzanotte G, Mariani L, Bonadonna G: **3H-thymidine-labeling index as a prognostic indicator in node-positive breast cancer.** *J Clin Oncol* 1990, **8**(8):1321-1326.
80. Silvestrini R, Daidone MG, Luisi A, Boracchi P, Mezzetti M, Di Fronzo G, Andreola S, Salvadori B, Veronesi U: **Biologic and clinicopathologic factors as indicators of specific relapse types in node-negative breast cancer.** *J Clin Oncol* 1995, **13**(3):697-704.
81. Silvestrini R, Daidone MG, Luisi A, Mastore M, Leutner M, Salvadori B: **Cell proliferation in 3,800 node-negative breast cancers: consistency over time of biological and clinical information provided by 3H-thymidine labelling index.** *Int J Cancer* 1997, **74**(1):122-127.
82. Volpi A, De Paola F, Nanni O, Granato AM, Bajorko P, Becciolini A, Scarpi E, Riccobon A, Balzi M, Amadori D: **Prognostic significance of biologic markers in node-negative breast cancer patients: a prospective study.** *Breast Cancer Res Treat* 2000, **63**(3):181-192.
83. Moran RE, Black MM, Alpert L, Straus MJ: **Correlation of cell-cycle kinetics, hormone receptors, histopathology, and nodal status in human breast cancer.** *Cancer* 1984, **54**(8):1586-1590.
84. Clark GM, Dressler LG, Owens MA, Pounds G, Oldaker T, McGuire WL: **Prediction of relapse or survival in patients with node-negative breast cancer by DNA flow cytometry.** *N Engl J Med* 1989, **320**(10):627-633.
85. Hedley DW, Clark GM, Cornelisse CJ, Killander D, Kute T, Merkel D: **Consensus review of the clinical utility of DNA cytometry in carcinoma of the breast. Report of the DNA Cytometry Consensus Conference.** *Cytometry* 1993, **14**(5):482-485.
86. Klintman M, Nilsson F, Bendahl PO, Ferno M, Liljegren G, Emdin S, Malmstrom P, Swedish Breast Cancer G: **A prospective, multicenter validation study of a prognostic index composed of S-phase fraction, progesterone receptor status, and tumour size predicts survival in node-negative breast cancer patients: NNBC, the node-negative breast cancer trial.** *Ann Oncol* 2013, **24**(9):2284-2291.
87. Sigurdsson H, Baldetorp B, Borg A, Dalberg M, Ferno M, Killander D, Olsson H: **Indicators of prognosis in node-negative breast cancer.** *N Engl J Med* 1990, **322**(15):1045-1053.
88. Gerdes J, Schwab U, Lemke H, Stein H: **Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation.** *Int J Cancer* 1983, **31**(1):13-20.
89. Lopez F, Belloc F, Lacombe F, Dumain P, Reiffers J, Bernard P, Boisseau MR: **Modalities of synthesis of Ki67 antigen during the stimulation of lymphocytes.** *Cytometry* 1991, **12**(1):42-49.

90. Scholzen T, Gerdes J: **The Ki-67 protein: from the known and the unknown.** *Journal of cellular physiology* 2000, **182**(3):311-322.
91. Bullwinkel J, Baron-Luhr B, Ludemann A, Wohlenberg C, Gerdes J, Scholzen T: **Ki-67 protein is associated with ribosomal RNA transcription in quiescent and proliferating cells.** *Journal of cellular physiology* 2006, **206**(3):624-635.
92. Rahmzadeh R, Huttmann G, Gerdes J, Scholzen T: **Chromophore-assisted light inactivation of pKi-67 leads to inhibition of ribosomal RNA synthesis.** *Cell proliferation* 2007, **40**(3):422-430.
93. Urruticoechea A, Smith IE, Dowsett M: **Proliferation marker Ki-67 in early breast cancer.** *J Clin Oncol* 2005, **23**(28):7212-7220.
94. Inwald EC, Klinkhammer-Schalke M, Hofstadter F, Zeman F, Koller M, Gerstenhauer M, Ortman O: **Ki-67 is a prognostic parameter in breast cancer patients: results of a large population-based cohort of a cancer registry.** *Breast Cancer Res Treat* 2013, **139**(2):539-552.
95. Yerushalmi R, Woods R, Ravdin PM, Hayes MM, Gelmon KA: **Ki67 in breast cancer: prognostic and predictive potential.** *Lancet Oncol* 2010, **11**(2):174-183.
96. Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, Ellis M, Henry NL, Hugh JC, Lively T, McShane L, Paik S, Penault-Llorca F, Prudkin L, Regan M, Salter J, Sotiriou C, Smith IE, Viale G, Zujewski JA, Hayes DF: **Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group.** *Journal of the National Cancer Institute* 2011, **103**(22):1656-1664.
97. Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnani M, Piccart-Gebhart M, Thurlimann B, Senn HJ, Panel M: **Tailoring therapies--improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015.** *Ann Oncol* 2015, **26**(8):1533-1546.
98. Polley MY, Leung SC, McShane LM, Gao D, Hugh JC, Mastropasqua MG, Viale G, Zabaglo LA, Penault-Llorca F, Bartlett JM, Gown AM, Symmans WF, Piper T, Mehl E, Enos RA, Hayes DF, Dowsett M, Nielsen TO, International Ki67 in Breast Cancer Working Group of the Breast International G, North American Breast Cancer G: **An international Ki67 reproducibility study.** *J Natl Cancer Inst* 2013, **105**(24):1897-1906.
99. Ekholm M, Beglerbegovic S, Grabau D, Lovgren K, Malmstrom P, Hartman L, Ferno M: **Immunohistochemical assessment of Ki67 with antibodies SP6 and MIB1 in primary breast cancer: a comparison of prognostic value and reproducibility.** *Histopathology* 2014, **65**(2):252-260.
100. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ: **Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011.** *Ann Oncol* 2011, **22**(8):1736-1747.
101. Luporsi E, Andre F, Spyrtos F, Martin PM, Jacquemier J, Penault-Llorca F, Tubiana-Mathieu N, Sigal-Zafrani B, Arnould L, Gompel A, Egele C, Poulet B, Clough KB, Crouet H, Fourquet A, Lefranc JP, Mathelin C, Rouyer N, Serin D,

- Spielmann M, Haugh M, Chenard MP, Brain E, de Cremoux P, Bellocq JP: **Ki-67: level of evidence and methodological considerations for its role in the clinical management of breast cancer: analytical and critical review.** *Breast Cancer Res Treat* 2012, **132**(3):895-915.
102. Varga Z, Diebold J, Dommann-Scherrer C, Frick H, Kaup D, Noske A, Obermann E, Ohlschlegel C, Padberg B, Rakozy C, Sancho Oliver S, Schobinger-Clement S, Schreiber-Facklam H, Singer G, Tapia C, Wagner U, Mastropasqua MG, Viale G, Lehr HA: **How reliable is Ki-67 immunohistochemistry in grade 2 breast carcinomas? A QA study of the Swiss Working Group of Breast- and Gynecopathologists.** *PLoS One* 2012, **7**(5):e37379.
 103. Ehinger A MP, Bendahl P-O, Elston C, Falck A-K, Forsare C, Grabau D, Rydén L, Stål O, Fernö M: **Histological grade provides significant prognostic information in addition to breast cancer subtypes defined according to St Gallen 2013 [In press].** *Acta Oncol* 2016.
 104. Seshadri R, Lee CS, Hui R, McCaul K, Horsfall DJ, Sutherland RL: **Cyclin D1 amplification is not associated with reduced overall survival in primary breast cancer but may predict early relapse in patients with features of good prognosis.** *Clin Cancer Res* 1996, **2**(7):1177-1184.
 105. Aaltonen K, Amini RM, Landberg G, Eerola H, Aittomaki K, Heikkila P, Nevanlinna H, Blomqvist C: **Cyclin D1 expression is associated with poor prognostic features in estrogen receptor positive breast cancer.** *Breast Cancer Res Treat* 2009, **113**(1):75-82.
 106. Lehn S, Tobin NP, Berglund P, Nilsson K, Sims AH, Jirstrom K, Harkonen P, Lamb R, Landberg G: **Down-regulation of the oncogene cyclin D1 increases migratory capacity in breast cancer and is linked to unfavorable prognostic features.** *Am J Pathol* 2010, **177**(6):2886-2897.
 107. Guo L, Liu S, Jakulin A, Yilamu D, Wang B, Yan J: **Positive expression of cyclin D1 is an indicator for the evaluation of the prognosis of breast cancer.** *International journal of clinical and experimental medicine* 2015, **8**(10):18656-18664.
 108. Yang C, Nan K, Zhang Y, Chen Y, Qin S: **High expression of cyclin D1 is correlated with the expression of estrogen receptor and good prognosis in breast cancer.** *Xi bao yu fen zi mian yi xue za zhi = Chinese journal of cellular and molecular immunology* 2016, **32**(1):84-87.
 109. Bieche I, Olivi M, Nogues C, Vidaud M, Lidereau R: **Prognostic value of CCND1 gene status in sporadic breast tumours, as determined by real-time quantitative PCR assays.** *Br J Cancer* 2002, **86**(4):580-586.
 110. Stendahl M, Kronblad A, Ryden L, Emdin S, Bengtsson NO, Landberg G: **Cyclin D1 overexpression is a negative predictive factor for tamoxifen response in postmenopausal breast cancer patients.** *Br J Cancer* 2004, **90**(10):1942-1948.
 111. Rogers S, Gloss BS, Lee CS, Sergio CM, Dinger ME, Musgrove EA, Burgess A, Caldon CE: **Cyclin E2 is the predominant E-cyclin associated with NPAT in breast cancer cells.** *Cell division* 2015, **10**:1.

112. Gudas JM, Payton M, Thukral S, Chen E, Bass M, Robinson MO, Coats S: **Cyclin E2, a novel G1 cyclin that binds Cdk2 and is aberrantly expressed in human cancers.** *Mol Cell Biol* 1999, **19**(1):612-622.
113. Keyomarsi K, Tucker SL, Buchholz TA, Callister M, Ding Y, Hortobagyi GN, Bedrosian I, Knickerbocker C, Toyofuku W, Lowe M, Herliczek TW, Bacus SS: **Cyclin E and survival in patients with breast cancer.** *N Engl J Med* 2002, **347**(20):1566-1575.
114. Lindahl T, Landberg G, Ahlgren J, Nordgren H, Norberg T, Klaar S, Holmberg L, Bergh J: **Overexpression of cyclin E protein is associated with specific mutation types in the p53 gene and poor survival in human breast cancer.** *Carcinogenesis* 2004, **25**(3):375-380.
115. Berglund P, Stighall M, Jirstrom K, Borgquist S, Sjolander A, Hedenfalk I, Landberg G: **Cyclin E overexpression obstructs infiltrative behavior in breast cancer: a novel role reflected in the growth pattern of medullary breast cancers.** *Cancer Res* 2005, **65**(21):9727-9734.
116. Berglund P, Landberg G: **Cyclin e overexpression reduces infiltrative growth in breast cancer: yet another link between proliferation control and tumor invasion.** *Cell Cycle* 2006, **5**(6):606-609.
117. Berglund P, Stighall M, Jirstrom K, Ryden L, Ferno M, Nordenskjold B, Landberg G: **Cyclin E confers a prognostic value in premenopausal breast cancer patients with tumours exhibiting an infiltrative growth pattern.** *J Clin Pathol* 2008, **61**(2):184-191.
118. Ahlin C, Gruhne B, Holmqvist M, Zetterberg A, Fjallskog ML: **Aberrant expression of cyclin E in low-risk node negative breast cancer.** *Acta Oncol* 2008, **47**(8):1539-1545.
119. Karakas C, Biernacka A, Bui T, Sahin AA, Yi M, Akli S, Schafer J, Alexander A, Adjapong O, Hunt KK, Keyomarsi K: **Cytoplasmic Cyclin E and Phospho-Cyclin-Dependent Kinase 2 Are Biomarkers of Aggressive Breast Cancer.** *Am J Pathol* 2016, **186**(7):1900-1912.
120. Yam CH, Fung TK, Poon RY: **Cyclin A in cell cycle control and cancer.** *Cell Mol Life Sci* 2002, **59**(8):1317-1326.
121. Michalides R, van Tinteren H, Balkenende A, Vermorken JB, Benraadt J, Huldij J, van Diest P: **Cyclin A is a prognostic indicator in early stage breast cancer with and without tamoxifen treatment.** *Br J Cancer* 2002, **86**(3):402-408.
122. Poikonen P, Sjostrom J, Amini RM, Villman K, Ahlgren J, Blomqvist C: **Cyclin A as a marker for prognosis and chemotherapy response in advanced breast cancer.** *Br J Cancer* 2005, **93**(5):515-519.
123. Aaltonen K, Ahlin C, Amini RM, Salonen L, Fjallskog ML, Heikkila P, Nevanlinna H, Blomqvist C: **Reliability of cyclin A assessment on tissue microarrays in breast cancer compared to conventional histological slides.** *Br J Cancer* 2006, **94**(11):1697-1702.
124. Ahlin C, Aaltonen K, Amini RM, Nevanlinna H, Fjallskog ML, Blomqvist C: **Ki67 and cyclin A as prognostic factors in early breast cancer. What are the optimal cut-off values?** *Histopathology* 2007, **51**(4):491-498.

125. Li L, Mu K, Zhou G, Lan L, Auer G, Zetterberg A: **Genomic instability and proliferative activity as risk factors for distant metastases in breast cancer.** *Br J Cancer* 2008, **99**(3):513-519.
126. Ahlin C, Zhou W, Holmqvist M, Holmberg L, Nilsson C, Jirstrom K, Blomqvist C, Amini RM, Fjallskog ML: **Cyclin A is a proliferative marker with good prognostic value in node-negative breast cancer.** *Cancer Epidemiol Biomarkers Prev* 2009, **18**(9):2501-2506.
127. Gong D, Ferrell JE, Jr.: **The roles of cyclin A2, B1, and B2 in early and late mitotic events.** *Molecular biology of the cell* 2010, **21**(18):3149-3161.
128. Aaltonen K, Amini RM, Heikkila P, Aittomaki K, Tamminen A, Nevanlinna H, Blomqvist C: **High cyclin B1 expression is associated with poor survival in breast cancer.** *Br J Cancer* 2009, **100**(7):1055-1060.
129. Nimeus-Malmstrom E, Koliadi A, Ahlin C, Holmqvist M, Holmberg L, Amini RM, Jirstrom K, Warnberg F, Blomqvist C, Ferno M, Fjallskog ML: **Cyclin B1 is a prognostic proliferation marker with a high reproducibility in a population-based lymph node negative breast cancer cohort.** *Int J Cancer* 2010, **127**(4):961-967.
130. Chae SW, Sohn JH, Kim DH, Choi YJ, Park YL, Kim K, Cho YH, Pyo JS, Kim JH: **Overexpressions of Cyclin B1, cdc2, p16 and p53 in human breast cancer: the clinicopathologic correlations and prognostic implications.** *Yonsei medical journal* 2011, **52**(3):445-453.
131. Patel H, Abduljabbar R, Lai CF, Periyasamy M, Harrod A, Gemma C, Steel J, Patel N, Busonero C, Jerjees D, Remenyi J, Smith S, Gomm JJ, Magnani L, Gyorffy B, Jones JL, Fuller-Pace FV, Shousha S, Buluwela L, Rakha EA, Ellis IO, Coombes RC, Ali S: **CDK7, cyclin H and MAT1 is elevated in breast cancer and is prognostic in estrogen receptor- positive breast cancer.** *Clin Cancer Res* 2016.
132. Skaland I, Janssen EA, Gudlaugsson E, Hui Ru Guo L, Baak JP: **The prognostic value of the proliferation marker phosphohistone H3 (PPH3) in luminal, basal-like and triple negative phenotype invasive lymph node-negative breast cancer.** *Cell Oncol* 2009, **31**(4):261-271.
133. Skaland I, Janssen EA, Gudlaugsson E, Klos J, Kjellevoid KH, Soiland H, Baak JP: **Phosphohistone H3 expression has much stronger prognostic value than classical prognosticators in invasive lymph node-negative breast cancer patients less than 55 years of age.** *Mod Pathol* 2007, **20**(12):1307-1315.
134. Skaland I, Janssen EA, Gudlaugsson E, Klos J, Kjellevoid KH, Soiland H, Baak JP: **Validating the prognostic value of proliferation measured by Phosphohistone H3 (PPH3) in invasive lymph node-negative breast cancer patients less than 71 years of age.** *Breast Cancer Res Treat* 2009, **114**(1):39-45.
135. Gudlaugsson E, Klos J, Skaland I, Janssen EA, Smaaland R, Feng W, Shao Z, Malpica A, Baak JP: **Prognostic comparison of the proliferation markers (mitotic activity index, phosphohistone H3, Ki67), steroid receptors, HER2, high molecular weight cytokeratins and classical prognostic factors in T(1)-()2)N(0)M(0) breast cancer.** *Polish journal of pathology : official journal of the Polish Society of Pathologists* 2013, **64**(1):1-8.

136. Zbytek B, Cohen C, Wang J, Page A, Williams DJ, Adams AL: **Nottingham-defined mitotic score: comparison with visual and image cytometric phosphohistone H3 labeling indices and correlation with Oncotype DX recurrence score.** *Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry* 2013, **21**(1):48-53.
137. Edge SB, Compton CC: **The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM.** *Ann Surg Oncol* 2010, **17**(6):1471-1474.
138. Tabar L, Fagerberg G, Day NE, Duffy SW, Kitchin RM: **Breast cancer treatment and natural history: new insights from results of screening.** *Lancet* 1992, **339**(8790):412-414.
139. Singletary SE, Allred C, Ashley P, Bassett LW, Berry D, Bland KI, Borgen PI, Clark G, Edge SB, Hayes DF, Hughes LL, Hutter RV, Morrow M, Page DL, Recht A, Theriault RL, Thor A, Weaver DL, Wieand HS, Greene FL: **Revision of the American Joint Committee on Cancer staging system for breast cancer.** *J Clin Oncol* 2002, **20**(17):3628-3636.
140. Fredholm H, Eaker S, Frisell J, Holmberg L, Fredriksson I, Lindman H: **Breast cancer in young women: poor survival despite intensive treatment.** *PLoS One* 2009, **4**(11):e7695.
141. Anders CK, Johnson R, Litton J, Phillips M, Bleyer A: **Breast cancer before age 40 years.** *Semin Oncol* 2009, **36**(3):237-249.
142. Fredholm H, Magnusson K, Lindstrom LS, Garmo H, Falt SE, Lindman H, Bergh J, Holmberg L, Ponten F, Frisell J, Fredriksson I: **Long-term outcome in young women with breast cancer: a population-based study.** *Breast Cancer Res Treat* 2016.
143. Kroman N, Jensen MB, Wohlfahrt J, Mouridsen HT, Andersen PK, Melbye M: **Factors influencing the effect of age on prognosis in breast cancer: population based study.** *BMJ* 2000, **320**(7233):474-478.
144. Foukakis T, Fornander T, Lekberg T, Hellborg H, Adolfsson J, Bergh J: **Age-specific trends of survival in metastatic breast cancer: 26 years longitudinal data from a population-based cancer registry in Stockholm, Sweden.** *Breast cancer research and treatment* 2011, **130**(2):553-560.
145. Bloom HJ, Richardson WW: **Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years.** *Br J Cancer* 1957, **11**(3):359-377.
146. Rakha EA, El-Sayed ME, Lee AH, Elston CW, Grainge MJ, Hodi Z, Blamey RW, Ellis IO: **Prognostic significance of Nottingham histologic grade in invasive breast carcinoma.** *J Clin Oncol* 2008, **26**(19):3153-3158.
147. Rakha EA, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, Fox SB, Ichihara S, Jacquemier J, Lakhani SR, Palacios J, Richardson AL, Schnitt SJ, Schmitt FC, Tan PH, Tse GM, Badve S, Ellis IO: **Breast cancer prognostic classification in the molecular era: the role of histological grade.** *Breast Cancer Res* 2010, **12**(4):207.

148. Frierson HF, Jr., Wolber RA, Berean KW, Franquemont DW, Gaffey MJ, Boyd JC, Wilbur DC: **Interobserver reproducibility of the Nottingham modification of the Bloom and Richardson histologic grading scheme for infiltrating ductal carcinoma.** *American journal of clinical pathology* 1995, **103**(2):195-198.
149. **Årsrapport från Nationella Bröstcancerregistret 2015**
[http://www.cancercentrum.se/globalassets/cancerdiagnoser/brost/kvalitetsregister/nationell_brostcancer_rapport_2015-2pdf.pdf]
150. Early Breast Cancer Trialists' Collaborative G, Davies C, Godwin J, Gray R, Clarke M, Cutter D, Darby S, McGale P, Pan HC, Taylor C, Wang YC, Dowsett M, Ingle J, Peto R: **Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials.** *Lancet* 2011, **378**(9793):771-784.
151. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL: **Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene.** *Science* 1987, **235**(4785):177-182.
152. Taneja P, Maglic D, Kai F, Zhu S, Kendig RD, Fry EA, Inoue K: **Classical and Novel Prognostic Markers for Breast Cancer and their Clinical Significance.** *Clinical Medicine Insights Oncology* 2010, **4**:15-34.
153. Ryden L, Haglund M, Bendahl PO, Hatschek T, Kolaric A, Kovacs A, Olsson A, Olsson H, Strand C, Ferno M, Swedish Her2 Analysis G: **Reproducibility of human epidermal growth factor receptor 2 analysis in primary breast cancer: a national survey performed at pathology departments in Sweden.** *Acta Oncol* 2009, **48**(6):860-866.
154. DeRisi J, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, Chen Y, Su YA, Trent JM: **Use of a cDNA microarray to analyse gene expression patterns in human cancer.** *Nat Genet* 1996, **14**(4):457-460.
155. Perou CM, Jeffrey SS, van de Rijn M, Rees CA, Eisen MB, Ross DT, Pergamenschikov A, Williams CF, Zhu SX, Lee JC, Lashkari D, Shalon D, Brown PO, Botstein D: **Distinctive gene expression patterns in human mammary epithelial cells and breast cancers.** *Proc Natl Acad Sci U S A* 1999, **96**(16):9212-9217.
156. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N: **A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer.** *The New England journal of medicine* 2004, **351**(27):2817-2826.
157. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, Livasy C, Carey LA, Reynolds E, Dressler L, Nobel A, Parker J, Ewend MG, Sawyer LR, Wu J, Liu Y, Nanda R, Tretiakova M, Ruiz Orrico A, Dreher D, Palazzo JP, Perreard L, Nelson E, Mone M, Hansen H, Mullins M, Quackenbush JF, Ellis MJ, Olopade OI, Bernard PS, Perou CM: **The molecular portraits of breast tumors are conserved across microarray platforms.** *BMC genomics* 2006, **7**:96.
158. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, Nordgren H, Farmer P, Praz V, Haibe-Kains B, Desmedt C, Larsimont D, Cardoso F, Peterse H, Nuyten D, Buyse M, Van de Vijver MJ, Bergh J, Piccart M, Delorenzi M: **Gene expression profiling**

- in breast cancer: understanding the molecular basis of histologic grade to improve prognosis.** *J Natl Cancer Inst* 2006, **98**(4):262-272.
159. Gruvberger S, Ringner M, Chen Y, Panavally S, Saal LH, Borg A, Ferno M, Peterson C, Meltzer PS: **Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns.** *Cancer Res* 2001, **61**(16):5979-5984.
 160. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lonning PE, Brown PO, Borresen-Dale AL, Botstein D: **Repeated observation of breast tumor subtypes in independent gene expression data sets.** *Proc Natl Acad Sci U S A* 2003, **100**(14):8418-8423.
 161. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J, Jatkoe T, Berns EM, Atkins D, Foekens JA: **Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer.** *Lancet* 2005, **365**(9460):671-679.
 162. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, Palazzo J, Marron JS, Nobel AB, Mardis E, Nielsen TO, Ellis MJ, Perou CM, Bernard PS: **Supervised risk predictor of breast cancer based on intrinsic subtypes.** *J Clin Oncol* 2009, **27**(8):1160-1167.
 163. Dubsy P, Brase JC, Jakesz R, Rudas M, Singer CF, Greil R, Dietze O, Luisser I, Klug E, Sedivy R, Bachner M, Mayr D, Schmidt M, Gehrman MC, Petry C, Weber KE, Fisch K, Kronenwett R, Gnant M, Filipits M, Austrian B, Colorectal Cancer Study G: **The EndoPredict score provides prognostic information on late distant metastases in ER+/HER2- breast cancer patients.** *Br J Cancer* 2013, **109**(12):2959-2964.
 164. Zhang Y, Schnabel CA, Schroeder BE, Jerevall PL, Jankowitz RC, Fornander T, Stal O, Brufsky AM, Sgroi D, Erlander MG: **Breast cancer index identifies early-stage estrogen receptor-positive breast cancer patients at risk for early- and late-distant recurrence.** *Clin Cancer Res* 2013, **19**(15):4196-4205.
 165. Desmedt C, Haibe-Kains B, Wirapati P, Buyse M, Larsimont D, Bontempi G, Delorenzi M, Piccart M, Sotiriou C: **Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes.** *Clin Cancer Res* 2008, **14**(16):5158-5165.
 166. Metzger Filho O, Ignatiadis M, Sotiriou C: **Genomic Grade Index: An important tool for assessing breast cancer tumor grade and prognosis.** *Critical reviews in oncology/hematology* 2011, **77**(1):20-29.
 167. Bogaerts J, Cardoso F, Buyse M, Braga S, Loi S, Harrison JA, Bines J, Mook S, Decker N, Ravdin P, Therasse P, Rutgers E, van 't Veer LJ, Piccart M, consortium T: **Gene signature evaluation as a prognostic tool: challenges in the design of the MINDACT trial.** *Nat Clin Pract Oncol* 2006, **3**(10):540-551.
 168. Wong WB, Ramsey SD, Barlow WE, Garrison LP, Jr., Veenstra DL: **The value of comparative effectiveness research: projected return on investment of the RxPONDER trial (SWOG S1007).** *Contemporary clinical trials* 2012, **33**(6):1117-1123.

169. Fried G, Moskovitz M: **Treatment decisions in estrogen receptor-positive early breast cancer patients with intermediate oncotype DX recurrence score results.** *SpringerPlus* 2014, **3**:71.
170. Aspegren K, Holmberg L, Adami HO: **Standardization of the surgical technique in breast-conserving treatment of mammary cancer.** *Br J Surg* 1988, **75**(8):807-810.
171. Fajdic J, Djurovic D, Gotovac N, Hrgovic Z: **Criteria and procedures for breast conserving surgery.** *Acta informatica medica : AIM : journal of the Society for Medical Informatics of Bosnia & Herzegovina : casopis Drustva za medicinsku informatiku BiH* 2013, **21**(1):16-19.
172. Leung JW: **Screening mammography reduces morbidity of breast cancer treatment.** *AJR American journal of roentgenology* 2005, **184**(5):1508-1509.
173. Veronesi U, Paganelli G, Galimberti V, Viale G, Zurrida S, Bedoni M, Costa A, de Cicco C, Geraghty JG, Luini A, Sacchini V, Veronesi P: **Sentinel-node biopsy to avoid axillary dissection in breast cancer with clinically negative lymph-nodes.** *Lancet* 1997, **349**(9069):1864-1867.
174. Tanis PJ, Nieweg OE, Valdes Olmos RA, Th Rutgers EJ, Kroon BB: **History of sentinel node and validation of the technique.** *Breast Cancer Res* 2001, **3**(2):109-112.
175. Kuehn T, Vogl FD, Helms G, Pueckler SV, Schirrmeister H, Strueber R, Koretz K, Kreienberg R, German multi-institutional t: **Sentinel-node biopsy for axillary staging in breast cancer: results from a large prospective German multi-institutional trial.** *Eur J Surg Oncol* 2004, **30**(3):252-259.
176. Giuliano AE, Hunt KK, Ballman KV, Beitsch PD, Whitworth PW, Blumencranz PW, Leitch AM, Saha S, McCall LM, Morrow M: **Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: a randomized clinical trial.** *Jama* 2011, **305**(6):569-575.
177. Galimberti V, Botteri E, Chifu C, Gentilini O, Luini A, Intra M, Baratella P, Sargenti M, Zurrida S, Veronesi P, Rotmensz N, Viale G, Sonzogni A, Colleoni M, Veronesi U: **Can we avoid axillary dissection in the micrometastatic sentinel node in breast cancer?** *Breast Cancer Res Treat* 2012, **131**(3):819-825.
178. Veronesi U, Cascinelli N, Mariani L, Greco M, Saccozzi R, Luini A, Aguilar M, Marubini E: **Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer.** *N Engl J Med* 2002, **347**(16):1227-1232.
179. Fisher B, Anderson S, Bryant J, Margolese RG, Deutsch M, Fisher ER, Jeong JH, Wolmark N: **Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer.** *N Engl J Med* 2002, **347**(16):1233-1241.
180. Clarke M, Collins R, Darby S, Davies C, Elphinstone P, Evans V, Godwin J, Gray R, Hicks C, James S, MacKinnon E, McGale P, McHugh T, Peto R, Taylor C, Wang Y, Early Breast Cancer Trialists' Collaborative G: **Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence**

- and 15-year survival: an overview of the randomised trials.** *Lancet* 2005, **366**(9503):2087-2106.
181. Simone NL, Dan T, Shih J, Smith SL, Sciuto L, Lita E, Lippman ME, Glatstein E, Swain SM, Danforth DN, Camphausen K: **Twenty-five year results of the national cancer institute randomized breast conservation trial.** *Breast Cancer Res Treat* 2012, **132**(1):197-203.
 182. Shou J, Massarweh S, Osborne CK, Wakeling AE, Ali S, Weiss H, Schiff R: **Mechanisms of tamoxifen resistance: increased estrogen receptor-HER2/neu cross-talk in ER/HER2-positive breast cancer.** *J Natl Cancer Inst* 2004, **96**(12):926-935.
 183. Shang Y, Brown M: **Molecular determinants for the tissue specificity of SERMs.** *Science* 2002, **295**(5564):2465-2468.
 184. Early Breast Cancer Trialists' Collaborative G, Dowsett M, Forbes JF, Bradley R, Ingle J, Aihara T, Bliss J, Boccardo F, Coates A, Coombes RC, Cuzick J, Dubsy P, Gnani M, Kaufmann M, Kilburn L, Perrone F, Rea D, Thurlimann B, van de Velde C, Pan H, Peto R, Davies C, Gray R: **Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials.** *Lancet* 2015, **386**(10001):1341-1352.
 185. Petrelli F, Coiu A, Cabiddu M, Ghilardi M, Lonati V, Barni S: **Five or more years of adjuvant endocrine therapy in breast cancer: a meta-analysis of published randomised trials.** *Breast Cancer Res Treat* 2013, **140**(2):233-240.
 186. Davies C, Pan H, Godwin J, Gray R, Arriagada R, Raina V, Abraham M, Medeiros Alencar VH, Badran A, Bonfill X, Bradbury J, Clarke M, Collins R, Davis SR, Delmestri A, Forbes JF, Haddad P, Hou MF, Inbar M, Khaled H, Kielanowska J, Kwan WH, Mathew BS, Mittra I, Muller B, Nicolucci A, Peralta O, Pernas F, Petruzella L, Pienkowski T, Radhika R, Rajan B, Rubach MT, Tort S, Urrutia G, Valentini M, Wang Y, Peto R, Adjuvant Tamoxifen: Longer Against Shorter Collaborative G: **Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial.** *Lancet* 2013, **381**(9869):805-816.
 187. Berruti A, Generali D, Kaufmann M, Puztai L, Curigliano G, Aglietta M, Gianni L, Miller WR, Untch M, Sotiriou C, Daidone M, Conte P, Kennedy D, Damia G, Petronini P, Di Cosimo S, Bruzzi P, Dowsett M, Desmedt C, Mansel RE, Olivetti L, Tondini C, Sapino A, Fenaroli P, Tortora G, Thorne H, Bertolini F, Ferrozzi F, Danova M, Tagliabue E, de Azambuja E, Makris A, Tampellini M, Dontu G, Van't Veer L, Harris AL, Fox SB, Dogliotti L, Bottini A: **International expert consensus on primary systemic therapy in the management of early breast cancer: highlights of the Fourth Symposium on Primary Systemic Therapy in the Management of Operable Breast Cancer, Cremona, Italy (2010).** *J Natl Cancer Inst Monogr* 2011, **2011**(43):147-151.
 188. Davies C, Godwin J, Gray R, Clarke M, Cutter D, Darby S, McGale P, Pan HC, Taylor C, Wang YC, Dowsett M, Ingle J, Peto R: **Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials.** *Lancet* 2011, **378**(9793):771-784.

189. Ryden L, Heibert Arnlind M, Vitols S, Hoistad M, Ahlgren J: **Aromatase inhibitors alone or sequentially combined with tamoxifen in postmenopausal early breast cancer compared with tamoxifen or placebo - Meta-analyses on efficacy and adverse events based on randomized clinical trials.** *Breast* 2016, **26**:106-114.
190. Goss PE, Ingle JN, Pritchard KI, Robert NJ, Muss H, Gralow J, Gelmon K, Whelan T, Strasser-Weippl K, Rubin S, Sturtz K, Wolff AC, Winer E, Hudis C, Stopeck A, Beck JT, Kaur JS, Whelan K, Tu D, Parulekar WR: **Extending Aromatase-Inhibitor Adjuvant Therapy to 10 Years.** *N Engl J Med* 2016, **375**(3):209-219.
191. Early Breast Cancer Trialists' Collaborative G, Peto R, Davies C, Godwin J, Gray R, Pan HC, Clarke M, Cutter D, Darby S, McGale P, Taylor C, Wang YC, Bergh J, Di Leo A, Albain K, Swain S, Piccart M, Pritchard K: **Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials.** *Lancet* 2012, **379**(9814):432-444.
192. Swain SM, Baselga J, Kim SB, Ro J, Semiglazov V, Campone M, Ciruelos E, Ferrero JM, Schneeweiss A, Heeson S, Clark E, Ross G, Benyunes MC, Cortes J, Group CS: **Pertuzumab, trastuzumab, and docetaxel in HER2-positive metastatic breast cancer.** *N Engl J Med* 2015, **372**(8):724-734.
193. Howell A, Cuzick J, Baum M, Buzdar A, Dowsett M, Forbes JF, Hochtin-Boes G, Houghton J, Locker GY, Tobias JS, Group AT: **Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer.** *Lancet* 2005, **365**(9453):60-62.
194. Holli K, Pitkanen M, Jarvenpaa R, Rajala J, Lahtela S, Hyodynmaa S, Ojala A: **Early skin and lung reactions in breast cancer patients after radiotherapy: prospective study.** *Radiother Oncol* 2002, **64**(2):163-169.
195. Killander F, Anderson H, Kjellen E, Malmstrom P: **Increased cardio and cerebrovascular mortality in breast cancer patients treated with postmastectomy radiotherapy--25 year follow-up of a randomised trial from the South Sweden Breast Cancer Group.** *Eur J Cancer* 2014, **50**(13):2201-2210.
196. Shapiro CL, Recht A: **Side effects of adjuvant treatment of breast cancer.** *N Engl J Med* 2001, **344**(26):1997-2008.
197. Chew HK: **Adjuvant therapy for breast cancer: who should get what?** *The Western journal of medicine* 2001, **174**(4):284-287.
198. Berruti A, Generali D, Kaufmann M, Puztai L, Curigliano G, Aglietta M, Gianni L, Miller WR, Untch M, Sotiriou C, Daidone M, Conte P, Kennedy D, Damia G, Petronini P, Di Cosimo S, Bruzzi P, Dowsett M, Desmedt C, Mansel RE, Olivetti L, Tondini C, Sapino A, Fenaroli P, Tortora G, Thorne H, Bertolini F, Ferrozzi F, Danova M, Tagliabue E, de Azambuja E, Makris A, Tampellini M, Dontu G, Van't Veer L, Harris AL, Fox SB, Dogliotti L, Bottini A: **International expert consensus on primary systemic therapy in the management of early breast cancer: highlights of the Fourth Symposium on Primary Systemic Therapy in the Management of Operable Breast Cancer, Cremona, Italy (2010).** *Journal of the National Cancer Institute Monographs* 2011, **2011**(43):147-151.
199. Rosell J, Nordenskjold B, Bengtsson NO, Fornander T, Hatschek T, Lindman H, Malmstrom PO, Wallgren A, Stal O, Carstensen J: **Time dependent effects of**

- adjuvant tamoxifen therapy on cerebrovascular disease: results from a randomised trial.** *Br J Cancer* 2011, **104**(6):899-902.
200. Dowsett M, Goldhirsch A, Hayes DF, Senn HJ, Wood W, Viale G: **International Web-based consultation on priorities for translational breast cancer research.** *Breast Cancer Res* 2007, **9**(6):R81.
 201. Ivshina AV, George J, Senko O, Mow B, Putti TC, Smeds J, Lindahl T, Pawitan Y, Hall P, Nordgren H, Wong JE, Liu ET, Bergh J, Kuznetsov VA, Miller LD: **Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer.** *Cancer Res* 2006, **66**(21):10292-10301.
 202. Malmstrom P, Bendahl PO, Boiesen P, Brunner N, Idvall I, Ferno M: **S-phase fraction and urokinase plasminogen activator are better markers for distant recurrences than Nottingham Prognostic Index and histologic grade in a prospective study of premenopausal lymph node-negative breast cancer.** *J Clin Oncol* 2001, **19**(7):2010-2019.
 203. Ryden L, Jonsson PE, Chebil G, Dufmats M, Ferno M, Jirstrom K, Kallstrom AC, Landberg G, Stal O, Thorstenson S, Nordenskjold B: **Two years of adjuvant tamoxifen in premenopausal patients with breast cancer: a randomised, controlled trial with long-term follow-up.** *Eur J Cancer* 2005, **41**(2):256-264.
 204. Ekholm M, Bendahl PO, Ferno M, Nordenskjold B, Stal O, Ryden L: **Two Years of Adjuvant Tamoxifen Provides a Survival Benefit Compared With No Systemic Treatment in Premenopausal Patients With Primary Breast Cancer: Long-Term Follow-Up (> 25 years) of the Phase III SBII:2pre Trial.** *J Clin Oncol* 2016, **34**(19):2232-2238.
 205. Swedish Breast Cancer Cooperative Group: **Randomized trial of two versus five years of adjuvant tamoxifen for postmenopausal early stage breast cancer. Swedish Breast Cancer Cooperative Group.** *J Natl Cancer Inst* 1996, **88**(21):1543-1549.
 206. Borgquist S, Holm C, Stendahl M, Anagnostaki L, Landberg G, Jirstrom K: **Oestrogen receptors alpha and beta show different associations to clinicopathological parameters and their co-expression might predict a better response to endocrine treatment in breast cancer.** *J Clin Pathol* 2008, **61**(2):197-203.
 207. Falck AK, Bendahl PO, Chebil G, Olsson H, Ferno M, Ryden L: **Biomarker expression and St Gallen molecular subtype classification in primary tumours, synchronous lymph node metastases and asynchronous relapses in primary breast cancer patients with 10 years' follow-up.** *Breast Cancer Res Treat* 2013, **140**(1):93-104.
 208. Hansen S, Grabau DA, Sorensen FB, Bak M, Vach W, Rose C: **The prognostic value of angiogenesis by Chalkley counting in a confirmatory study design on 836 breast cancer patients.** *Clin Cancer Res* 2000, **6**(1):139-146.
 209. Ryden S, Ferno M, Borg A, Hafstrom L, Moller T, Norgren A: **Prognostic significance of estrogen and progesterone receptors in stage II breast cancer.** *J Surg Oncol* 1988, **37**(4):221-226.

210. Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G, Kallioniemi OP: **Tissue microarrays for high-throughput molecular profiling of tumor specimens.** *Nat Med* 1998, **4**(7):844-847.
211. Falck AK, Ferno M, Bendahl PO, Ryden L: **Does analysis of biomarkers in tumor cells in lymph node metastases give additional prognostic information in primary breast cancer?** *World J Surg* 2010, **34**(7):1434-1441.
212. Baak JP, van Diest PJ, Ariens AT, van Beek MW, Bellot SM, Fijnheer J, van Gorp LH, Kwee WS, Los J, Peterse HC, et al.: **The Multicenter Morphometric Mammary Carcinoma Project (MMMCP). A nationwide prospective study on reproducibility and prognostic power of routine quantitative assessments in The Netherlands.** *Pathology, research and practice* 1989, **185**(5):664-670.
213. Chebil G, Bendahl PO, Ferno M: **Estrogen and progesterone receptor assay in paraffin-embedded breast cancer--reproducibility of assessment.** *Acta Oncol* 2003, **42**(1):43-47.
214. Bring J. TA: **Introduktion till medicinsk statistik**, vol. 1:4: Studentlitteratur 2006; 2006.
215. Bowers: **Medical statistics from scratch: an introduction for health professionals**: John Wiley & Sons Ltd; 2008.
216. Dupont: **Statistical modeling for biomedical researchers. A simple introduction to the analysis of complex data**: Cambridge University Press; 2009.
217. Vach W: **Regression models as a tool in medical research**: CRC Press, Taylor & Francis Group, LLC; 2013.
218. Schoenfeld DA: **Sample-size formula for the proportional-hazards regression model.** *Biometrics* 1983, **39**(2):499-503.
219. Royston P, Altman DG: **External validation of a Cox prognostic model: principles and methods.** *BMC medical research methodology* 2013, **13**:33.
220. Harrell FE, Jr., Lee KL, Mark DB: **Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors.** *Stat Med* 1996, **15**(4):361-387.
221. Royston P, Altman DG, Sauerbrei W: **Dichotomizing continuous predictors in multiple regression: a bad idea.** *Stat Med* 2006, **25**(1):127-141.
222. Royston PaS, W: **Multivariable Model-Building: A pragmatic approach to regression analysis based on fractional polynomials for modelling continuous variables**: John Wiley & Son,s Ltd; 2008.
223. Harrell FE, Jr.: **Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis**: Springer; 2001.
224. Sotiriou C, Pusztai L: **Advances in translational research in breast cancer; a bridge to future therapies.** *Eur J Cancer* 2008, **44**(18):2705-2706.
225. Nielsen TO, Jensen MB, Burugu S, Gao D, Tykjaer Jorgensen CL, Balslev E, Ejlersen B: **High risk premenopausal Luminal A breast cancer patients derive no benefit from adjuvant cyclophosphamide-based chemotherapy: results from the DBCG77B clinical trial.** *Clin Cancer Res* 2016.

226. Viale G, Giobbie-Hurder A, Regan MM, Coates AS, Mastropasqua MG, Dell'Orto P, Maiorano E, MacGrogan G, Braye SG, Ohlschlegel C, Neven P, Orosz Z, Olszewski WP, Knox F, Thurlimann B, Price KN, Castiglione-Gertsch M, Gelber RD, Gusterson BA, Goldhirsch A, Breast International Group T: **Prognostic and predictive value of centrally reviewed Ki-67 labeling index in postmenopausal women with endocrine-responsive breast cancer: results from Breast International Group Trial 1-98 comparing adjuvant tamoxifen with letrozole.** *J Clin Oncol* 2008, **26**(34):5569-5575.
227. Ruiz C, Seibt S, Al Kuraya K, Siraj AK, Mirlacher M, Schraml P, Maurer R, Spichtin H, Torhorst J, Popovska S, Simon R, Sauter G: **Tissue microarrays for comparing molecular features with proliferation activity in breast cancer.** *Int J Cancer* 2006, **118**(9):2190-2194.
228. Kyndi M, Sorensen FB, Knudsen H, Overgaard M, Nielsen HM, Andersen J, Overgaard J: **Tissue microarrays compared with whole sections and biochemical analyses. A subgroup analysis of DBCG 82 b&c.** *Acta Oncol* 2008, **47**(4):591-599.
229. Batistatou A, Televantou D, Bobos M, Eleftheraki AG, Kouvaras E, Chrisafi S, Koukoulis GK, Malamou-Mitsi V, Fountzilas G: **Evaluation of current prognostic and predictive markers in breast cancer: a validation study of tissue microarrays.** *Anticancer Res* 2013, **33**(5):2139-2145.
230. Knutsvik G, Stefansson IM, Aziz S, Arnes J, Eide J, Collett K, Akslen LA: **Evaluation of Ki67 expression across distinct categories of breast cancer specimens: a population-based study of matched surgical specimens, core needle biopsies and tissue microarrays.** *PLoS One* 2014, **9**(11):e112121.
231. Grasic-Kuhar C, Bracko M, Zakotnik B: **Risk factors for late relapse and death in patients with early breast cancer.** *Neoplasma* 2008, **55**(5):416-420.
232. Ignatiadis M, Azim HA, Jr., Desmedt C, Veys I, Larsimont D, Salgado R, Lyng MB, Viale G, Leyland-Jones B, Giobbie-Hurder A, Kammler R, Dell'Orto P, Rothe F, Laios I, Ditzel HJ, Regan MM, Piccart M, Michiels S, Sotiriou C: **The Genomic Grade Assay Compared With Ki67 to Determine Risk of Distant Breast Cancer Recurrence.** *JAMA oncology* 2016, **2**(2):217-224.
233. Cserni G, Voros A, Liepniece-Karele I, Bianchi S, Vezzosi V, Grabau D, Sapino A, Castellano I, Regitnig P, Foschini MP, Zolota V, Varga Z, Figueiredo P, Decker T, Focke C, Kulka J, Kaya H, Reiner-Concin A, Amendoeira I, Callagy G, Caffrey E, Wesseling J, Wells C: **Distribution pattern of the Ki67 labelling index in breast cancer and its implications for choosing cut-off values.** *Breast* 2014, **23**(3):259-263.
234. Aleskandarany MA, Green AR, Ashankyty I, Elmouna A, Diez-Rodriguez M, Nolan CC, Ellis IO, Rakha EA: **Impact of intratumoural heterogeneity on the assessment of Ki67 expression in breast cancer.** *Breast Cancer Res Treat* 2016, **158**(2):287-295.
235. Leung S NT, Zabaglo L, Arun I, Badve S, Bane A, Bartlett J, Borgquist S, Chang M, Dodson A, Enos R, Fineberg S, Focke C, Gao D, Gown A, Grabau D, Gutierrez C, Hugh J, Kos Z, Länkhholm A-V, Lin M-G, Mastropasqua M, Moriya T, Nofech-Mozes S, Osborne K, Penault-Llorca F, Piper T, Sakatani T, Salgado R, Starczynski J, Viale G, Hayes D, McShane L, Dowsett M, on behalf of the International Ki67 in

- Breast Cancer Working Group of the Breast International Group and North American Breast Cancer Group (BIG-NABCG): **Analytical validation of a standardized scoring protocol for Ki67: phase 3 of an international multicenter collaboration.** *Nature partner journals Breast cancer* 2016, **2**.
236. Focke CM, van Diest PJ, Decker T: **St Gallen 2015 subtyping of luminal breast cancers: impact of different Ki67-based proliferation assessment methods.** *Breast Cancer Res Treat* 2016, **159**(2):257-263.
 237. Denkert C, Budczies J, von Minckwitz G, Wienert S, Loibl S, Klauschen F: **Strategies for developing Ki67 as a useful biomarker in breast cancer.** *Breast* 2015, **24 Suppl 2**:S67-72.
 238. Polley MY, Leung SC, Gao D, Mastropasqua MG, Zabaglo LA, Bartlett JM, McShane LM, Enos RA, Badve SS, Bane AL, Borgquist S, Fineberg S, Lin MG, Gown AM, Grabau D, Gutierrez C, Hugh JC, Moriya T, Ohi Y, Osborne CK, Penault-Llorca FM, Piper T, Porter PL, Sakatani T, Salgado R, Starczynski J, Laenholm AV, Viale G, Dowsett M, Hayes DF, Nielsen TO: **An international study to increase concordance in Ki67 scoring.** *Mod Pathol* 2015, **28**(6):778-786.
 239. Kim KI, Lee KH, Kim TR, Chun YS, Lee TH, Park HK: **Ki-67 as a predictor of response to neoadjuvant chemotherapy in breast cancer patients.** *Journal of breast cancer* 2014, **17**(1):40-46.
 240. Ekholm M, Grabau D, Bendahl PO, Bergh J, Elmberger G, Olsson H, Russo L, Viale G, Ferno M: **Highly reproducible results of breast cancer biomarkers when analysed in accordance with national guidelines - a Swedish survey with central re-assessment.** *Acta Oncol* 2015, **54**(7):1040-1048.
 241. Zhao X, Rodland EA, Sorlie T, Naume B, Langerod A, Frigessi A, Kristensen VN, Borresen-Dale AL, Lingjaerde OC: **Combining gene signatures improves prediction of breast cancer survival.** *PLoS One* 2011, **6**(3):e17845.
 242. Prat A, Parker JS, Fan C, Cheang MC, Miller LD, Bergh J, Chia SK, Bernard PS, Nielsen TO, Ellis MJ, Carey LA, Perou CM: **Concordance among gene expression-based predictors for ER-positive breast cancer treated with adjuvant tamoxifen.** *Ann Oncol* 2012, **23**(11):2866-2873.
 243. Webster LR, Lee SF, Ringland C, Morey AL, Hanby AM, Morgan G, Byth K, Mote PA, Provan PJ, Ellis IO, Green AR, Lamoury G, Ravdin P, Clarke CL, Ward RL, Balleine RL, Hawkins NJ: **Poor-prognosis estrogen receptor-positive breast cancer identified by histopathologic subclassification.** *Clin Cancer Res* 2008, **14**(20):6625-6633.
 244. Ellis MJ, Tao Y, Luo J, A'Hern R, Evans DB, Bhatnagar AS, Chaudri Ross HA, von Kameke A, Miller WR, Smith I, Eiermann W, Dowsett M: **Outcome prediction for estrogen receptor-positive breast cancer based on postneoadjuvant endocrine therapy tumor characteristics.** *J Natl Cancer Inst* 2008, **100**(19):1380-1388.
 245. Rakha EA, Reis-Filho JS, Ellis IO: **Combinatorial biomarker expression in breast cancer.** *Breast Cancer Res Treat* 2010, **120**(2):293-308.
 246. Green AR, Powe DG, Rakha EA, Soria D, Lemetre C, Nolan CC, Barros FF, Macmillan RD, Garibaldi JM, Ball GR, Ellis IO: **Identification of key clinical**

- phenotypes of breast cancer using a reduced panel of protein biomarkers.** *Br J Cancer* 2013, **109**(7):1886-1894.
247. Aleskandarany MA, Soria D, Green AR, Nolan C, Diez-Rodriguez M, Ellis IO, Rakha EA: **Markers of progression in early-stage invasive breast cancer: a predictive immunohistochemical panel algorithm for distant recurrence risk stratification.** *Breast Cancer Res Treat* 2015, **151**(2):325-333.
248. Eden P, Ritz C, Rose C, Ferno M, Peterson C: **"Good Old" clinical markers have similar power in breast cancer prognosis as microarray gene expression profilers.** *Eur J Cancer* 2004, **40**(12):1837-1841.
249. Cuzick J, Dowsett M, Wale C, Salter J, Quinn E, Zabaglo L, Howell A, Buzdar A, Forbes J: **Prognostic value of a combined ER, PgR, Ki67, HER2 immunohistochemical (IHC4) score and comparison with the GHI recurrence score - results from TransATAC.** *Cancer Res* 2009, **69**(Suppl)(24):503s.
250. Synnestvedt M, Borgen E, Russnes HG, Kumar NT, Schlichting E, Giercksky KE, Karesen R, Nesland JM, Naume B: **Combined analysis of vascular invasion, grade, HER2 and Ki67 expression identifies early breast cancer patients with questionable benefit of systemic adjuvant therapy.** *Acta Oncol* 2013, **52**(1):91-101.
251. Rakha EA, Soria D, Green AR, Lemetre C, Powe DG, Nolan CC, Garibaldi JM, Ball G, Ellis IO: **Nottingham Prognostic Index Plus (NPI+): a modern clinical decision making tool in breast cancer.** *Br J Cancer* 2014, **110**(7):1688-1697.
252. Green AR, Soria D, Powe DG, Nolan CC, Aleskandarany M, Szasz MA, Tokes AM, Ball GR, Garibaldi JM, Rakha EA, Kulka J, Ellis IO: **Nottingham prognostic index plus (NPI+) predicts risk of distant metastases in primary breast cancer.** *Breast Cancer Res Treat* 2016, **157**(1):65-75.
253. Green AR, Soria D, Stephen J, Powe DG, Nolan CC, Kunkler I, Thomas J, Kerr GR, Jack W, Cameron D, Piper T, Ball GR, Garibaldi JM, Rakha EA, Bartlett JM, Ellis IO: **Nottingham Prognostic Index Plus: Validation of a clinical decision making tool in breast cancer in an independent series.** *The journal of pathology Clinical research* 2016, **2**(1):32-40.
254. Stalhammar G, Fuentes Martinez N, Lippert M, Tobin NP, Molholm I, Kis L, Rosin G, Rantalainen M, Pedersen L, Bergh J, Grunkin M, Hartman J: **Digital image analysis outperforms manual biomarker assessment in breast cancer.** *Mod Pathol* 2016, **29**(4):318-329.
255. Skaland I, Ovestad I, Janssen EA, Klos J, Kjellevoid KH, Helliesen T, Baak JP: **Digital image analysis improves the quality of subjective HER-2 expression scoring in breast cancer.** *Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry* 2008, **16**(2):185-190.
256. Gudlaugsson E, Skaland I, Janssen EA, Smaaland R, Shao Z, Malpica A, Voorhorst F, Baak JP: **Comparison of the effect of different techniques for measurement of Ki67 proliferation on reproducibility and prognosis prediction accuracy in breast cancer.** *Histopathology* 2012, **61**(6):1134-1144.

257. Hornberger J, Alvarado MD, Rebecca C, Gutierrez HR, Yu TM, Gradishar WJ: **Clinical validity/utility, change in practice patterns, and economic implications of risk stratifiers to predict outcomes for early-stage breast cancer: a systematic review.** *J Natl Cancer Inst* 2012, **104**(14):1068-1079.
258. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, He X, Perou CM: **Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer.** *Breast Cancer Res* 2010, **12**(5):R68.
259. Jonsson G, Staaf J, Vallon-Christersson J, Ringner M, Holm K, Hegardt C, Gunnarsson H, Fagerholm R, Strand C, Agnarsson BA, Kilpivaara O, Luts L, Heikkila P, Aittomaki K, Blomqvist C, Loman N, Malmstrom P, Olsson H, Johannsson OT, Arason A, Nevanlinna H, Barkardottir RB, Borg A: **Genomic subtypes of breast cancer identified by array-comparative genomic hybridization display distinct molecular and clinical characteristics.** *Breast Cancer Res* 2010, **12**(3):R42.
260. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, Pietenpol JA: **Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies.** *J Clin Invest* 2011, **121**(7):2750-2767.
261. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, Graf S, Ha G, Haffari G, Bashashati A, Russell R, McKinney S, Group M, Langerod A, Green A, Provenzano E, Wishart G, Pinder S, Watson P, Markowitz F, Murphy L, Ellis I, Purushotham A, Borresen-Dale AL, Brenton JD, Tavare S, Caldas C, Aparicio S: **The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups.** *Nature* 2012, **486**(7403):346-352.
262. Haibe-Kains B, Desmedt C, Loi S, Culhane AC, Bontempi G, Quackenbush J, Sotiriou C: **A three-gene model to robustly identify breast cancer molecular subtypes.** *J Natl Cancer Inst* 2012, **104**(4):311-325.
263. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, Pierga JY, Brain E, Causeret S, DeLorenzi M, Glas AM, Golfinopoulos V, Goulioti T, Knox S, Matos E, Meulemans B, Neijenhuis PA, Nitz U, Passalacqua R, Ravdin P, Rubio IT, Saghatelyan M, Smilde TJ, Sotiriou C, Stork L, Straehle C, Thomas G, Thompson AM, van der Hoeven JM, Vuylsteke P, Bernards R, Tryfonidis K, Rutgers E, Piccart M, Investigators M: **70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer.** *N Engl J Med* 2016, **375**(8):717-729.
264. Altman DG, Royston P: **The cost of dichotomising continuous variables.** *BMJ* 2006, **332**(7549):1080.
265. Sauerbrei W, Royston P, Bojar H, Schmoor C, Schumacher M: **Modelling the effects of standard prognostic factors in node-positive breast cancer. German Breast Cancer Study Group (GBSG).** *Br J Cancer* 1999, **79**(11-12):1752-1760.
266. Ejlersen B, Jensen MB, Mouridsen HT, Danish Breast Cancer Cooperative G: **Excess mortality in postmenopausal high-risk women who only receive adjuvant endocrine therapy for estrogen receptor positive breast cancer.** *Acta Oncol* 2014, **53**(2):174-185.

267. Kalderstam J, Eden P, Bendahl PO, Strand C, Ferno M, Ohlsson M: **Training artificial neural networks directly on the concordance index for censored data using genetic algorithms.** *Artificial intelligence in medicine* 2013, **58**(2):125-132.
268. **Memorial Sloan Kettering Cancer Center Breast Cancer Nomogram: Ductal Carcinoma In Situ (DCIS) Recurrence**
[<http://nomograms.mskcc.org/breast/DuctalCarcinomaInSituRecurrencePage.aspx>]
269. Martin M, Brase JC, Ruiz A, Prat A, Kronenwett R, Calvo L, Petry C, Bernard PS, Ruiz-Borrego M, Weber KE, Rodriguez CA, Alvarez IM, Segui MA, Perou CM, Casas M, Carrasco E, Caballero R, Rodriguez-Lescure A: **Prognostic ability of EndoPredict compared to research-based versions of the PAM50 risk of recurrence (ROR) scores in node-positive, estrogen receptor-positive, and HER2-negative breast cancer. A GEICAM/9906 sub-study.** *Breast Cancer Res Treat* 2016, **156**(1):81-89.
270. Saal LH, Vallon-Christersson J, Hakkinen J, Hegardt C, Grabau D, Winter C, Brueffer C, Tang MH, Reutersward C, Schulz R, Karlsson A, Ehinger A, Malina J, Manjer J, Malmberg M, Larsson C, Ryden L, Loman N, Borg A: **The Sweden Cancerome Analysis Network - Breast (SCAN-B) Initiative: a large-scale multicenter infrastructure towards implementation of breast cancer genomic analyses in the clinical routine.** *Genome medicine* 2015, **7**(1):20.