TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

LOCOREGIONAL STAGE ASSESSMENT IN CLINICALLY NODE NEGATIVE BREAST CANCER

Clinical, imaging, pathologic and statistical methods

Vilma Madekivi

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1547 | MEDICA – ODONTOLOGICA | TURKU 2021

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

LOCOREGIONAL STAGE ASSESSMENT IN CLINICALLY NODE NEGATIVE BREAST CANCER

Clinical, imaging, pathologic and statistical methods

Vilma Madekivi

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1547 | MEDICA – ODONTOLOGICA | TURKU 2021

University of Turku

Faculty of Medicine
Department of Clinical Medicine
Clinical Oncology
Doctoral programme in Clinical Research DPCR

Supervised by

MD, Ph. D, Research Professor Eeva Salminen
Finnish Nuclear and Radiation Safety
Helsinki, Finland

Adjunct professor
Department of Oncology
University of Turku and Turku University Hospital
Turku, Finland

MD, Ph.D Pia Boström
Department of Pathology
Turku University Hospital and
University of Turku
Turku, Finland

Reviewed by

Docent Päivi Auvinen
Department of Oncology
Kuopio University Hospital and
University of Eastern Finland
Kuopio, Finland

Docent Teijo Kuopio
Department of Pathology
Central Finland Health Care District and
University of Jyväskylä
Jyväskylä, Finland

Opponent

Professor Arja Jukkola
Faculty of Medicine and Health Technology
University of Tampere

Department of Oncology
Tampere University Hospital
Tampere, Finland

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-8440-4 (PRINT)
ISBN 978-951-29-8441-1 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)
Painosalama, Turku, Finland 2021

To my family

UNIVERSITY OF TURKU
Faculty of Medicine
Department of Clinical Medicine
Clinical Oncology
VILMA MADEKIVI: Locoregional stage assessment in clinically node negative breast cancer. Clinical, imaging, pathologic, and statistical methods.
Doctoral Dissertation, 111 pp.
Doctoral Programme in Clinical Research DPCR
May 2021

ABSTRACT

The locoregional staging remains an essential part of prognostication in breast cancer. Tumour size and biology, together with the number of lymph node metastases, guide the planning of appropriate treatments. Accurate clinical, imaging, pathologic, and statistical staging is needed as the surgical staging diminishes.

In this study, 743 clinically lymph node negative breast cancer patients treated in 2009–2017 were evaluated. Clinopathological factors were investigated in association with the number of lymph node metastases, the use of preoperative imaging methods and the surgical treatment method. A nomogram was developed and tested to predict the number of lymph node metastases after sentinel lymph node positivity. Three previously published models were validated to confirm their feasibility in the current population to predict nodal stage pN2a or pN3a.

Tumour size, biologic subtype and proliferation associated with higher numbers of lymph node metastases. To predict stage pN2a or pN3a, the machine learning algorithms identified tumour size, invasive ductal histology, multifocality, lymphovascular invasion, oestrogen receptor status and the number of positive sentinel lymph nodes as risk factors. The nomograms performed well with favourable discrimination. Clinopathological factors seemed to guide preoperative magnetic resonance imaging (MRI) prior to more extensive surgery. MRI estimated the increasing tumour size more accurately than mammography or ultrasound.

According to this study, clinopathological factors, additional preoperative MRI and modern statistics can be utilized in breast cancer staging without extensive surgical interference. The importance of non-surgical investigations in staging is growing in the planning of surgical, systemic and radiation treatments. Thus, maintaining the impressive survival outcomes of clinically node negative breast cancer patients can be achieved.

KEYWORDS: breast cancer, stage, axillary lymph node dissection, nomogram, magnetic resonance imaging

TURUN YLIOPISTO
Lääketieteellinen tiedekunta
Kliininen laitos
Kliininen syöpätautioppi
VILMA MADEKIVI: Kliinisesti imusolmukenegatiivisen rintasyövän paikallislevinneisyyden arvioiminen. Kliiniset, kuvantamisen, patologian alan ja tilastotieteen menetelmät.
Väitöskirja, 111 s.
Turun kliininen tohtoriohjelma TKT
Toukokuu 2021

TIIVISTELMÄ

Kasvaimen paikallinen levinneisyys on tärkeä rintasyövän ennustetekijä. Kasvaimen koko ja biologia sekä imusolmukemetastaasien lukumäärä ohjaavat syöpähoitojen suunnittelua. Levinneisyyden selvittelyssä tarvitaan huolellista kliinistä tutkimusta sekä tarkkoja kuvantamisen, patologian alan ja tilastotieteen menetelmiä, kun kirurginen levinneisyysluokittelu vähenee.

Tutkimuksessa arvioitiin vuosina 2009–2017 hoidettujen 743 kliinisesti imusolmukenegatiivisen suomalaisen potilaan tietoja. Työssä selvitettiin kliinispatologisten tekijöiden ja kainaloimusolmukemetastaasien lukumäärän, leikkausta edeltävien kuvantamistutkimusten sekä leikkausmenetelmien yhteyttä. Ennustemalli kehitettiin ja koekäytettiin positiivisen vartijaimusolmuketutkimuksen jälkeisen imusolmukemetastaasien määrän arvioimiseksi. Kolme aiemmin julkaistua mallia validoitiin, jotta niiden käyttökelpoisuus imusolmukeluokan pN2a tai pN3a ennustamisessa varmistuisi tässä aineistossa.

Kasvainkoko, biologinen alatyyppi ja jakautumisnopeus olivat yhteydessä suurempaan imusolmukemetastaasien määrään. Koneoppimisalgoritmit määrittivät levinneisyysluokan pN2a tai pN3a ennustamiseksi tarvittaviksi tekijöiksi kasvainkoon, invasiivisen duktaalisen histologian, monipesäkkeisyyden, suoni-invaasion, estrogeenireseptoristatuksen sekä positiivisten vartijaimusolmukkeiden määrän. Ennustemallit toimivat aineistossa hyvin osoittaen suotuisaa erotuskykyä. Kliinispatologiset tekijät näyttivät ohjaavan magneettikuvauspäätöstä ennen laajaa kirurgista hoitoa. Magneettikuvaus oli tarkin kuvantamismenetelmä suurenevan kasvainkoon arvioinnissa.

Tämän tutkimuksen perusteella kliinispatologiset tekijät, leikkausta edeltävä täydentävä magneettikuvaus ja nykyaikaiset tilastotieteen menetelmät voivat hyödyttää rintasyövän levinneisyysluokittelua ilman laajoja kirurgisia toimenpiteitä. Kajoamattomien tutkimusten asema levinneisyysluokittelussa on vahvistumassa kirurgisten, lääkkeellisten ja sädehoitojen suunnittelun yhteydessä. Tarkka levinneisyysluokittelu edesauttaa kliinisesti imusolmukenegatiivisten rintasyöpäpotilaiden erinomaista ennustetta.

AVAINSANAT: rintasyöpä, levinneisyys, kainaloevakuaatio, nomogrammi, magneettikuvaus

Table of Contents

Abbreviations

AJCC	The American Joint Committee on Cancer
ALN	Axillary lymph node
ALND	Axillary lymph node dissection
AUC	Area under the curve
BRCA	Breast cancer susceptibility gene
cTNM	Clinically determined tumour, nodes and metastases classification
CI	Confidence intervals
DCIS	Ductal carcinoma in situ
ECE	Extracapsular extension
EGFR	Epidermal growth factor receptor
ER	Oestrogen receptor
HE	Haematoxylin and eosin
HER2	Human epidermal growth factor receptor 2
IQR	Interquartile range
ITC	Isolated tumour cells
Ki-67	Ki-67 proliferation index
MRI	Magnetic resonance imaging
PAD	Pathological anatomic diagnosis
PR	Progesterone receptor
pTNM	Pathologically determined tumor, nodes and metastases classification
ROC	Receiver operating characteristic
SD	Standard deviation
SLN	Sentinel lymph node
SLNB	Sentinel lymph node biopsy
TNM	Tumour, nodes, and metastases classification
UICC	The Union for International Cancer Control
WHO	World Health Organization

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I Vilma Madekivi, Pia Boström, Riitta Aaltonen, Tero Vahlberg, Eeva Salminen. The sentinel node with isolated tumor cells or micrometastases. Benefits and risks of axillary dissection. *Anticancer Research*, 2017; 37: 3757–3762.

II Vilma Madekivi, Pia Boström, Antti Karlsson, Riitta Aaltonen, Eeva Salminen. Can a machine-learning model improve the prediction of nodal stage after a positive sentinel lymph node biopsy in breast cancer? *Acta Oncologica*, 2020; 59: 689–695.

III Vilma Madekivi, Antti Karlsson, Pia Boström, Eeva Salminen. Are breast cancer nomograms still valid to predict the need for axillary dissection? *Oncology,* 2021; 1–5. [In press]

IV Vilma Madekivi, Pia Boström, Tero Vahlberg, Riitta Aaltonen, Eeva Salminen. Characteristics of clinically node negative breast cancer patients needing preoperative MRI. *Surgical Oncology*, 2021; 38: 101552. [In press]

1 Introduction

Being the most common cancer in women and causing over 600 000 deaths globally in 2019, breast cancer affects the lives of millions. In Finland, breast cancer is the fifth most common cause of death for women. As nationwide screening programs have succeeded in detecting breast cancer at early stage, and oncologic treatments have evolved, breast cancer mortality has been decreasing. A growing number of breast cancer survivors will return to work and their daily lives. Therefore, breast cancer research for less and less morbid treatment schemes without compromising safety have been under great interest, leading to diminishing both breast and axillary lymph node surgeries.

Surgical removal of the malignant tissue is not only the most important curative treatment in breast cancer but also provides with valuable information on tumour extent and metastatic deposits in regional lymph nodes. Staging the local tumour and the regional lymph nodes has been the foundation for most of the advances in breast cancer treatment since the introduction of the tumour (T), node (N) and metastasis (M) classification in the 1950s. Staging has reflected the level of tumour spread and together with tumour biology guided the need for more radical, heavy or targeted treatments in order to result in the best possible prognostic outcome.

Accurate knowledge on the preoperative cancer stage and biology is requested as the extent of the primary surgery and the application of neoadjuvant systemic treatment is decided before the definite pathologic anatomical diagnosis is completed. Even after surgery, further staging investigations may be needed, if there is uncertainty of the residual tumour burden. Non-invasive staging methods such as predictive models and different imaging methods can help estimating the tumour and lymph node stage.

Predictive models, nomograms, can estimate the risk for additional lymph node metastases after positive sentinel lymph node biopsy. Yet, their validity in variable populations has not been uniform. Many institutions have developed their own predictive tools or validated previous models on a local patient population. Patient and tumour related characteristics have been shown to associate with regional lymph node stage. Mammography, ultrasound-guided needle biopsies and magnetic resonance imaging (MRI) should determine the tumour stage and location as

accurately as possible. However, the role of routine multimodality imaging that includes MRI in addition to standard mammography and ultrasound examinations is not confirmed in preoperative locoregional breast cancer staging. The association between preoperative clinical stage and postoperative pathologic stage remains indefinite, even though the clinical stage determines the primary treatment methods. The clinical TNM stage is dependent on the sensitivity and specificity of various physical and imaging examinations.

2 Review of the Literature

2.1 Detecting breast cancer

2.1.1 Background on the diagnostics in breast cancer

Breast cancer is the most common cancer and the leading cause of cancer death for women (Global Cancer Observatory 2020). In Finland, the cumulative risk for female breast cancer during lifetime is approximately 14 % (Danckert et al. 2019). Breast carcinomas are a heterogenous group of malignancies which have infiltrated through the basement membrane at their site of origin (Jones 2019). Anatomically, cancerous cells from a primary breast tumour migrate to axillary or other regional lymph nodes due to the lymphatic drainage, causing regional lymph node metastases (Tanis et al. 2001). Early breast cancer includes the cancer stages in which the cancer only involves the breast or the regional lymph nodes (National Cancer Institute [a]). In locally advanced breast cancer, clinically fixed or matted regional lymph node metastases are present, lymph node metastases in internal mammary nodes, supraclavicular or infraclavicular nodes are detected, the primary tumour exceeds 5 cm in size with regional lymph node metastases, the tumour is inflammatory or the tumour infiltrates the chest wall or the skin. Moreover, locally advanced cancers can be classified according to their eligibility for surgical treatment. (Garg & Prakash 2015).

Approximately 60 % of patients are free from regional or distant metastases at the time of primary cancer diagnosis. Up to 30–42 % of all patients are lymph node positive i.e. present with axillary lymph node (ALN) metastases at the time of diagnosis (National Cancer Institute [b], Coburn et al. 2004; Sarkeala et al. 2014). It is rare that ALN metastases occur without a detectable primary tumour, a condition known as occult breast cancer (Ofri & Moore 2020). Lymphadenopathy in the axillary or supraclavicular regions can be a sign of locally advanced breast cancer, but also of infections, injuries, sarcoidosis, other malignancies such as lymphomas, or non-specific reactivity (Bazemore et al. 2002). The probability of distant metastases at the time of diagnosis is low: globally, only 5 to 10 % of breast cancer tumours have observably spread beyond regional lymph nodes or via blood

circulation at the time of diagnosis (National Cancer Institute [b]; Cardoso et al. 2018).

In Europe, most breast cancer patients are first diagnosed in a screening program or by the clinical discovery of a palpable breast tumour or other breast symptoms (Biganzoli et al. 2020). The incidence and the mortality of breast cancer increases with age (Danckert et al. 2019). In Finland, 86.5 % of new breast cancer patients were aged 50 years old or older in the years 2014 to 2018 (Figure 1, Finnish Cancer Registry [a]). In addition to age, the risk factors for developing breast cancer include oestrogen exposure during lifetime, genetics, some benign breast conditions and dietary or lifestyle matters (Collaborative Group on Hormonal Factors in Breast Cancer 2012; Dafni et al. 2019; Dyrstad et al. 2015; McTiernan 2003).

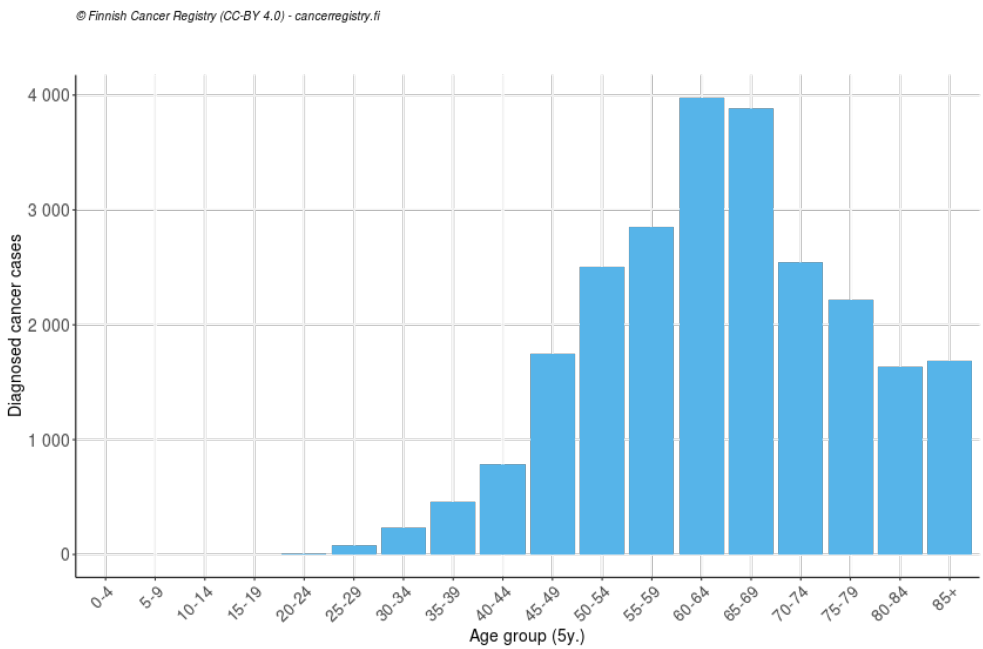

Figure 1. Breast cancer incidence among Finnish women by age in 2014–2018. (Finnish Cancer Registry [a], accessed 9.11.2020. Reproduced with permission from Finnish Cancer Registry.)

The current international guidelines state that the primary work-up for breast cancer diagnosis should consist of the physical examination of breasts and local lymph node regions, the imaging of breast and axillary regions and the pathologic assessment of needle biopsies of tumours and suspicious ALNs (Cardoso et al. 2019). The overall diagnostic and treatment responsibility is often provided by a multidisciplinary breast cancer group, considered as an important communication between breast

cancer specialists (Burstein et al. 2019). The multidisciplinary breast cancer groups may in fact improve cancer-related survival and harmonize differences between hospitals (Kesson et al. 2012). However, there is uncertain evidence that multidisciplinary breast cancer groups' work alone benefits cancer survival, since the research in multidisciplinary meetings is often retrospective in nature and the reasons for improved outcomes may be unclear (Houssami et al. 2006; Shao et al, 2019).

At Turku University Hospital, the process of breast cancer diagnosis and primary treatment starts with tumour detection by screening or clinically, accompanied by full mammography, and ultrasound-guided core needle biopsy of the tumour. The diagnosis is often made in the primary health care, after which the patient is referred to the Department of Surgery. Physical examination may be performed after imaging if the tumour is detected by screening. Before or after the appointment with the surgeon, additional staging investigations by imaging and needle biopsies may be ordered through multidisciplinary discussions. In 2020, the time from referral from primary health care to primary surgical treatment was 21–25 days. Postoperatively, the final pathological anatomic results are viewed in relation to other investigations, forming the comprehensive summary and guiding the further treatment plan in the multidisciplinary breast cancer meeting. (Tyks Cancer Centre webpage)

2.1.2 Breast cancer screening

Breast cancer screening programs are available and recommended biannually in most European Union countries for 50 to 69 years old women, with no previous cancer diagnosis or history of familial breast cancer. Breast cancer screening is recommended to be conducted as screening mammography. (Basu et al. 2018; Cardoso et al. 2019). The benefits and drawbacks of breast cancer screening have been thoroughly discussed (Myers et al. 2015). Due to cancer detection in earlier stage, screening mammography programs have been shown to improve cancer mortality (Zielonke et al. 2020), although criticism of overall benefits has been presented (Gøtzsche et al. 2013). Nevertheless, women tend to value the benefits of breast cancer screening over the possible psychological distress or overtreatment caused by the screening (Mathioudakis et al. 2019).

For breast cancer susceptibility 1 and 2 (*BRCA1* and *2*) gene mutation carriers, clinical examination every 6–12 months is recommended at minimum from the age 25 together with annual MRI screening, and annual mammography from the age 30 years (Paluch-Shimon et al. 2016).

The current Finnish screening program for breast cancer covers approximately 80 % of the 50–69 years old women that have been invited to the free bilateral mammography screening (Heinävaara et al. 2016). At screening age, screening

detects 55–62 % of new cancer cases. Of these cases, 29 % have regional lymph node metastases according to a Finnish population. (Davidson et al. 2013; Sarkeala et al. 2014) It has been estimated that screening reduces breast cancer mortality for 33 % among the Finnish participants (Heinävaara et al. 2016). However, international research has suggested a lower reduction of 15 % in mortality, and emphasized to better acknowledge the harmful effects of screening (Gøtzsche et al. 2013). All Finnish women in screening age are invited to the mammography appointment. These results are viewed by two breast radiologists, and in suspicious cases the woman is invited to further investigations: additional mammography images, ultrasound examination and needle biopsies are often performed (Finnish Cancer Registry [b]).

Screening younger than 50 years old women is not recommended due to no benefit in cancer mortality, and high rates of over-diagnosis (van den Ende et al. 2017). However, 44 % of breast cancers detected in Finland before screening age are staged with regional lymph node metastases, and 6 % with distant metastases. Most breast cancer deaths occur on patients diagnosed after the screening age, as distant metastases in this age group are more common than in younger patients at the time of diagnosis. Including the age group of 70–74 years old women in the screening program have been discussed. (Sarkeala et al. 2014) Since the 1990s, tumours detected especially at screening age are becoming smaller in size, and tumours in any age group are becoming oestrogen receptor (ER) positive more frequently (Aromaa-Häyhä et al. 2018).

2.2 Predictive and prognostic factors in breast cancer

2.2.1 Predictive factors in breast cancer

When cancer occurs, predictive and prognostic factors are assessed to evaluate the expected effects of treatments and the prognosis for cancer recurrence after treatment. Only few predictive factors are suitable clinical indicators of specific treatments in breast cancer: ER, progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) status, and in metastatic triple-negative tumours, programmed death-ligand 1 (PD-L1). (Perry et al. 2006; WHO Classification of Tumours Editorial Board 2019) However, many predictive factors have been noted for other uses in breast cancer management, such as predicting the number of ALN metastases (van la Parra et al. 2011).

2.2.2 Prognostic factors in breast cancer

Prognostic factors are patient or tumour-related characteristics which have individual and multivariable associations with patient prognosis i.e., cancer recurrence or survival. Moons et al. (2009) summarized that "prognosis is estimating the risk of future outcomes in individuals based on their clinical and non-clinical characteristics". In breast cancer, young age, tumour size, tumour multifocality, histological type, histological grade, the number and size of ALN metastases, lymphovascular invasion, tumour-infiltrating lymphocytes, excision margins, ER and PR status, HER2 overexpression and the proliferation activity of tumour cells have been established as the most important prognosticators (Anders et al. 2008; Burstein et al. 2019; Merkel & Osborne 1989; Perry et al. 2006; Pinto et al. 2001). Many prognostic models have been developed to calculate multivariable effects of individual prognostic factors (Phung et al. 2019).

Reflecting the cancer aetiology, genetic variance, and the above-mentioned tumour biology by hormone receptors, HER2 and proliferation, four biological cancer subtypes have been introduced and also employed as prognostic factors: luminal A-like, luminal B-like, HER2-enriched and triple-negative subtypes (Table 1) (Cardoso et al. 2019; WHO Classification of Tumours Editorial Board 2019; Nguyen et al. 2008). Overlapping with triple-negative tumours, basal type cancers are distinguished by cytokeratin CK5/6 or epidermal growth factor receptor (EGFR) positivity (Pal et al. 2011). Other subgroups of triple-negative cancer include immunomodulatory, mesenchymal, mesenchymal stem-like and luminal androgen receptor cancer types (Yin et al. 2020)

Table 1. Biologic breast tumour subtypes. Modified from WHO Classification of Tumours Editorial Board 2019.

Biological subtype	Hormone receptor status	HER2 status	Proliferation
Luminal A-like	ER and PR positive	Negative	Low Ki-67
Luminal B-like (HER2 negative)	ER positive PR negative or low	Negative	High Ki-67
Luminal B-like (HER2 positive)	ER positive Any PR	Positive	Any Ki-67
HER2-enriched	Negative	Positive	Any Ki-67
Triple-negative	Negative	Negative	Any Ki-67

The hormone receptor positive luminal subtypes are associated with less recurrences and increased overall survival compared to the hormone receptor negative subtypes. However, luminal B-like cancers are more aggressive than luminal A-like cancers, with higher recurrence rates. (Prat et al. 2015) Even though targeted therapies have improved the prognosis for patients with HER2 positive cancers, their outcome is still poorer than for those with HER2 negative cancers. Triple-negative tumours are very aggressive and associate with the worst prognosis among different breast cancer subtypes. (Waks et al. 2019)

Also, genotyping tumour samples have brought out differences between breast malignancies. The genetic variance conveys prognostic information (Guo et al. 2015). As the germline mutations in *BRCA1* or *BRCA2* genes remain important subjects for genetic testing, 16 other genes will also be tested in the current practice if hereditary breast cancer is suspected (Paluch-Shimon et al. 2016; Turku University Hospital instructions 2020). Women with strong familial cancer predisposition, young age at time of breast cancer diagnosis or previous ovarian cancer diagnosis should be offered genetic testing for *BRCA1* and *2* mutations (Cardoso et al. 2019).

2.2.3 TNM classification of breast tumours

Cancer staging aims to predicting patient prognosis and planning cancer treatments (Sawaki et al. 2019). The anatomical stage classification by The American Joint Committee on Cancer (AJCC) and The Union for International Cancer Control (UICC) for tumours (T), nodes (N) and metastases (M) has guided clinical decision-making for decades and is agreed with by the World Health Organization (WHO) (Table 2). In these publications, stage has been described as a summary of the locoregional and distant extent of the disease. Stage is traditionally reported from 0 to IV (Table 3), and is utilized in planning treatments, conducting research, estimating prognosis, and transmitting comparable information. Staging takes advantage of both clinically detected (referred to with prefix “c”) and pathological (referred to with prefix “p”) information. The clinical cTNM stage is determined by physical and imaging examinations, whereas the pathological pTNM stage concludes the findings microscopically and by histopathological methods. The final pathological stage is attained after surgically removing all cancerous breast tissue and examining the regional lymph nodes by sentinel lymph node biopsy (SLNB) or axillary lymph node dissection (ALND). (Giuliano et al. 2017 [a]; WHO Classification of Tumours Editorial Board 2019)

Table 2. TNM classification of breast tumours by World Health Organization. Modified from WHO Classification of Tumours Editorial Board 2019.

T: Primary tumour size		**N: Regional lymph nodes**	
TX	Primary tumour cannot be assessed	pNX	Regional lymph nodes cannot be assessed
T0	No evidence of primary tumour	pN0	No regional lymph node metastasis or ITC only
Tis	Ductal carcinoma in situ or Paget disease of the nipple	pN1	Metastasis in 1–3 ipsilateral axillary lymph nodes; and/or internal mammary nodes detected by sentinel lymph node biopsy but not clinically detected
T1	2 cm or less in greatest dimension	pN1mi	Micrometastasis only
T1mi	Microinvasion 0.1 cm or less	pN1a	Metastasis in 1–3 ipsilateral axillary lymph nodes, including at least 1 larger than micrometastasis
T1a	> 0.1 cm but ≤ 0.5 cm	pN1b	Internal mammary lymph node metastasis
T1b	> 0.5 cm but ≤ 1 cm	pN1c	Metastasis in 1–3 ipsilateral axillary lymph nodes and internal mammary lymph nodes
T1c	> 1 cm but ≤ 2 cm	pN2a	Metastasis in 4–9 ipsilateral axillary lymph nodes
T2	More than 2 cm but not more than 5 cm in greatest dimension	pN2b	Clinically detected internal mammary lymph node metastasis without axillary lymph node metastases
T3	More than 5 cm in greatest dimension	pN3a	Metastasis in more than 10 ipsilateral axillary lymph nodes or metastasis in infraclavicular level III lymph nodes
T4	Any size with direct extension to chest wall (ribs, intercostal muscles, or serratus anterior, but not pectoral muscle) and/or to skin. Invasion of the dermis alone does not qualify as T4.	pN3b	Clinically detected internal mammary lymph node metastasis with axillary lymph node metastasis; or internal mammary lymph node metastasis detected in sentinel lymph node biopsy with more than 3 axillary lymph node metastases
T4a	Extension to chest wall	pN3c	Metastasis in ipsilateral supraclavicular lymph node(s)
T4b	Ulceration, ipsilateral satellite skin nodules, or skin oedema		
T4c	Both T4a and T4b		
T4d	Inflammatory carcinoma		

M: Distant metastasis	
M0	No distant metastasis
M1	Distant metastasis

Table 3. Breast cancer stages 0 to IV World Health Organization. Modified from WHO Classification of Tumours Editorial Board 2019.

Stage 0	Tis	N0	M0
Stage IA	T1	N0	M0
Stage IB	T0, T1	N1mi	M0
Stage IIA	T0, T1	N1	M0
	T2	N0	M0
Stage IIB	T2	N1	M0
	T3	N0	M0
Stage IIIA	T0, T1, T2	N2	M0
	T3	N1, N2	M0
Stage IIIB	T4	N0, N1, N2	M0
Stage IIIC	Any T	N3	M0
Stage IV	Any T	Any N	M1

Some confusion in the classification and terminology has, however, been aroused between the AJCC and UICC manuals (Cserni et al. 2018). The latest 2017 edition of AJCC cancer staging manual has enhanced differences between these publications by introducing a new prognostic approach. Anatomical stage is now extended with biological prognosticators: ER and PR status, HER2 status and grade, together with commercial multigene assays, resulting in a "clinical prognostic stage group" prior to operation, and a "pathologic prognostic stage group" after tumour resection and distinct pathologic analyses (Cserni et al. 2018; Giuliano et al. 2017 [a]; 2018).

After the introduction of the biomarkers into the new classification, its superiority over the former anatomical classification has been confirmed (Weiss et al. 2018). It has been estimated that the new classification will upstage 19.3 % and downstage 23.8 % of patients with locoregional cancer compared to the former classifications (Plichta et al. 2018). For example, patients with the former stage IA have generally favourable prognosis, yet in some cases result in unexpected recurrence or mortality. The unfavourable outcomes in this stage group may be predicted with high Ki-67 proliferation index (Ki-67) values, low percentage of ER positive cells or HER2 positivity (Kronqvist et al. 2004), and in the new TNM classification these patients are upstaged into stage IB if the tumour presents with grade II–III and triple-negativity (NCCN Clinical Practice Guidelines 2020).

2.3 Clinical locoregional stage assessment

2.3.1 Physical examination

The preoperative investigations are necessary for determining the treatment since the treatment decisions depend on the estimated locoregional cancer stage. The extent of the surgery and neoadjuvant systemic treatments rely on preoperative physical examination, imaging, and pathological needle biopsies, i.e., the triple assessment. (Biganzoli et al. 2020).

During breast cancer diagnostics a crude locoregional stage estimation is done by a breast specialist including visual investigation and bimanual palpation of breasts and regional superficial lymph nodes: the axillary region and supraclavicular areas (Cardoso et al. 2019; Willet et al. 2010). At the time of cancer detection, it is approximated that 25 to 35 % of breast malignancies are non-palpable and the median tumour diameter is smaller than 2 cm (Cady et al. 1996; Lovrics et al. 2011). Tumour palpation can both overestimate and underestimate tumour size (Bosch et al. 2003; Katz et al. 2017). Concerning healthy women, physical examination can increase false positive findings if it is used in breast cancer screening (Myers et al. 2015). In fact, the regular self-examination of breasts does not increase survival but results in more diagnostic procedures (Kösters & Gøtzsche 2003). The palpation of regional lymph nodes has a high false negative rate of 45 % (Sacre 1986). The physical examination of locoregional mamillary and lymph node areas is complemented with an assessment of possible distant metastases (Cardoso et al. 2019).

2.3.2 Imaging in early and locally advanced breast cancer

Imaging in early and locally advanced breast cancer includes bilateral mammography and ultrasound examination for all eligible patients (Cardoso et al. 2019). Imaging should define the size and localization of the primary tumour and its every focus, and the involvement of regional lymph nodes. An accurate clinical cTNM stage depends on the imaging examinations. In mammography, ultrasound and MRI, breast cancer tumours reflect the biological cancerous processes and visualize as irregularly shaped masses and indistinct or spiculated margins (Monticciolo 2011).

2.3.2.1 Mammography

In mammography, space-occupying breast cancer masses, asymmetry and calcifications can be visualized by ionizing radiation (X-rays). The radiation doses

used in mammograms are low, and for example the regular screening mammography is estimated to result in 10 breast cancers per 100 000 women during lifetime (Hauge et al. 2014). The lesions are reported according to the breast imaging reporting and database system (BI-RADS). (Rao et al. 2016; Shah et al. 2014) Although bilateral mammography is the routine imaging method for most patients, its accuracy in estimating tumour size is inferior to ultrasound and MRI (Berg et al. 2004; Bosch et al. 2003; Hata et al. 2004; Hieken et al. 2001). Mammography is more sensitive for older than younger women, due to its better performance on fatty rather than dense breast tissue (von Euler-Chelpin et al. 2019). Among women younger than the current screening age, false negative mammography findings occur in 35 % of cases (Joensuu et al. 1994). Mammography can reveal lymphadenopathy but is often unable to distinguish between cancerous and benign causes. For example, if the size of homogeneously dense nonfatty nodes is used to distinguish abnormal lymph nodes, the sensitivity of mammography to detect ALN metastases is only 30 %. (Walsh et al. 1997).

2.3.2.2 Ultrasound examination

Ultrasound examination of the primary breast lesion is the first imaging choice for women under 30–35 years of age, pregnant or breast-feeding (Finnish Breast Cancer Group 2019; Shah et al. 2014). The regional lymph nodes are most often examined with ultrasound, leading to needle biopsies in suspicious cases. The current practice suggests that in tumour stage T1–2, preoperative needle biopsies from ALNs are performed only if several ALNs present with abnormal features (Finnish Breast Cancer Group 2019). Imaging the axilla should in minimum include lymph node levels I and II (Figure 2) (Chang et al. 2020). Suspicious lymph nodes in ultrasound examination visualize with round or irregular shape, asymmetrically thick cortex, or loss of fatty hilum (Chang et al. 2020; Cho et al. 2009). A systematic review and a meta-analysis have presented that the sensitivity of ultrasound in assessing ALN involvement is 50 %, and the false negative rate is 25 % (Diepstraten et al. 2014). Currently, suspicious lymph nodes in imaging are examined with either ultrasound-guided fine needle aspiration or core needle biopsy. Yet, ultrasound-guided core needle biopsy has been proven to be more accurate in axillary staging than fine needle aspiration (Balasubramanian et al. 2018). In fact, contrast-enhanced ultrasound and a guided core needle biopsy could be an alternative for a surgical sentinel lymph node biopsy in the preoperative staging of regional lymph nodes (Nielsen et al. 2017).

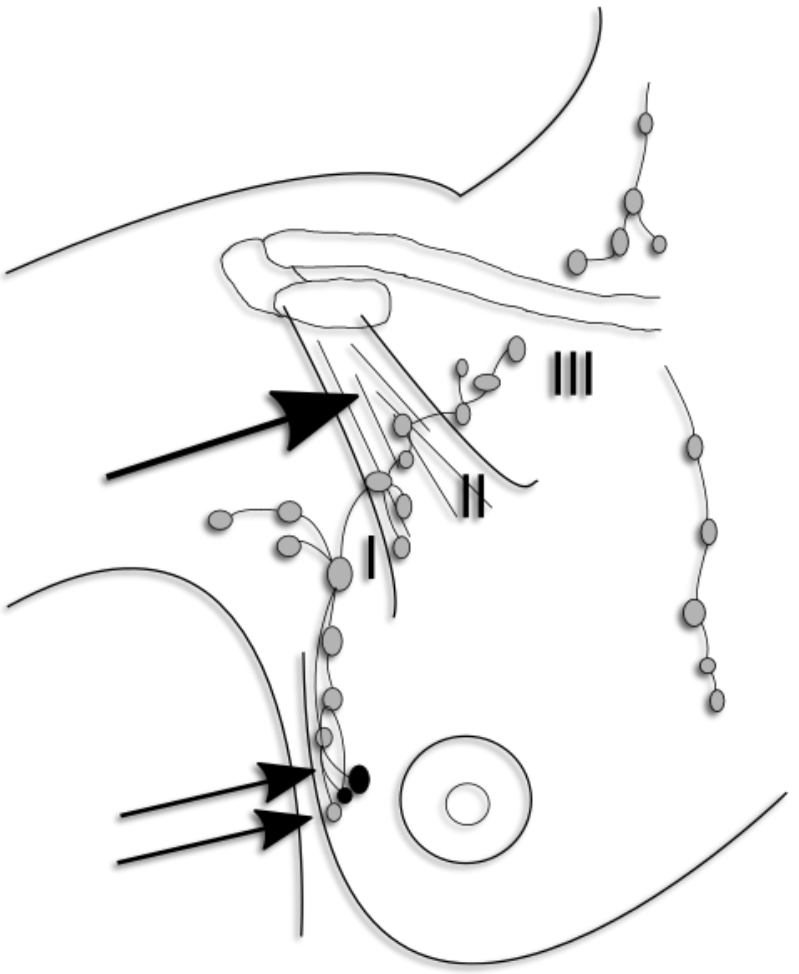

Figure 2. Regional lymph nodes. Modified from Chang et al. (2020). Axillary nodal chains at levels I, II and III, in relation to musculus pectoralis minor (one arrow). Sentinel lymph nodes are often located at the lower part of level I lymph nodes (two arrows).

2.3.2.3 Magnetic resonance imaging

MRI is the most accurate imaging method in tumour size and multifocality estimations (Sardanelli et al. 2010), and it can benefit in discriminating HER2 overexpression, biological tumour subtypes and ductal carcinoma in situ (DCIS) (Elias et al. 2014; Holli-Helenius et al. 2017; Tajima et al. 2019). Yet, routine MRI in preoperative tumour assessment is only recommended for patients with mutated *BRCA1* or *2* genes, invasive lobular histology, dense breast tissue, suspicion of multifocality, conflicting imaging results and physical examination, occult breast cancer, or neoadjuvant systemic therapies; or in other cases after a multidisciplinary

breast cancer group's decision (Bakker et al. 2019; Cardoso et al. 2019; Causer et al. 2017; Mann et al. 2008; Marinovich et al. 2015; Sardanelli et al. 2010). A second-look ultrasound examination should be used to further characterize the changes seen in MRI (Mazzei et al. 2019). Adding MRI to presurgical imaging investigations can delay surgical treatment for 11 to 20 days (Bleicher et al. 2009; Zhang et al. 2017) and increase mastectomies or prophylactic procedures (Houssami et al. 2017).

In MRI, same features are feasible for suspicious ALNs as in ultrasound examination, together with heterogenous enhancement, "perifocal oedema" and "rim enhancement", as particular MRI findings (Ecanow et al. 2013). Although MRI can detect positive lymph nodes, discriminating between nodal stages N1 and N2-3 might be limited (van Nijnatten et al. 2016).

2.3.2.4 Other imaging methods

Other imaging methods used for locoregional breast cancer include digital tomosynthesis, and infrared breast thermography (Alabousi et al. 2020; Singh & Singh 2020). Computed tomography of the chest can be suggested for patients with clinically detected ALN metastases, large tumours, or aggressive tumour biology (Cardoso et al. 2019). To improve the diagnostic imaging, computer aided detection of malignancies by different machine-learning methods has been under growing interest in breast cancer research, showing promising results in both increasing sensitivity and specificity (Yassin et al. 2018).

2.3.3 Surgical treatment and stage assessment

In earlier years, the most important treatment for breast cancer was wide excisions of mammary and axillary areas, in order to remove all cancerous tissue with substantial margins (Sakorafas et al. 2010). By removing the entire primary tumour and pathologic lymph nodes, an accurate locoregional stage could be specified. Today, breast conserving surgery can be suggested for up to 80 % of new breast cancer patients, tumour-free dissection surface is considered a sufficient margin for invasive carcinomas, and axillary management continues to renew towards minimal invasiveness (Buchholz et al. 2014; Cardoso et al. 2019, Veronesi et al. 2012).

Breast conserving surgery refers to partial resection of the breast, which includes the malignant tumour. It should be the first choice of surgical treatment for most stage I–II breast cancer patients (Association of Breast Surgery at Baso 2009; Ayala de la Peña et al. 2019). In breast conserving surgery, an essential goal is to dissect the entire primary tumour with clear margins because margin positivity associates with more local recurrences (Houssami et al. 2014). To achieve this, breast surgeons can employ intraoperative tumour palpation, ultrasound guidance, wire localization,

radioactive seed localization, radioguided occult lesion localization, radiography of the resected specimen, or intraoperative pathologic assessment of the specimen (Gray et al. 2018). "No ink on tumour" is the new standard for an adequate tumour margin for invasive carcinomas (Buchholz et al. 2014). For DCIS, a margin of 2 mm is recommended (Kuerer at al. 2017). If the tumour excision in one sample fails, precise tumour staging is impossible even after additional excisions, because the relation of multiple biopsies cannot be reliably determined (Stamatakos 2011). Breast conserving surgery is always followed by whole breast radiation therapy to meet with – and even improve – the survival rates of patients undergoing mastectomy (Clarke et al. 2005; van Maaren et al. 2016). Nowadays, oncoplastic methods are often preferred due to good cosmetic outcomes and the possibility to wider excisions, although research methodology on assessing aesthetic views has been varying (Papanikolau et al. 2019). Oncoplastic breast conserving surgery is safe, with comparable survival outcomes to other techniques (de Lorenzi et al. 2016).

By mastectomy, the entire mammary tissue is removed as precisely as possible, resulting in as accurate tumour staging as possible. For reconstruction purposes, skin and nipple can be spared (Goethals & Rose 2020). As mastectomy is conducted, some patients undergo contralateral breast surgery for volume symmetry (Nahabedian 2008). Mastectomy is recommended if the tumour is very large in relation to breast size, tumour foci extend to multiple breast quadrants, surgical margins remain positive after multiple resections, there are contraindications to radiotherapy, or if the patient wishes for radical surgery (Ayala et al. 2019; Cardoso et al. 2019). If gene mutations of *BRCA1* or *2* are present, contralateral prophylactic mastectomy can be offered (Paluch-Shimon et al. 2016).

Staging the regional lymph nodes has been done by surgically removing at least ten of the approximately 40 ALNs from levels I–II (Figure 2) (Ebner at al. 2019; Lee & Sabel 2011). This can result in postoperative arm-swelling, pain, numbness, and shoulder dysfunction (Castelo et al. 2020; Wang et al. 2020). Nowadays, to obtain fewer complications compared to axillary lymph node dissection, clinically node negative patients are referred to SLNB, which results in better postoperative quality of life (Mansel et al. 2006; Veronesi et al. 2003; Wang et al. 2011). In fact, staging the axilla should be done by SLNB instead of ALND on clinically node negative patients (Bromham et al. 2017). Sentinel lymph nodes (SLN) are defined as the first lymph nodes which are met by afferent lymph fluid and migrating tumour cells from the primary tumour, and which act as "filters" for tumour cells (Harrison & Brock 2018; Nieweg et al. 2001). They can be detected by the breast surgeon by blue dye lymphatic mapping and radio-guidance (Gipponi et al. 2004). If the SLN is free from cancerous deposits, it is unlikely that other lymph nodes will be involved, and the axillary recurrence rate is as low as 0.3 % (van der Ploeg 2008).

Sometimes the lymphatic fluid flows towards non-palpable internal mammary lymph nodes located parasternally, which then act as SLNs (Heuts et al. 2009). Internal mammary lymph node metastases are present in 5–17 % of ALN negative patients, and in 28–52 % of ALN positive patients (Gong et al. 2017). Internal mammary lymph nodes or supraclavicular lymph nodes are not routinely assessed during staging investigations or surgery (Heuts et al. 2009). These areas are typically treated with postoperative radiation therapy instead of surgery, with a curative intention, if multiple ALN metastases are found (Poortmans et al. 2015). Supraclavicular lymph node metastases are usually a presentation of late locally advanced cancer if distant metastases are not yet present, their incidence being up to 4.3 % of locoregional breast cancer. Supraclavicular lymph node metastases may be treated with supraclavicular lymph node dissection and not only with oncologic treatments, especially in non-luminal breast cancer. (Ai et al. 2020)

The last decade of clinical trials has changed the look on ALN control in early breast cancer: NSABP B-32 showed that SLNB alone is as safe as ALND when SLNB shows no metastases (Krag et al. 2010). IBCSG 23-01 showed that women with early breast cancer, tumour diameter of 5 cm or less and only isolated tumour cells (ITC) or micrometastases (0.2–2 mm in size) in SLNB do not benefit from additional ALND (Galimberti et al. 2018). AMAROS showed that clinically node negative patients with T1–2 disease and at least one SLN metastasis can be treated as safely with axillary irradiation as with ALND, according to six years of follow-ups of axillary recurrence as the primary endpoint. The long-term observations of the AMAROS trial are yet to be published. In AMAROS trial, the maximum number of positive SLNs was not limited although only 1 % of the study and control groups had four or more positive SLNs. (Donker et al. 2014)

ACOSOG Z0011 showed that clinically node negative patients with T1–2 disease treated with breast conserving surgery and whole-breast irradiation do not benefit from ALND even if SLNB shows 1–2 metastases. Unlike the AMAROS trial, the primary endpoint in the ACOSOG Z0011 trial was overall survival (Giuliano et al. 2017 [b]). Even though the results of ACOSOG Z0011 have been widely acknowledged in clinical practice, the trial has also received notable criticism due to limited methodology (Latosinsky et al. 2012). Still, the most recent recommendations have stated that the ACOSOG Z0011 results could be applied to patients undergoing mastectomy and axillary radiation therapy (Burstein et al. 2019; Cardoso et al. 2019).

Based on international clinical trials, Finnish Breast Cancer Group has published national breast cancer diagnostic and treatment recommendations in Finland for over a decade (Finnish Breast Cancer Group 2019). The updated 2019 recommendation takes a more precise stand on routine use of ALND on breast cancer patients than the international guidelines do (Cardoso et al. 2019; Lyman et al. 2016).

Accordingly, even if SLNs show macrometastases residual disease can for most cases be managed with axillary irradiation instead of ALND. Derived from the Finnish recommendation most early breast cancer patients with tumour size up to 5 cm and pN0–1 disease will not undergo ALND. ALND is recommended for those with a positive SLN and intraoperative suspicion of N2–3 disease, a bigger tumour, recurrent cancer, neoadjuvant therapies received or inability to receive radiotherapy (Finnish Breast Cancer Group 2019). The sufficiency of SLNB as the only surgical method of staging clinically node negative axilla and finding axillary nodal burden of N2–3 remains debatable for example in a situation with three positive SLNs.

In the future, the SOUND trial will give new information about omitting SLNB and other surgical staging of the axilla (Gentilini et al. 2012; 2016), and the POSNOC trial will randomise patients with T1–2 disease and 1–2 macrometastases in SLNB into no axillary surgery, ALND or axillary irradiation (Goyal et al. 2015).

2.4 Pathologic locoregional stage assessment

2.4.1 Needle biopsies

During breast cancer diagnostics and preoperative stage assessment, needle biopsies from suspected primary tumours and regional lymph nodes are evaluated by breast pathologists. Core needle biopsy is more sensitive than fine needle aspiration, retaining the architectural look of the tissue, but both have good clinical performance (Wang et al. 2017). If core biopsy is not available, fine needle aspiration may be performed as an affordable diagnostic procedure, being also less time-consuming and less risky for complications than core biopsy (Field et al. 2019).

The core needle specimen of the breast lesion is analysed and the histology, hormone receptor status, HER2 expression, grade, and Ki-67 can be determined (Cardoso et al. 2019). The agreement of core needle biopsy and larger surgical samples is good for ER, PR and HER2 status (Chen et al. 2012; Li et al. 2012) but only moderate for tumour grade (Knuttel et al. 2016). It is often advisory to repeat hormone receptor and HER2 testing on the surgical sample, if core biopsy results are negative (Ellis et al. 2016). If core needle biopsy shows DCIS, approximately 25 % of the lesions have invasive histology in the final pathological report (Brennan et al. 2011). If the lesion is very small in size, core needle biopsy may remain the only detected histological sample of the primary tumour. Biopsies from all suspected breast foci are recommended (Cardoso et al. 2019).

Suspicious ALNs are investigated with ultrasound-guided fine needle aspiration cytology or core needle biopsy. The core biopsy is more sensitive than fine needle aspiration in diagnosing lymph node metastases. The core biopsy can result in more complications, but fine needle aspiration is associated with more frequent need for

additional diagnostic procedures. (Wang et al. 2012; Balasubramanian et al. 2018; Pyo et al. 2020). Fine needle aspiration samples are direct smears, with air-dried Giemsa staining and alcohol-fixed Papanicolaou staining. They result in classifying the cytology samples into five categories: insufficient/inadequate, benign, atypical, suspicious of malignancy, or malignant. (Field et al. 2019)

2.4.2 Surgical resection specimens

As the resected breast sample or the entire mastectomy sample is evaluated in the pathology laboratory, a final and conclusive assessment of tumour stage is determined. To ensure reliability, the specimen is carefully processed. Excisional biopsy surfaces are freshly inked to preserve anatomical and three-dimensional orientation. The breast is high in adipose tissue and therefore proper fixation in adequate amount of formalin is necessary, and slicing should preferably be done within hours from surgery. Haematoxylin and eosin (HE) staining, complemented with immuhistochemical stainings, is the basis of histological analysis. (Ellis et al. 2016; Stamatakos 2011)

Tumour size is determined from the invasive component of the tumour, measured as the largest dimension both macroscopically and microscopically (Ellis et al. 2016; Stamatakos 2011). Multifocal tumours are staged according to the size of the largest lesion (Lakhani et al. 2012). Multifocality refers to multiple invasive foci within the same breast quadrant, clearly separate (often more than 5 mm apart) and not connected with DCIS; in comparison to multicentricity, which has been defined as tumour foci in multiple breast quadrants (Ellis et al. 2016; Middleton et al. 2002).

The microscopic view together with immunohistochemical and other analyses reveal growth patterns and tumour biology, which are particular for several different invasive breast cancer types. The current WHO classification of breast tumours lists 30 different invasive epithelial tumour types with a specific morphology code. The most common types are the invasive carcinoma of no special type (also known as invasive ductal carcinoma) and invasive lobular carcinoma. DCIS is a precancerous finding, with no invasive characteristics, and theoretically no metastatic tendency or mortality. (WHO Classification of Tumours Editorial Board 2019). Unlike most invasive carcinomas, DCIS appears rarely as a precise round lump since it affects the inside of the breast duct system. Therefore, determining the size of DCIS lesions can often lead to underestimation (Ellis et al. 2016).

Together with tumour histology and local extent, the final pathological anatomic diagnosis (PAD) concludes all other predictive and prognostic factors: ER and PR status are counted as the percentage of immunoreactive tumour nuclei cells; < 1 % negative, ≥ 1 % positive, 1-10 % low positive (Allison et al. 2020). HER2 status is

determined by protein overexpression in immunohistochemistry, or by in situ hybridization (Wolff et al. 2018). Proliferation marker Ki-67 is a protein detected by immunohistochemistry in proliferating tumour cells. Peritumoural lymphovascular invasion refers to tumour cells present in lymphatic or venous channels. (Ellis et al. 2016) Grade is determined by tubule formation, pleomorphism of the nuclei and mitotic count (Lakhani et al. 2012). Tumour-infiltrating lymphocytes show the immune system's lymphocytic activity in the tumour stroma, which is counted as the percentage of affected stromal area (McCart et al. 2019). Also, excision margins are reported (Ellis et al. 2016). Table 4 shows a summary of the current histopathological chart used in Turku University Hospital for all routine pathological analyses in breast cancer.

Table 4. Routine pathological anatomic analyses of dissected breast cancer samples at Turku University Hospital. Modified from Ellis et al. (2016).

Histological tumour type
Largest tumour diameter (mm)
Shortest surgical margin (mm)
Direction of the shortest surgical margin (clockwise)
Distance to skin (mm)
Lymphovascular invasion (present / not present)
Histological grade (1–3)
Nuclear grade if DCIS or invasive lobular carcinoma (1–3)
Number of ALN metastases / Number of all examined ALNs
Number of SLN metastases / Number of all examined SLNs
Isolated tumour cells (Present / Not present)
Definition of prognostic markers (From needle biopsy / From surgical dissection sample)
Percentage of ER positive cells
Percentage of PR positive cells
HER2/neu staining (0, 1+, 2+, 3+)
HER2/neu in situ hybridisation (+/-)
Ki-67 proliferation index (%)

2.4.3 Axillary lymph node samples

During the primary breast cancer operation, detected SLNs can be sent to the pathology laboratory as frozen section samples. These samples are investigated intraoperatively in order to determine the need for ALND immediately. (Lee & Sabel 2011) One in four clinically node negative patients will have a positive SLN (Mansel et al. 2006). SLN frozen section has high false negative rates, especially in minimally involved lymph nodes, with a sensitivity of 40 % for ITC or micrometastases, and 94 % for macrometastases; and specificity of 100 % (Liu et al. 2011). It has been hypothesized that tissue loss in the cutting process of frozen section samples could result in under-staging (Layfield et al. 2011). The recent development in axillary management has led to decreasing use of frozen section, and the SLNs are now mostly examined postoperatively as paraffin-embedded samples (Jorns & Kidwell. 2016; Lombardi et al. 2018). Therefore, the number of positive SLNs and the size of the metastases are evaluated postoperatively according to the current recommendations, and the need for ALND in a second operation is decided thereafter.

After axillary clearance, collecting all lymph nodes from the dissected material and slicing them thinly for microscopic evaluation is preferable in order to detect all metastatic deposits and to result in the most accurate nodal staging. The nodes are fixed in formalin and stained with haematoxylin and eosin, by standard. The size of the lymph node metastasis is categorised into ITC (≤ 0.2 mm in size, or < 200 cells), micrometastasis (> 200 cells, or > 0.2 mm but no more than 2 mm in size) and macrometastasis (> 2 mm in size). (WHO Classification of Tumours Editorial Board 2019) A meta-analysis has shown that 61.4–83.3 % of patients with a SLN metastasis do not have other non-SLN metastases (Wang et al. 2011). In the subgroup of patients with SLN micrometastases or ITC, only 7.5–15 % will have further metastases (Cserni et al. 2004; Meretoja et al. 2011). Also, extracapsular extension (ECE) of lymph node metastases is reported (Ellis et al. 2016).

2.5 Statistics in locoregional stage assessment

2.5.1 Nomograms in breast cancer

Nomograms are mathematical models, formulas or scoring systems which aim to statistically predict the probability of a certain outcome. By strict definition, nomograms are risk-estimation tools of multiple variables and with continuous numerical results (Matsen & Van Zee 2014). The most important components of cancer nomograms are presented in Table 5. In breast cancer, nomograms have been used to estimate prognosis, treatment outcomes, histologic tumour size, distant

metastases, and the number of ALN metastases. (Balachandran et al. 2015; Bosch et al. 2003; Boutros et al. 2015) Concerning primary tumour size and tumour stage, some research has been conducted on developing predictive models: Bosch et al (2003) developed a formula based on imaging features to predict the pathological tumour size. Several other studies have developed nomograms based on clinopathological factors that statistically associate with decline in tumour size, or pathological complete response after neoadjuvant systemic therapy (Keam et al. 2011; Kim et al. 2015; Rouzier et al. 2006; Yan et al. 2020). Kondo et al (2015) presented a model to predict upstaging of preoperative DCIS diagnosis into invasive carcinoma in postoperative evaluation; factors predicting invasive histology were sclerosing adenosis, pleomorphic calcifications in mammography, suspicion of invasiveness in ultrasound or MRI, and tumour size larger than 2 cm in ultrasound examination.

Table 5. Components of cancer nomogram performance (Balachandran et al. 2015).

Discrimination	➢ The ability to differentiate patients undergoing a certain event from patients not undergoing the event. ➢ Reported as the area under the receiver operating characteristic curve (AUC, ROC). ➢ The AUC value 0.5 interprets as not better than chance, and AUC value 1 as perfect discrimination. AUC values > 0.75 are recommended for clinical prediction tools.
Calibration	➢ How precisely the nomogram can estimate risks in comparison to the observed, real-life risk. ➢ Varies in different levels of resulted probabilities. ➢ Should include confidence intervals.
Validation	➢ Testing the nomograms performance in external, unbiased, and separate populations, or as internal validation within the original dataset.
Clinical utility	➢ Does the nomogram improve decision-making and patient outcomes?

2.5.2 Nomograms predicting axillary lymph node status

2.5.2.1 The history and current trends of developing nomograms

During the twentieth century, ALNs were noted as an important prognostic factor and included in the TNM classification (Lakhani et al. 2012). After SLNB was established in breast cancer management, a need to identify the patients who were at low risk of additional lymph node metastases emerged (Matsen & Van Zee 2014). In the 1990s, associations of clinopathological features such as tumour size and lymphovascular invasion were investigated in relation to non-SLN metastases after a positive SLNB (Reynolds et al. 1999; Turner et al. 2000). Later, these studies were followed by multivariable nomograms to answer to the multifactorial aetiology of lymph node metastases (Bevilacqua et al. 2007). A meta-analysis has indicated that tumour size, lymphovascular invasion, metastasis size in the SLN, more than one metastatic SLN, 0–1 negative SLNs, SLN positivity in more than 50 % of SLNs, and ECE are the strongest predictors of non-SLN metastases (van la Parra et al. 2011).

The first multivariable nomogram was the model developed at the Memorial Sloan-Kettering Cancer Center, which has been widely validated afterwards. The model included tumour size, tumour type and grade, lymphovascular invasion, multifocality, ER status, the method of detecting SLN metastasis and the number of positive and negative SLNs. (Matsen & Van Zee 2014; Van Zee et al. 2003) Afterwards, it was noticed that nomograms at the time did not perform adequately if the SLN metastasis was minimal, presenting with ITC or micrometastasis. Therefore, new nomograms were needed and developed (Meretoja et al. 2012) until randomized controlled trials discovered it was unnecessary to conduct ALND after ITC or micrometastasis in the SLNB (Galimberti et al. 2018).

Since the inception of nomograms predicting non-SLN metastases, numerous models have been developed around the world, by using different population cohorts. In validation studies, it was evident that the nomograms performed best in the centre where they were developed, and needed to be validated in other local populations before clinical use (Cserni et al. 2013; Tvedskov et al. 2014; van den Hoven et al. 2015; Wu et al. 2018). Around 2010s, multicentre studies and huge cohort studies gained attention in order to improve heterogenicity of the cohorts, and the external validity of the nomograms (Chen et al. 2017; Meretoja et al. 2012; Werkoff et al. 2009). The ACOSOG Z011 trial (Giuliano et al. 2017 [b]) set a new standard for axillary management, and nomogram development shifted towards predicting the high risk of 3–4 or more lymph node metastases, i.e., the patients that still needed ALND or extensive adjuvant therapies. In these models, preoperatively available clinopathological factors have often been preferred in the nomogram development (Chen et al. 2017; Cserni et al. 2013; Kim et al. 2017; Meretoja et al. 2013).

At least one meta-analysis has tried to determine the best and generally acceptable nomogram. Yet, it has shown that the accuracy of nomograms is often not adequate. (Zhu et al. 2013; Chen et al. 2012) The wide interest for nomograms has resulted in many poor-quality models and studies with insufficient transparency (Siesling et al. 2019). To improve the quality and understanding of cancer nomograms, advisory evaluations have been published, addressing the importance of discrimination, calibration, clinical utility, and sample size (Balachandran et al. 2015; Vergouve et al. 2005). Today, many models exist but none have been established in the international breast cancer guidelines as a part of the preoperative diagnostic and staging work-up (Cardoso et al. 2019; Lyman et al. 2016). The early 2000s and 2010s nomograms remain the most commonly validated models, even though "a nomogram can become less accurate with time for a variety of reasons, such as improvements in therapy, earlier detection, and changes in natural history" (Balachandran et al. 2015).

2.5.2.2 Reasoning behind nomograms

To select appropriate treatments for individual patients, nomograms have been considered helpful tools in determining the nodal stage without surgical interference. Whenever ALND is omitted, an accurate nodal stage cannot be known. Therefore, planning the adjuvant systemic and radiation therapies becomes more difficult. A reliable nomogram can be of good value when SLNB is positive and the need for ALND is considered in a second operation. In the future, nomograms could help determining which patients can avoid all axillary operations, including SLNB. (Meretoja et al. 2014; Ahn et al. 2017; Houvenaeghel et al. 2019). However, literature on the nomograms' clinical utility as randomized controlled trials, or as reviews of their safety, is scarce.

It could be beneficial to extend the landmark clinical trials', such as the ACOSOG Z0011, results safely into other populations, and nomograms could aid in this extension (Chen et al. 2017). Recently, two widely accepted nomograms have been used to estimate risks of non-SLN metastases, to compare the similarity and balance of cohorts in the two treatments arms of the ACOSOG Z0011 trial, and to evaluate whether radiation oncologists treated nomogram-estimated higher-risk patients differently. In this study, the authors concluded that the two nomograms were able to identify patients with higher risk of additional ALN metastases after a positive SLNB, and that higher risk estimations were associated with the more frequent use of supraclavicular radiation fields, but not with high tangent radiation. (Katz et al. 2020). Nomograms can help predict locoregional recurrence (Pepels et al. 2013), and their employment in clinical work could benefit economically (Bonsang-Kitzis et al. 2017).

The limitations of predictive models include the varying set of factors used in the nomograms, the varying methodology and ranging probability results, the limited clinical utility, and the varying accuracy (Chen et al. 2011).

2.5.3 Machine-learning in locoregional stage prediction

Machine-learning is a part of artificial intelligence. It means "programming computers to optimize a performance criterion using example data or past experience" (Alpaydin 2020). A machine-learning model can use even huge amounts of past information to predict the future, or to describe the data, by learning to detect features and patterns in the data (Alpaydin 2020). In contrast to traditional logistic regression, machine-learning uses "automatic learning from data" (Christodoulou et al. 2019). In breast cancer, machine-learning algorithms have been used for example to detect and diagnose cancerous breast lesions in imaging (Fusco et al. 2016; Yassin et al. 2018) and in microscopic cytology samples (Saha et al. 2016), to estimate the risk for breast cancer (Nindrea et al. 2018), to estimate breast cancer recurrence (Izci et al. 2020) and survival (Montazeri et al. 2016), and to detect lymphoedema (Fu et al. 2018) or depression on breast cancer survivors (Cvetković 2017).

The most promising and practical of the above-mentioned applications might be the image analysing models. In mammography screening, machine-learning may assist in automated cancer detection. (Le et al. 2019) Machine-learning could possibly reduce the amount of work and cost needed to interpret MRI images, reduce observer-related variance, and improve specificity. The average sensitivity and specificity of different machine-learning algorithms in detecting breast cancer in MRI were 90 % and 85 % in a previous systematic review. The construction of a machine learning model is however not entirely straightforward, as the features which classify for example cancerous and benign lesions in MRI must be carefully chosen, and the classifier results need to be understandable for the users. Moreover, internationally standardized imaging protocols and image databases could improve the research in cancer detection by machine-learning. The introduction of machine-learning into the clinical work can be challenging as the lack of time and effort from physicians to adopt these tools in their work, together with costs and false positive or negative findings, hinder the onset of wide machine-learning use (Fusco et al. 2016; Yassin et al. 2018)

In predicting ALN metastases, most nomograms have been developed with well-known logistic regression. During the recent years, machine-learning methods have been employed more frequently: artificial neural networks (Nowikiewicz et al. 2014; 2017), support-vector machines (Ding et al. 2013; Kim et al. 2017) and alternative decision tree method (Sugimoto et al. 2014) have been used. Some studies have also applied machine-learning in creating radiomic signatures (digital images converted

into data) from medical imaging examinations, and then by using the radiomics and logistic regression developed a nomogram (Han et al. 2019; Qiu et al. 2020). Nevertheless, a systematic review of clinical prediction models has noted that machine-learning models do not perform superiorly to logistic regression, and comparative studies of these two methods should improve in quality by i.a. reporting the results of calibration (Christodoulou et al. 2019).

3 Aims

The aim of this study was to evaluate how locoregional stage can be assessed in clinically node negative breast cancer, emphasising the factors which associate with non-SLN metastases, and how current clinical, imaging, pathological and statistical modalities relate to tumour and nodal stage. The hypothesis of this study was that clinopathological characteristics can be used to predict how widely breast cancer spreads to local lymph nodes, and to select patients for additional MRI in breast cancer diagnostics. MRI was hypothesized to affect patients' primary surgical treatment chronologically and in extensiveness. The aims of studies I–IV were:

I To evaluate which predictive factors are related to higher numbers of lymph node metastases, and whether there are metastatic findings present when ALND is conducted in a second, delayed operation. The survival and the treatment-related morbidity of patients treated with a delayed ALND was evaluated.

II To create a mathematical model, a nomogram, that would predict four or more ALN metastases using patient and tumour related clinopathological variables, indicating high nodal tumour burden after a positive SLNB. The aim was to use modern machine-learning methods to create the nomogram.

III To validate predictive nomograms that have been developed in other institutions globally several years ago, and that cannot be used in external populations without further validation. The aim was to confirm the nomograms' validity in a current patient population to predict the nodal stage pN2a or pN3a.

IV To assess how current imaging modalities associate with locoregional stage evaluation and treatment in breast cancer, by investigating whether additional MRI of breasts and axillary regions delays the primary surgery or is followed by more frequent employment of mastectomy rather than breast-conserving surgery in the primary operation, and whether the primary tumour size and ALNs can be reliably estimated by different imaging modalities in relation to final histopathological findings.

4 Materials and Methods

4.1 Study population

The patient samples for this study consisted of newly diagnosed primary breast cancer patients treated in Turku University Hospital in 2009–2017. The study was conducted under permissions granted by the University of Turku, The Hospital District of Southwest Finland, and Turku Clinical Research Centre (T119/2017). At the time, the local ethics committee's approval was not mandatory for register studies. The committee was consulted about the study design of studies II–IV (ETMK:56/1801/2017). The hospital medical records of departments of oncology, surgery, radiology, and pathology were searched to identify eligible patients:

I Patients treated for primary breast cancer surgery, and including SLNB, between September 2009 and August 2012 were evaluated. Inclusion criterion was at least one detected SLN metastasis. Recipients of neoadjuvant systemic therapies were excluded. Additionally, comparative patients with negative SLNs were enrolled. Records for eligible N=162 patients with 169 operated breasts from unilateral and bilateral cases and 168 SLNBs were further analysed. Patients with a delayed ALND were evaluated again, 4–7 years after the primary surgeries in late 2016, for cancer recurrence and possible adverse effects.

II–III Patients treated for primary breast cancer between September 2012 and December 2017 were evaluated. Of these 3215 patients, an eligible cohort of N=581 patients was determined according to the criteria: primary breast cancer diagnosis, SLNB and a following ALND, with at least one confirmed metastatic lymph node. ALND was either a part of the primary operation or conducted as a delayed second operation. After removing 10 or more lymph nodes, the patient in question was included even if ALND was not technically performed. Patients with male gender, recipients of neoadjuvant systemic therapies, bilateral breast cancer, only benign nodes, four positive SLNs, or relevant clinopathological information missing were excluded. Therefore, the patient samples in studies II and III were N=530 and N=529, respectively.

Figure 3 summarizes patient inclusion in studies II–III. In study II, two sub-cohorts were set; patients treated in 2012–2016 (N=460) formed the training cohort, and patients treated in 2017 (N=70) formed the validation cohort. In study III, alternative two sub-cohorts were set; a validation cohort for the nomograms developed by Chagpar et al. (2007) and Katz et al. (2008), N=529, and after excluding patients with ITC or micrometastasis in SLNB, a validation cohort for the nomogram developed by Meretoja et al. (2013), N=351.

IV Patients eligible in study II and treated in 2012–2016 were included (as described above and in Figure 3), and in addition, one more patient with required information was returned from previous exclusions, resulting in a cohort of N=461 patients. A sub-cohort of patients, N=96, who underwent preoperative MRI together with standard imaging, were evaluated to investigate the performance of preoperative imaging on ALN metastases.

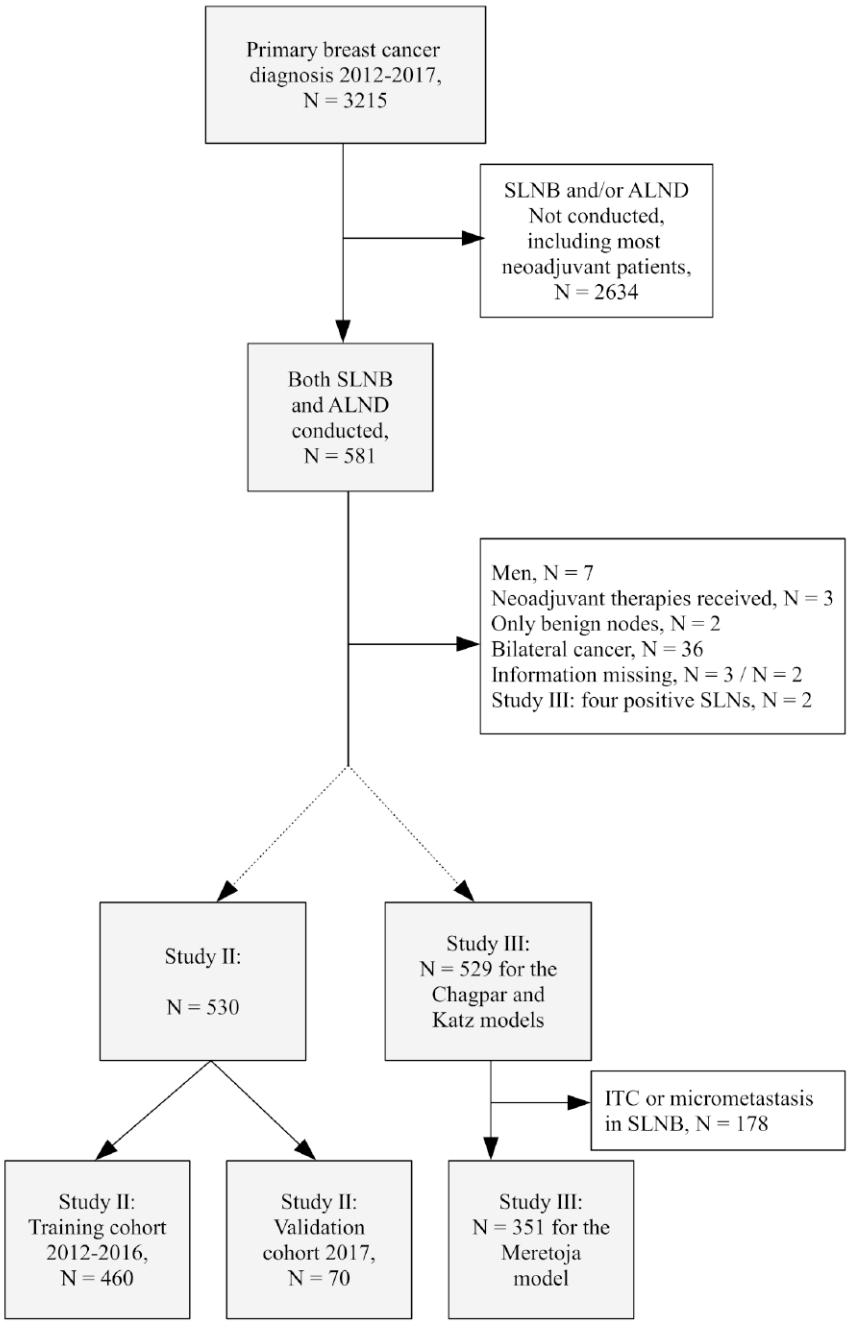

Figure 3. Flowchart of the patient populations included in studies II and III. Modified from Original publications II and III.

4.2 Patient, tumour, and nodal characteristics

Clinopathological variables were collected in order to assess their associations to lymph node metastases, surgical treatment and imaging examinations. Collected variables in each study are shown in Table 6. In study III, clinopathological variables were recorded according to the three nomograms chosen for validation (Chagpar et al. 2007; Katz et al. 2008; Meretoja et al. 2013).

Table 6. Recorded clinopathological variables in studies I–IV

	Study I	Study II	Study III	Study IV
Patient age	●	●	●	●
Method of detecting tumour		●		●
Method of surgery	●	●		●
Dates of tumour detection, surgery and final PAD				●
Tumour palpability		●		●
Tumour size in pathology[a]	●	●	●	●
Tumour size in mammography				●
Tumour size in ultrasound				●
Tumour size in MRI				●
ALN status in imaging				●
Tumour laterality and location within the breast	●	●	●	
Tumour histology	●	●	●	●
Tumour multifocality		●	●	●
DCIS presentation	●			
Tumour grade	●	●	●	●
ER status	●	●		●
PR status	●	●		●
HER2 status	●	●	●	●
Ki-67	●	●		●
Lymphovascular invasion	●	●	●	●
Biologic subtype	●	●		●
SLN metastasis size[b]	●	●	●	
Number of positive and negative SLNs	●	●	●	
Number of positive and negative ALNs	●	●	●	●
Extracapsular extension			●	

[a] Multifocal tumour size was calculated as the combined diameter of the lesions.
[b] Categorized as ITC, micrometastasis or macrometastasis.

The surgical method included breast conserving surgery with SLNB and ALND, or mastectomy with SLNB and ALND. Delayed ALNDs were recorded. SLNB findings in both frozen section and paraffin embedding were collected. The extend of malignant deposits in SLNB was categorized into ITC, micrometastases and macrometastases both in intraoperative and postoperative analyses. Mere ITC findings were not considered ALN metastases in studies II–IV although they were marked as lymph node positivity in study inclusion.

Tumours were categorized into prognostic biologic subtypes accordingly: luminal A-like, luminal B-like, HER2-enriched type and triple-negative type. The distinction between luminal A and HER2 negative luminal B-like tumours was made according to the St. Gallen International Consensus Guidelines of the time, in which luminal A-like tumours have low Ki-67 values of under 14 % and HER2 negative luminal B-like tumours have high Ki-67 values of over 14 % (Goldhirsch et al. 2011). Otherwise, the biologic subtypes were determined as presented in Table 1, and therefore HER2 positive tumours were categorized either as luminal B-like tumours or as HER2-enriched type tumours depending on hormone receptor positivity. In study I, basal type tumours were distinguished from triple-negative tumours based on CK5/6 and EGFR positivity.

Multifocality referred to two or more invasive tumour lesions at any separating distance. Contrary to the TNM classification, multifocal tumour size was calculated as the combined diameter of the lesions, to reflect the entire tumour burden. Tumour location within the breast was adapted from a previous publication (Desai et al. 2018).

4.3 Imaging methods

According to national and international instructions, all patients underwent preoperative bilateral mammography and ultrasound examination of breast and axillary regions. Ultrasound-guided preoperative fine needle aspiration was executed on suspicious lymph nodes. In study IV, MRI scanners of 1.5 T and 3.0 T (Magnetom Avanto and Magnetom Aera, Siemens) were used to examine selected patients after a multidisciplinary breast cancer group's recommendation. The national recommendations at the time suggested preoperative MRI if the triple investigation was inconclusive or mammography together with ultrasound examination was difficult to interpret prior to breast-conserving surgery. Additionally, preoperative MRI was suggested for patients with occult breast cancer. (Finnish Breast Cancer Group 2013).

4.4 Histopathological methods for axillary lymph nodes

The original histopathological records of the study population were evaluated to collect the pathologic information. During the primary breast cancer diagnostics and treatment, the histopathological analyses had been conducted followingly: Intraoperative frozen section and postoperative paraffin embedding were used to investigate the SLNs. SLNB consisted of 1–6 SLNs for every patient. Lymph node slicing of 2–3 mm intervals was employed, and two HE slides (Haematoxylin solution according to Delafield, SIGMA ALDRICH 03971; Eosin solution, Reagens 180072) and two rapid immunohistochemistry slides for pan-anticytokeratins (Cytonel-Plus ULTRARAPID IHC, Jilab, Finland) were examined during the frozen section analysis. The final histopathological diagnosis determined the ALN status as lymph nodes were fixed in formalin and embedded in paraffin with HE staining and cytokeratin immunohistostaining (Anti-Pan Keratin AE1/AE3/PCK26, Primary Antibody, Ventana, Roche).

4.5 Nomograms for validation

The multi-institutional model from the United States and Canada by Chagpar et al. (2007), the nomogram by Katz et al. (2008) from two academic centres in the US, and the European international multicentre model by Meretoja et al. (2013) were selected to be validated. Two of these authors note that their nomogram should be validated externally prior to clinical use elsewhere (Katz et al. 2007; Meretoja et al. 2013). The patient inclusions of these studies were conducted from 1998 to 2004 for the Chagpar model and from 2004 to 2011 for the Meretoja model. The study by Katz et al. lacked this information.

The Chagpar model (2007) incorporates the number of positive SLNs, the ratio of positive and negative SLNs (50 % or more as a risk factor) and the T stage. It scores 1–5 points, and a low < 5 % risk of stage N2–3 is indicated by one point. The Katz model (2008) includes seven variables: tumour size, number of positive and negative SLNs, lymphovascular invasion, tumour histology, ECE of the SLN metastasis and SLN metastasis size categorised as ITC, micrometastasis or macrometastasis. It performs as a mathematical scoring system with a resulting risk estimation of 0.005–0.99 for stage N2–3. The Meretoja model (2013) incorporates the prevalence of N2 stage in the patient series, tumour size, number of positive and negative SLNs and ECE. It is available as an online calculator and as a mathematical equation with a resulting continuous risk estimation level.

4.6 Statistical methods

In study I, the correlation between continuous variables and the number of lymph node metastases was calculated by using Spearman rank-order correlation coefficient. To compare differences in the number of lymph node metastases between the levels of categorical variables, Mann-Whitney U-test or Kruskall-Wallis test with Dwass-Steel method in pairwise comparisons were employed. To assess skewed distributions, non-parametric methods were used. The patients with no ALN metastases were analysed so that the associations of different variables with the number of ALN metastases would better be presented among the study population.

In studies II–IV, patient and tumour characteristics were presented with means and standard deviations (SD) for continuous variables which followed normal distribution, otherwise with medians and interquartile range (IQR). Visual method, Shaphiro-Wilk's test and Q-Q plot were used to ensure the normality distribution assumption of continuous variables. To report categorical variables, counts and percentages were used.

In study II, the characteristics of the training cohort were analysed in order to develop a nomogram which estimates the probability of four or more lymph node metastases (nodal stage pN2a or pN3a). To determine which clinopathological factors were associated with stage pN2a or pN3a, machine-learning methods and logistic regression analysis were used. In machine-learning, self-learning abilities were utilized in nomogram development. The gradient boosted trees model (XGBoost) was employed to investigate which variables formed the best combination for the model, and to develop the nomogram. Manual selection of variables by trial and error was used after XGBoost's employment to decrease the number of variables in the final nomogram. Therefore, an accurate model was obtained with fewer variables. Multivariable logistic regression model was created from the variables selected by the XGBoost method.

The training cohort was five-time cross-validated, and the cross-validation was repeated 100 times with different randomization. The discrimination of the XGBoost was investigated and hyperparameters were tuned by determining the mean area under the receiver operating characteristic curve (AUC value) and the 95 % confidence intervals (CI) for the AUC from the 500 cross-validation runs. The XGBoost model was tested in the validation cohort and bootstrap analysis was used 10,000 times to present a mean AUC value and 95 % CIs for the validation cohort. For logistic regression model, the training cohort was also five-time cross-validated, and the AUC values were determined similarly. Calibration curves for the XGBoost and logistic regression models were calculated to investigate how well the nomograms were able to predict the real-life observed number of lymph node metastases. Calibration was reported as risk estimations in quantiles (0–0.25, 0.25–0.5, 0.5–0.75, 0.75–1).

In study III, the predicted risk for four or more lymph node metastases (stage pN2a or pN3a) was calculated for individual patients by using predictive models developed by Chagpar et al. (2007), Katz et al. (2008) and Meretoja et al. (2013). The Katz and Meretoja models were employed as mathematical equations, whereas the Chagpar model is a scoring system. Discrimination was determined and reported as receiver operating characteristic curves (ROC curves) and AUC values. The scores of the Chagpar model were mapped to probabilities determined in the original study, to draw the ROC curves and calculate AUC values. Calibration curves were presented to demonstrate the actual risk-estimation levels of the models in quantiles (0-0.25, 0.25-0.50, 0.50-0.75, 0.75-1.0).

In study IV, patient age was reported as a continuous variable as well as following screening age categories, < 50 years old or ≥ 50 years old. The time between tumour detection and surgical treatment or final postoperative PAD was determined from the first tumour detection date in any imaging examination to the breast surgery date of the primary tumour, and to the date of final PAD report of the breast cancer dissection specimens. The frequencies of categorial variables were compared by Chi-square test or Fisher exact test. Two-sample t-test, or 1-way analysis of variance using Tukey's method in pairwise comparisons, were employed to compare normally distributed continuous variables. The differences in non-normally distributed continuous variables were investigated with Mann-Whitney U test and Kruskal-Wallis tests with Dwass-Steel-Critchlow-Fligner method in pairwise comparisons. Wilcoxon signed rank was employed to evaluate mean differences between tumour size imaging modalities and final PAD. Spearman correlation coefficients were used to estimate correlations between tumour size in final PAD and the difference in tumour size between final PAD and imaging modalities. The frequencies of tumour-free and suspicious lymph nodes in imaging modalities were compared with McNemar's test. A multivariable logistic regression analysis was conducted based on statistically significant clinopathological variables in univariable analyses. Odds ratios (OR) with 95 % CIs were used to report the independent associations of clinopathological factors with MRI imaging in logistic regression analysis.

In studies II–IV, JMP Pro 14 for Windows was used for statistical reporting of patient and tumour characteristics. In studies II and III, Python 3.6: Anaconda, Inc. programming language and XGBoost 0.90 were used for the machine-learning analysis and sklearn 0.21.2 was used for logistic regression analysis. In studies I and IV, the SAS System for Windows, release 9.4 (SAS Institute Inc., Cary, NC) was used for statistical analyses. Significance level (p-value) less than 0.05 was considered statistically significant.

5 Results

5.1 Factors predictive for axillary lymph node metastases (I)

The characteristics of 162 patients and 169 tumours treated in 2009–2012 are reported in Table 7. In addition to standard SLN metastasis detection, three patients had malignant tumour cells found in fatty tissue surrounding the SLNs, resulting in ALND. Negative frozen section results remained minimal (negative 47 %, cytokeratin positive only 1 %, ITC 23 % or micrometastases 29 %) in paraffin section. ITC in frozen section were upstaged to micrometastases in 3 (17 %) cases. Micrometastases in frozen section, however, were upstaged to macrometastases in 4 (20 %) cases. The number of cases with non-SLN metastases detected in ALND was 35 (33 % of positive SLN cases).

Table 7. Patient and tumour characteristics in study I, modified from Original publication I. A selected cohort of clinically node negative patients treated in 2009–2012, N=162 patients and N=169 operated breasts.

Variable	
Mean age, years (range; SD)	62 (28–89; 12.2)
Tumour laterality	
Left breast	76 (45 %)
Right breast	93 (55 %)
Multifocality present	39 (23 %)
Primary breast operation	
Breast-conserving surgery with SLNB	51 (30 %)
Breast-conserving surgery with SLNB and immediate ALND	66 (39 %)
Mastectomy with SLNB	14 (8 %)
Mastectomy with SLNB and immediate ALND	38 (23 %)
Delayed ALND performed	29 (17 %)
Median tumour size (IQR)	1.8 cm (1.5)

Variable	
Tumour histology	
Invasive ductal carcinoma	135 (80 %)
Invasive lobular carcinoma	17 (10 %)
Invasive ductal and lobular carcinoma	4 (2 %)
DCIS only	8 (5 %)
Presence of DCIS	
Present	98 (58 %)
Not present	71 (42 %)
Tumour grade	
I	25 (15 %)
II	92 (54 %)
III	53 (31 %)
Lymphovascular invasion present	17 (10 %)
ER positive tumours	148 (92 %)
PR positive tumours	126 (78 %)
HER2 positive tumours	21 (13 %)
Ki-67 > 14 %	111 (67 %)
Biologic subtype	
Luminal A-like	50 (30 %)
Luminal B-like	98 (58 %)
HER2-enriched	5 (3 %)
Triple-negative	6 (4 %)
Basal type	2 (1 %)
SLN metastasis size in frozen section	
No metastases	66 (39 %)
ITC	18 (11 %)
Micrometastasis	20 (12 %)
Macrometastasis	62 (37 %)
Positive cytokeratin staining only	2 (1 %)
SLN metastasis size in paraffin block	
No metastases	40 (24 %)
ITC	27 (16 %)
Micrometastasis	39 (23 %)
Macrometastasis	61 (36 %)
Positive cytokeratin staining only	1 (1 %)
Total number of ALN metastases	
0	62 (36 %)
1–3	90 (53 %)
4 or more	17 (10 %)

Tumour size was associated with more ALN metastases (Spearman, r=0.22, p=0.004). Luminal B-like tumours had more ALN metastases than luminal A-like tumours (median 1.0 vs. 0.5, p=0.002), and high Ki-67 was a risk factor for more ALN metastases (r=0.23, p=0.003). However, with Ki-67 values higher than 30 %, the difference was not statistically significant. Patient age, tumour grade, HER2 positivity, presence of DCIS, or ER and PR status were not statistically significantly associated with the number of ALN metastases.

5.2 Prediction of four or more non-sentinel lymph node metastases (II, III)

5.2.1 Developing and testing a predictive nomogram (II)

Of the 460 patients treated in 2012–2016 and used as the training cohort for the nomogram development, 380 (83 %) were staged pN0–1 and 80 (17 %) were staged pN2a or pN3a. Of these 460 patients, 261 (57 %) and 199 (43 %) were operated with breast-conserving surgery and mastectomy, respectively. Of the patients with stage pN0–1, 128 (34 %) had tumour detection by screening mammography, 300 (79 %) had palpable tumours, 311 (82 %) had pure invasive ductal histology and 53 (14 %) had pure invasive lobular histology, 114 (30 %) had multifocal tumours, 358 (94.2 %) had ER positivity, 354 (93.2 %) had PR positivity, median Ki-67 value was 20 (IQR 18), 108 (28 %) had luminal A-like tumours, 248 (65 %) had luminal B-like tumours, 6 (1.6 %) had HER2 type tumours, 18 (4.7 %) had triple-negative tumours, 38 (10 %) had lymphovascular invasion, 252 (66 %) had one metastatic SLN, 154 (41 %) had ITC or micrometastases in SLNB, the median tumour size was 2.3 cm (IQR 1.7), and the mean age was 63.3 (SD 12.7) years.

Of the 80 patients with stage pN2a or pN3a in the training cohort, 23 (29 %) had tumour detection by screening mammography, 64 (80 %) had palpable tumours, 55 (69 %) had pure invasive ductal histology and 22 (28 %) had pure invasive lobular histology, 43 (54 %) had multifocal tumours, 70 (88 %) had ER and PR positivity, median Ki-67 value was 20 (IQR 21.75), 23 (29 %) had luminal A-like tumours, 49 (61 %) had luminal B-like tumours, 1 (1.3 %) had HER2 type tumours, 7 (8.8 %) had triple-negative tumours, 20 (25 %) had lymphovascular invasion, 30 (38 %) had one metastatic SLN, 2 (2.5 %) had ITC or micrometastases in SLNB, the median tumour size was 3.3 cm (IQR 3.2), and the mean age was 62.6 (SD 12.5) years.

In the validation cohort of 70 patients treated in 2017, 61 (87 %) patients were staged pN0–1 whereas 9 (13 %) patients were staged pN2a or pN3a. The validation cohort was similar to the training cohort concerning most clinopathological factors, except the proportion of macrometastases in SLNB which was 79 % in the training cohort and 65 % in the validation cohort.

The XGBoost modelling resulted in the best combination of clinopathological factors which had the highest association with stage pN2a or pN3a. The variables in the model were tumour size (in mm), tumour histology (pure invasive ductal other), tumour multifocality (present or not present), percentage of ER positive cells, number of positive SLNs and number of positive SLNs multiplied by tumour size. Using these variables, the multivariable logistic regression model for the probability p of stage pN2a or pN3a was constructed:

$$\text{logit}(p) = -1.732 - 0.732*a + 0.534*b + 0.792*c - 0.017*d + 1.863*e + 0{,}029*f - 0{,}004*g$$

In this equation, variables are placed as follows: a = tumour histology (pure invasive ductal carcinoma = 1, other = 0), b = multifocality (multifocal = 1, unifocal = 0), c = lymphovascular invasion (present = 1, not present = 0), d = percentage of ER positive cells, e = number of positive SLNs, f = tumour size in mm, g = number of positive SLNs multiplied by tumour size in mm. The number of positive SLNs does not include ITC.

By using the XGBoost model and the logistic regression model, the probability (0–1.0) for stage pN2a or pN3a was counted for each patient in the training and validation series. In the training cohort, the AUC values for the XGBoost and logistic regression models were 0.80 (95 % CI 0.71–0.89) and 0.85 (95 % CI 0.77–0.93), respectively. This means that the models performed well in discriminating between patients with low and high nodal stage. To validate this notion, the AUC values in the separate validation cohort were 0.80 (95 % CI 0.65–0.92) for the XGBoost model and 0.75 (95 % CI 0.58–0.89) for the logistic regression model.

The nomograms' ability to predict the real-life observed number of lymph node metastases was evaluated by drawing calibration curves. According to the calibration process, the models are not ideally calibrated in the validation cohort. This can be explained by the small number of cases in each quantile. The risk for nodal stage pN2a or pN3a predicted by these models cannot therefore be interpreted as the true risk, according to the small validation sample of this study. The model calibration varies on different predicted risk levels. However, calibration alone is not a measure of accuracy. When the model aims to result in "yes or no" categories, unfavourable calibration can be accepted.

5.2.2 Validation of previous predictive models (III)

In study III, the patient cohort form 2012–2017 was used in the validation of the Chagpar, Katz and Meretoja models to detect the stage pN2a or pN3a. In final histopathology, the number of metastatic ALNs was 0 to 3 on most clinically node

negative patients (N = 442, 83.6 %), while four or more metastatic SLNs was present in 87 (16.4 %) patients. If the final nodal stage was pN2a or pN3a, all these patients had micro- or macrometastases in the SLNB, most of them (N = 85, 97.7 %) being macrometastases. Hence, ITC in SLNB did not result in stage pN2a or pN3a in any cases. Micrometastases in SLNB resulted in stage pN2a or pN3a in 2 (1.9 %) cases, both of which being of pure invasive ductal histology. The mean age of patients with stage pN0–1 was 63.2 (SD 12.7) and 64.2 (SD12.9) with stage pN2a or pN3a.

Of the 442 patients with stage pN0–1 treated in 2012–2017, 72 (16.3 %) had ITC only in SLNB, 297 (67.2 %) had one positive SLN, 68 (15.4 %) had two positive SLNs and 5 (1.1 %) had three positive SLNs. Most of these patients (N = 164, 37.1 %), had metastases in all removed SLNs, or one negative SLN among the removed SLNs (N = 146, 33.0 %). Of these 442 patients, 266 (60.2 %) had macrometastases in SLNs. Furthermore, 358 (81.0 %) patients had pure invasive ductal histology, 65 (14.7 %) had pure invasive lobular histology, 400 (90.5 %) had stage T1 or T2, 134 (30.3 %) had multifocal tumours, 47 (10.6 %) had lymphovascular invasion, 18 (4.1 %) had ECE, 47 (10.7 %) had HER2 positive tumours, 319 (72.2 %) had tumour grade I or II, and 184 (41.6 %) had tumour location in the outer quadrants. The median tumour size was 2.3 cm (IQR 1.7).

Of the 87 patients with stage pN2a or pN3a treated in 2012–2017, 0 (0 %) had ITC only in SLNB, 34 (39.1 %) had one positive SLN, 16 (18.4 %) had two positive SLNs and 13 (14.9 %) had three positive SLNs. The majority of these patients (N = 66, 75.9 %) had metastases in all removed SLNs, and 85 (97.7 %) patients had macrometastases in SLNs. Additionally, 58 (66.7 %) patients had pure invasive ductal histology, 25 (28.7 %) had pure invasive lobular histology, 64 (73.6 %) had stage T1 or T2, 44 (50.6 %) had multifocal tumours, 22 (25.3 %) had lymphovascular invasion, 14 (16.1 %) had ECE, 7 (8.0 %) had HER2 positive tumours, 57 (65.5 %) had tumour grade I or II, and 37 (42.5 %) had tumour location in the outer quadrants. The median tumour size was 3.3 cm (IQR 3.1).

In this cohort, the models performed well and the discrimination between patients with and without nodal stage pN2a or pN3a was satisfactory. Table 8 reports the AUC values of the original studies and the current study. The discrimination was presented as ROC curves of these models. According to the curves and AUC values, the Katz model was slightly more reliable than the other two. Calibration curves were drawn, showing moderate calibration in the validation cohort. The risk estimations in the Chagpar model were placed between 0 to 0.5, and therefore the calibration might be considered poor or affected by the small sample size. On the other hand, the calibration curve for the Meretoja model was acceptable.

Table 8. Discrimination of the three validated nomograms in the current patient cohort (Original publication III) and in the original studies, with reported 95 % confidence intervals.

	AUC in the current patient series (95 % CI)	AUC in the training cohort of the original study (95 % CI)	AUC in the validation cohort of the original study (95 % CI)
Chagpar et al. 2007	0.79 (0.74 to 0.83)	0.88 (0.86 to 0.91)	0.90 (0.87 to 0.92)
Katz et al. 2008	0.87 (0.83 to 0.91)	0.83	0.81
Meretoja et al. 2013	0.82 (0.76 to 0.86)	0.77 (0.72 to 0.82)	0.77 (0.74 to 0.81)

5.3 Lymph node findings and outcomes after delayed axillary lymph node dissections (I)

Of the 162 patients treated in 2009–2012, 28 (17 %) had an ALND performed in a delayed second operation, one patient on both axillas. However, this sample was rich in delayed ALNDs because of the study question of interest. All delayed ALNDs in study I were performed due to ITC or micrometastasis in SLNB according to the clinical practice of the time. After ALND, a total of 10–31 ALNs were removed on these patients. Only three (10 %) of these patients had further non-sentinel lymph node metastases in the dissected axillary specimen. Two (7 %) of these 28 patients, both with triple-negative grade III invasive ductal tumours, had deceased of breast cancer by the follow-up in 2016. Four (14 %) had significant adverse effects after the ALND, including lymphoedema, pain and tightness. In 2012–2016, 27 (5.9 %) ALNDs were performed in a delayed second operation in study II. During these years, 291 (63.3 %) cases of ALNDs resulted in no further ALN metastases.

5.4 Preoperative imaging investigations on clinically node negative patients (IV)

In the group of clinically node negative patients treated for breast cancer in 2012–2016, MRI was conducted on 198 (42.9 %) patients with the mean age of 58.7 (SD 11.2) years, in comparison to the 263 patients without MRI and with the mean age of 66.5 (SD 12.6) years ($p < 0.0001$).

Patients with MRI were more often younger than screening age, than those without MRI (20.7 % vs. 11.4 %, $p = 0.0062$). Similarly, patients with MRI had bigger tumours (2.8 cm vs. 2.3 cm, $p = 0.0016$), more often invasive lobular histology (27.8 % vs. 7.6 %, $p < 0.0001$), higher grade tumours ($p = 0.0459$), HER2 positivity (14.7 % vs. 6.8 %, $p = 0.0058$), and multifocal tumours (40.9 % vs. 28.9

%, p = 0.0071). Lymphovascular invasion, ER or PR status, Ki-67 value, biologic tumour subtype or tumour palpability did not differ between patients with or without preoperative MRI.

5.4.1 From tumour detection to operation and postoperative pathological anatomic diagnosis

Most tumours were detected by palpation (60.7 %) or by screening mammography (32.8 %), the rest (6.5 %) being for example incidental findings on other imaging investigations. Patients diagnosed with screening were younger (mean age 60.3 years, SD 5.8) than patients diagnosed with palpation (mean age 64.0 years, SD 14.8) (p = 0.001) or other ways (mean age 69.5 years, SD 11.7) (p=0.0001). Screening detected tumours were smaller than those detected by palpation (p < 0.0001) or other ways (p = 0.359). Patients who had tumours detected by screening, had additional non-SLN metastases in 53 (35.1 %) cases, and patients who had tumours detected by palpation had non-SLN metastases in 108 (38.7 %) cases. The median number of ALN metastases for patients with tumours detected by screening was 1 (IQR 1, range 0 to 30), as well as with tumours detected by palpation (IQR 2, range 0 to 23). There was no statistically significant difference between different detection methods and number of ALN metastases (p = 0.076), multifocality (p = 0.875) or tumour histology (p = 0.230).

Patients who underwent preoperative MRI had a longer time period from tumour detection to the primary breast cancer operation and the final postoperative PAD. For patients with MRI, the median times from tumour detection to operation and final PAD were 43 and 64 days for patients with MRI, and 36 and 57 days for patients without MRI (p < 0.0001 and p = 0.0004, respectively). If a delayed ALND was performed, the median time from tumour detection to final PAD was 81 days, when it otherwise was only 60 days (p = 0.0044).

If MRI was used, mastectomy as the primary surgery was more likely than on patients without MRI (p < 0.0001). Axillary reoperation rates were not different between patients with or without MRI (p = 0.574). Breast reoperation rates were not investigated.

5.4.2 Tumour size in preoperative imaging and postoperative histopathology

In comparison to final PAD, median tumour size was 0.6 cm (IQR 2.3) larger in MRI (p < 0.0001), 0.4 cm (IQR 1.4) smaller in ultrasound examination (p < 0.0001) and 0.2 cm (IQR 1.2) smaller in mammography (p < 0.0001). As tumour size increased, the disagreement increased between tumour size in imaging and histopathology. Yet, this

disagreement was the smallest between tumour size in MRI and tumour size in histopathology. The difference between tumour size in histopathology and tumour size in imaging was most favourable when tumour size was measured in MRI: the Spearman correlation between tumour size in final PAD, and the difference in tumour size between final PAD and MRI was 0.36 ($p < 0.0001$). The comparative Spearman correlations for mammography and ultrasound examination were 0.52 and 0.66, respectively ($p < 0.0001$ for both). Multifocal tumours were present in 157 (34.1 %) cases, their median size being 4.3 cm (IQR 3.8) in MRI, 2.0 cm (IQR 1.6) in ultrasound examination, 2.0 cm (IQR 1.7) in mammography and 3.3 cm (IQR 2.4) in final PAD. Multifocal tumours were larger than unifocal tumours in MRI ($p = 0.0029$), ultrasound examination ($p = 0.0129$) and final PAD ($p < 0.0001$), but not in mammography ($p = 0.2022$). If the tumour consisted of different histological types, it was very likely (75–100 %) to be multifocal. Pure invasive lobular carcinomas had multifocality in 50.7 % of cases, whereas pure invasive ductal carcinomas were multifocal in 29.7 % of cases.

5.4.3 Axillary lymph nodes in preoperative imaging

All clinically node negative patients in the cohort, treated in 2012–2016, were diagnosed with ALN positivity in final PAD. This was due to patient selection, as patients were included in the study only if SLNB was positive and resulted in ALND. Positive ALN findings included ITC, micrometastases and macrometastases. The median number of positive ALNs in histopathology was one (IQR 2, range 0–30), when only micrometastases and macrometastases in ALNs were considered metastases. Even though preoperative conclusions on the nodal stage was cN0 on all patients, some patients had suspicion of positive ALNs in preoperative imaging modalities. Among selected 96 patients with preoperative use of MRI, 10 (10.5 %) patients had suspicion of ALN metastasis in MRI, and 9 (9.5 %) in ultrasound examination; there was no statistically significant difference between these two modalities ($p = 0.782$). Additionally, needle biopsies of the suspicious ALNs did not result in confirming malignancy either, and the patients proceeded to surgical staging by SLNB. Preoperative imaging was unable to detect ALN metastases in most cases of this patient population, as approximately 90 % of the patients with pathologically confirmed ALN metastases were clinically considered node negative.

5.4.4 Predictive factors for the more frequent use of preoperative magnetic resonance imaging and mastectomy in the primary operation

The factors with a p value < 0.05 (patient age, primary operation method, tumour size, tumour histology, multifocality, tumour grade and HER2 status) were included

in a multivariable logistic regression analysis in order to evaluate their independent associations with the use of MRI. According to the analysis, patient age (OR 0.93, 95% CI 0.91 to 0.95, $p < 0.0001$), operation method (OR 2.57, 95% CI 1.60 to 4.13, $p < 0.0001$), tumour histology (invasive lobular vs invasive ductal carcinoma, OR 5.74, 95% CI 2.94 to 11.20, $p < 0.0001$ and other histology vs invasive ductal carcinoma, OR 13.04, 95% CI 2.09 to 81.31, $p = 0.0060$) and HER2 status (OR 2.82, 95% CI 1.34 to 5.93, $p = 0.0064$) had independent associations with the more frequent use of MRI, when age was used as a continuous variable. The result was comparable when age was categorized as < 50 years old or ≥ 50 years old.

All chosen preoperative variables (age, tumour size, tumour histology, multifocality and the use of triple imaging) were statistically significantly associated with the primary operation method in the univariate analysis ($p < 0.05$) (Table 9). The mean age of patients operated with breast-conserving surgery was 62.5 years (SD 10.5), and 63.9 years (SD 14.9) with mastectomy. In the multivariable logistic regression analysis, age (OR 1.03, 95% CI 1.01 to 1.04, $p = 0.0053$), the use of triple imaging (OR 2.39, 95% CI 1.52 to 3.76, $p = 0.0002$) and tumour size (OR 1.05, 95% CI 1.03 to 1.06, $p < 0.0001$) were independently associated with mastectomy rate in the primary operation when age was determined as continuous. When age was categorized as < 50 years old or ≥ 50 years old, the result was similar except age did not remain significant.

Table 9. Characteristics of patients with breast-conserving surgery and mastectomy as the primary operation.

	Breast-conserving surgery, N=262	**Mastectomy, N=199**	**p**
Age: < 50 years old 50 years or older	 32 (12.2 %) 230 (87.8 %)	 39 (19.6 %) 160 (80.4 %)	0.0296
Median tumour size, cm (IQR)	2.1 (1.5)	3.0 (3.0)	< 0.0001
Tumour histology: Pure invasive ductal Pure invasive lobular Mixed invasive ductal and lobular Other	 220 (84.0 %) 30 (11.4 %) 7 (2.7 %) 5 (1.9 %)	 147 (73.9 %) 45 (22.6 %) 5 (2.5 %) 2 (1.0 %)	0.0133
Multifocality: Present Not present	 73 (27.9 %) 189 (72.1 %)	 84 (42.2 %) 115 (57.8 %)	0.0013
Triple imaging: Applied Not applied	 91 (34.7 %) 171 (65.3 %)	 107 (53.8 %) 92 (46.2 %)	< 0.0001

6 Discussion

6.1 Predicting the presence of axillary lymph node metastases (I)

The number of ALN metastases is an important prognostic factor and a crucial part of cancer staging, often leading to decisions in the extent of surgery, systemic therapies, reconstructive methods, and postoperative radiation treatment (Caudle et al. 2014). This study showed that minimal SLN metastases are rarely indicators of multiple ALN metastases, that delayed ALND results in complications on some patients, and that commonly known tumour features play a significant role in metastasising into regional lymph nodes. In this study, one third of positive SLN findings preceded non-SLN metastases in ALND; these patients may have benefitted from complete ALND.

In this study, the decision to complete the surgical treatment with a delayed ALND was made due to ITC or micrometastases in the SLNB, according to the practice of the time. Nowadays, it is known that the minimal SLN metastases do not necessitate ALND and ALND can often be replaced with radiation therapy (Donker et al. 2014; Galimberti et al. 2018). Accordingly, the delayed ALNDs in this study resulted in non-SLN metastases in only 10 % of the cases. This is in line with previous reports: non-SLN metastases are prevalent in 7.2–15 % cases of ITC or micrometastases in SLNB (Cserni et al. 2004; Meretoja et al. 2011). Consequently, most patients with minimal SLN metastases do not need ALND for removing cancerous tissue, and even the possible residual tumour burden can be managed with adjuvant therapies. The failure of locating the SLN is rather rare, only 2 % (Mansel et al. 2006) and may occur if internal mammary lymph nodes are not examined. A meta-analysis has reported the internal mammary SLNB to result in positivity in 15 % of cases, and with six times higher positivity rates when ALN metastases are present (Gong et al. 2019). Therefore, considering the possibility of internal mammary lymph node metastases might need more attention in staging than in the current study and in the current clinical practice in general.

Since the inclusion of the current study population, the relevance of either immediate or delayed ALND in nodal management has declined even for patients with SLN macrometastases (Donker et al. 2019; Giuliano et al. 2017 [b]).

Intraoperative SLN frozen sectioning has been limited in the recent years, and ALNDs are rarely performed in a second separate operation after a positive postoperative SLNB result. Yet, additional lymph node metastases after a positive SLNB should be considered in all patients to distinguish between those with low and high risks for poorer prognosis. If a patient is in high risk for reduced survival, indicated by the prognostic factors, more morbid treatments must be accepted as for patients with delayed ALNDs in this study. ALND does not only serve as removing cancerous tissue or cells but also as the basis of staging and planning further treatments and follow-ups. An alternative for surgical staging of patients without ALND after a positive SLNB is needed when minimizing the invasive care.

Clinopathological features could be used to select patients for radiation treatment after positive SLNB (Morrow 2018). Currently, if a positive SLN is detected but ALND is omitted, radiation treatment for the regional lymph node areas can be suggested followingly: no radiation treatment in low-risk cases, level I–II lymph nodes irradiated in average risk cases, and level I–III lymph nodes together with supraclavicular and internal mammary lymph nodes irradiated in high risk cases. However, the need for radiation treatment and its extent is inconclusive if the SLN presents with ITC or micrometastasis (Cardoso et al. 2019; Finnish Breast Cancer Group 2019). As the therapeutic radiation of the regional lymph nodes has evolved simultaneously with surgical treatment diminishing, the adequate regional treatment in different stages of cancer spread may have been compromised.

A previous meta-analysis has reported that tumour size, lymphovascular invasion, SLN metastasis size, extra-capsular extension, more than one positive SLN, one or no negative SLNs and ratio of positive SLNs > 50 % are the strongest predictive factors for additional ALN metastases after a positive SLNB. In this meta-analysis, tumour biology was noted as ER, PR, HER2 positivity, and grade but biologic tumour subtypes were not specified. (van la Parra et al. 2011). The current study resulted in only one of these factors, as tumour size, biologic tumour subtype, and proliferation index associated with higher numbers of ALN metastases. Nevertheless, it is evident that regional lymph node metastases in breast cancer have a multifactorial origin, and the metastasising can be predicted with established measurable clinopathological factors. Introducing these associations into clinical work require knowledge of statistical thinking, and development of practical multivariable prediction tools.

One clinopathological factor of interest in studies I, II and IV was tumour subtype. Luminal A and B -like tumours were distinguished with a cut-off value of 14 % for Ki-67. Hence most luminal tumours were classified into the luminal B -like subgroup. However, the low and high Ki-67 values in this context should be determined according to laboratory variance, and the Ki-67 cut-off value of 20–30 % might be more appropriate at our institution.

The harms after delayed ALND included some of the commonly described effects: lymphoedema, pain and tightness. Other studies have also reported lowered quality of life and performance skills, slow returning to normal activity, impaired shoulder function and neurologic problems after ALND (Lyman et al. 2016; Mansel et al. 2006; Mejdahl et al. 2013; Rao et al. 2013). ALND also results in more adverse effects than axillary radiation treatment (Donker et al. 2014). Therefore, in the light of long-term patient satisfaction, omitting unnecessary delayed ALNDs is an advantageous goal as most delayed ALNDs in this study did not result in finding additional ALN metastases. However, it has become challenging to recommend treatments for individual patients or patients even with the same cancer stage: new treatments can offer a less morbid alternative for previous practices, but the overall benefits and risks should be carefully weighed (Burstein et al. 2019).

6.2 Predicting stage pN2−3 in clinically node negative patients (II)

In this study, nodal stage pN2a or pN3a was estimated in clinically node negative patients by using a new designated nomogram, developed at the University of Turku and Turku University Hospital. The machine-learning model was compared with logistic regression analysis, and both performed well in discriminating patients with high and low nodal stage. The AUC values of the machine-learning model were well above the recommended 0.75 for clinical prediction tools (Balachandran et al. 2015). The machine-learning model outperformed the logistic regression model in maintaining good performance in a separate validation cohort.

Prediction of stage pN2–3 is clinically relevant in the current practice where ALND is only recommended for those with a suspicion of stage N2–3 after a positive SLNB, and some other subgroups of patients (Finnish Breast Cancer Group 2019). According to this recommendation, 78 patients in this study with confirmed nodal stage pN2a or pN3a may not have undergone ALND in the current clinical practice. Detecting the high nodal stage is important in order to select all patients with four or more ALN metastases for extensive postoperative regional nodal irradiation (Burstein et al. 2019). To be exact, the irradiated lymph node levels are determined according to the overall risk, including the number of ALN metastases (Cardoso et al. 2019). In the current study, 17 % of clinically node negative patients with SLN positivity had four or more ALN metastases. According to the current knowledge, these patients should be selected for complete ALND, and for more extensive regional irradiation than what is offered in nodal stage pN0–1. The selection of these patients may be compromised if the omission of ALND or its replacement with radiation treatment is liberally applied in other patient groups than those meeting the criteria of recent trials behind clinical guidelines (Donker et al. 2019; Giuliano et al.

2017 [b]). The St. Gallen's international consensus guidelines are in favour of expanding the eligible patient groups in order to benefit from these trial results more widely (Burstein at al. 2019).

Nomograms may be helpful in the transitional period before complementary clinical trials are available for general, heterogenous patient populations (Tapia et al. 2019). To succeed in differentiating between stage pN0–1 and pN2–3, the nomogram variables chosen by the machine-learning algorithm in this study included some familiar predictive factors: tumour size, tumour histology, tumour multifocality, lymphovascular invasion, ER status and the number of positive SLNs. In addition, the number of SLN metastases multiplied by tumour size was inserted as a tailored variable. In the future, this nomogram could be refined into a clinical computational risk calculator, in which the presented tumour factors can easily be placed. The logistic regression model can be used as a mathematical equation reported in the results section of this study although the interpretation of its results oblige familiarization with the subject.

The practicality and comprehensibility of the nodal stage calculator are essential for its clinical use. The calculator results could benefit in determining which patients need ALND to be cured, and which could benefit from certain adjuvant treatments. The nomogram estimations should, however, be used as a complimentary tool together with treatment recommendations based on international randomized clinical trials (Balachandran et al. 2015). Future research may determine whether statistically predicted nodal stage is in fact as prognostically accurate as the traditionally determined nodal stage. To accomplish this, randomized controlled clinical trials must be launched to clarify nomograms' clinical utility. Today, cancer care relies on follow-up trials, but time will show whether the ever-evolving artificial intelligence can in fact replace some of the costly and time-consuming research methods.

Nomograms predicting the nodal stage have had challenges in external validity in heterogenous populations. The modern artificial intelligence and self-learning features of machine-learning could help in developing more generally applicable models for global use. However, machine-learning models may produce promising results in small populations, yet similarly to other nomograms still need validation in larger datasets (Fusco et al. 2016). In this study, the XGBoost model outperformed the logistic regression model in balancing the small sample size of the validation cohort by seemingly learning more over-all characteristics from the data. In fact, machine-learning performs best in very large datasets, where it can discover even unexpected interactions between descriptive characteristics.

6.3 Validation of former nomograms to predict the nodal stage pN2–3 (III)

In addition to developing our own nomogram in study II, in study III it was legitimate to evaluate some of the other nomograms developed globally to predict the ALN status. In this study, three previous prediction models were validated in a current patient population. The patients included in our study were treated for locoregional breast cancer for up to 13–19 years later than the patients in the original nomogram studies. Therefore, it can be interpreted that the nomograms developed many years ago are still valid among the 2010s patients.

The commonly validated prediction models are rather old having been developed in the early twenty-first century. Their cohorts are mainly formed from patients diagnosed in the 1990s (Chen and Wu 2011; Hwang et al. 2003; Van Zee et al. 2003). Balachandran et al. (2015) noted that "a nomogram can become less accurate with time for a variety of reasons, such as improvements in therapy, earlier detection, and changes in natural history" although they presented no studies supporting this notion. While these possible changes were not investigated in the current study, it was clear that variables used in the validated nomograms reflected the very established predictors in breast cancer, thus probably not easily influenced by minor alterations. In a previous study by Ngô et al. (2012), it was also shown that a nomogram developed in patients from 2000–2007, performed similarly on patients treated in 2009.

Although a meta-analysis of nomogram validation has indicated that the external performance of prediction tools for non-SNL metastases is not adequate (Zhu et al. 2013), more recent studies have discussed that nomograms are still relevant in the era after ACOSOG Z0011 and other notable trials (Chen et al. 2017; Tapia et al. 2019). The current study is not entirely in line with the publications showing weaker nomogram discrimination in external populations than in original training populations. The AUC value of the Chagpar model in the current study did not meet with the high AUC values of the original study. However, the AUC values of Katz and Meretoja models were in fact higher in our cohort than in the original studies. This may be due to cohort homogeneity in our single institution study, favouring good statistical model performance, compared to the original multi-institutional patient populations. According to recommendations the sample size in the current validation study came close to adequacy with 87 events and over 100 non-events at least for the Katz model (Vergouwe et al. 2005). Still, a larger patient cohort may have improved the reliability of the results.

Validating nomograms from other institutions globally is a way to evaluate the local population in relation to international populations. Hence validation studies could support the safe applying of international guidelines and study results to local populations. They also provide alternative and additional tools to be used by

multidisciplinary breast cancer groups, promoting safe and individualized breast cancer care. Compared to the nomogram developed at our own institution, the AUC values for the three presented nomograms performed surprisingly well. The machine-learning nomogram developed on our own patient population had the AUC value of 0.80 (95 % CI 0.65 to 0.92) in a separate validation cohort, while the AUC values for the Chagpar, Katz and Meretoja models were 0.79 (95% CI 0.74 to 0.83), 0.87 (95 % CI 0.83 to 0.91) and 0.82 (95% CI 0.76 to 0.86), respectively. The AUC values are not fully comparable due to different sample sizes.

This study confirms the status of the former twenty-first century nomograms in the prediction of nodal stage pN2a or pN3a, supporting their reliability in today's context. On the other hand, the extend of clinical nomogram use remains unclear.

6.4 Preoperative imaging in the context of staging (IV)

This study focused on evaluating which patients are referred to MRI for additional preoperative staging, how tumour size is presented in different imaging methods and which effects preoperative MRI may have on treatment timing and the extent of primary surgery at our institution.

In comparison to standard imaging, preoperative MRI was more frequently conducted for patients with younger age, mastectomy rather than breast-conserving surgery as the primary surgery, invasive lobular tumours rather than invasive ductal tumours, and HER2 positive tumours. These results are in line with another study investigating factors that associated with more frequent staging by MRI or thomosynthesis (Mariscotti et al. 2019). The results indicate that the clinical guidelines suggesting preoperative MRI for lobular tumours and for patients with neoadjuvant systemic treatment in HER2 positive cases have become well accustomed in the clinical practice at our institution (Cardoso et al. 2019). Younger age of patients with MRI might partly reflect hereditary breast cancer cases or dense breast tissue. Additionally, large, and multifocal tumours were associated with more frequent use of MRI in the univariate setting. Hence, the results of this study are an example of quality control at our institution, showing good concordance between guidelines and clinical practice.

Tumour size was estimated by mammography, ultrasound, and MRI rather well. These methods can thus be safely used in preoperative tumour staging. In harmony with previous studies, MRI estimated the varying tumour size most accurately among preoperative imaging investigations yet overestimating the tumour size (Sardanelli et al. 2010; Katz et al. 2017; Haraldsdóttir et al. 2017). Smaller tumours were detected by the screening mammography more frequently than by palpation or other means. Furthermore, patients diagnosed by screening mammography were

younger than patients diagnosed after finding a palpable tumour. This is naturally one of the goals of the screening program: to detect cancer in early stage which does not yet present with clinical symptoms. Extending the screening into covering older age groups, some of the palpable tumours could be detected at an earlier stage. Tumour size in the MRI reports of evaluated medical records were often described with the entire abnormal area of the breast, resulting in large dimensions. Yet, tumour size in the final PAD was calculated from the invasive component, and surrounding DCIS was not included in the current study. This can be the explanation for MRI overestimates.

Preoperative MRI caused reasonable delays in primary breast cancer operation and the final pathological anatomic diagnosis, compared to some other studies investigating the MRI waiting time (Bleicher et al. 2009; Zhang et al. 2017). The routine preoperative MRI with certain indications before a multidisciplinary meeting as a part of the diagnostics could shorten this delay. Anyhow, if reoperations are prevented by MRI, a short delay in surgery will be of a lesser disadvantage for patients. Previous research has suggested that preoperative MRI may change treatment plans, but the issue is complex (Karlsson et al. 2019; Lehman et al. 2019). Some reports have stated that preoperative MRI does not reduce reoperations (Houssami et al. 2017; Turnbull et al. 2010), but also conflictive conclusions have been presented (Lai et al. 2016). In another, prospective study at Turku University Hospital, there was no statistical difference between breast reoperation rates with or without preoperative MRI (Brück et al. 2018). Unfortunately, the need to perform mastectomies instead of breast-conserving surgeries after MRI could not be definitely described in the current study. Patient wishes or surgeon preferences were unknown concerning more radical surgical methods. The need for breast reoperations were also not noted in this study.

Plichta et al. (2018) summarized that there is no definite evidence that breast MRI could result in less positive margins or improved outcomes of breast cancer patients. The need to surgically remove all cancer foci detected by MRI is controversial as irradiation after breast-conserving surgery might suffice as a treatment for minor residual tumour foci (Brennan et al. 2009). If these statements are true, it is inevitable that patients selected for preoperative MRI must be carefully described by clinopathological characterization. On the other hand, clinically node negative patients are those with the highest potential for curative outcome, which should not be degraded by inadequate tumour staging. It seems that the preoperative staging by mammography and ultrasound is enough for many clinically node negative patients, yet a considerable number of patients may need additional MRI for improved accuracy.

SLNB is a more frequent staging method for the axilla after MRI than after standard imaging (Lai et al. 2016). The current study showed that after positive

SLNB, ALN reoperation rate was similar between patients with and without preoperative MRI. In this study, patients with more apparent nodal metastases possibly visualized by MRI or ultrasound were excluded as they did not undergo SLNB. However, this study shows that there are many breast cancer patients with clinically negative nodes, yet metastases confirmed in SLNB. Due to patient selection, the study population displays the patients whose nodal metastases were undetected by physical examination, ultrasound examination, fine-needle aspiration, or MRI. That said, preoperative staging of the axilla by imaging and other investigations may have an increasing importance as the surgical treatment diminishes. It is worrying that at the same time as surgical treatment is in transition, some international institutions have omitted preoperative axillary ultrasound. Their reasoning for this is to prevent overtreatment since most patients with palpation negative ALNs do not need ALND. (Chang et al. 2020; Wallis et al. 2018). Hence the routine preoperative axillary ultrasound examination has been questioned (Morrow 2018; Plichta et al. 2018). Concerning neoadjuvant systemic therapies and the MRI evaluation of treatment response, it has been suggested that controlling the ALN status with MRI after few treatment cycles may predict cancer survival (Myller et al. 2020). According to the current study, MRI may however not be a reliable imaging method for the nodal stage.

Finally, it can be discussed that the evolving knowledge on early and locally advanced breast cancer has resulted in advancing outcomes with lesser treatment-related morbidity. On the other hand, the decreasing use of invasive methods has led to the increasing importance of non-invasive methods such as preoperative imaging and statistics to stage breast cancer, together with the knowledge of tumour biology in pathological samples. The core and the historical basis of cancer care, classifying and staging the cancer, should not be forgotten as modern achievements of breast cancer research are introduced to clinical work. International treatment guidelines are based on randomized controlled clinical trials and large amounts of evidence supporting specific recommendations. To overcome the lack of staging information from omitted surgical procedures, guidelines rely on outcome results. However, the missing knowledge on the accurate stage may be more complex than just measuring survival or recurrence rates. If anatomical locoregional staging has guided the treatment planning since the 1950s, a few years or a decade of follow-ups for a certain trial population may not be enough to abandon accurate locoregional staging. Multidisciplinary breast cancer experts are encountering a demanding task to navigate through these changing times towards more and more individualized cancer care.

6.5 Study limitations

This study has some limitations. Some of the sample sizes were small even though the best obtainable cohorts at the time were used. Concerning machine-learning, there is no straightforward way for sample size estimation. Sample sizes influenced for example the calibrations of the nomograms. To support the calibrations in this study, for example false negative rates could have been calculated as by Tapia et al. (2019). The delayed ALNDs were infrequent in study I, although they were the subject of special interest.

The results of this study may not be completely applicable in today's context due to change in clinical practice. The long time period of patient inclusions from 2009 to 2017 can benefit in more generally applicable results, but also cause some bias due to the changes. This is a single institution study which may be difficult to repeat elsewhere due to some local practices. During the years, the hospital clinics and individual professionals may have had varying practices in recording information relevant to this study. Furthermore, multifocality was defined as the combined diameter of all foci although the TNM classification recommends using the diameter of the largest foci of multifocal tumours for staging (WHO Classification of Tumours Editorial Board 2019). The actual T stage which was determined during cancer diagnostics was not verified during data collection. Follow-up information on breast reoperations was not collected to evaluate the true effect of MRI on surgical treatment. In study III, ECE was not a reliable variable in this cohort. Our institution has not had a routine reporting scheme for ECE.

This study did not include patients with given neoadjuvant systemic treatments, nor can discuss the results in relations to neoadjuvant therapies. Yet, neoadjuvant systemic treatment has been an important factor in diminishing ALNDs. In general, the results of this study are only appropriate for clinically node negative patients.

7 Conclusions

1. Minimal metastases in sentinel lymph nodes are rarely an indication of non-sentinel lymph node metastases. Delayed axillary lymph node dissections are unnecessary for most patients with isolated tumour cells or micrometastases in sentinel lymph node biopsy.
2. Clinopathological factors can be used to predict increasing numbers of lymph node metastases and nodal stage. Machine-learning techniques are beneficial when developing mathematical tools for nodal stage prediction.
3. Through the validation process, nomograms from other institutions and decades can be confirmed for prediction of the nodal stage, and for the clinical use to help the stage-related decision on appropriate surgical and adjuvant treatments.
4. For selected patients such as those with young age, large tumour size, invasive lobular or HER2 positive tumours, preoperative magnetic resonance imaging is a noteworthy addition to the staging investigations to guarantee a tumour-free outcome in the first-line breast cancer surgery. However, magnetic resonance imaging does not seem to decrease axillary reoperations after a positive sentinel lymph node biopsy.

Acknowledgements

This work was carried out in the Department of Clinical Oncology, University of Turku, and Turku University Hospital during January 2017 – April 2021. I would like to express my sincere gratitude for Professor Heikki Minn for the opportunity to conduct this research in the Department of Clinical Oncology.

My special thanks belongs to my supervisors Research Professor, Adjunct Professor Eeva Salminen, and MD, Ph.D Pia Boström, who have kindly offered their expertise and support throughout the research process. They have enabled the completion of this thesis by their continuously valuable advice and input.

I am overly thankful for the follow-up committee members Ph.D Antti Karlsson for his irreplaceable expertise on statistical points of view, and Docent Pauliina Kronqvist for her precious views on the study. I am also highly grateful for my co-authors, biostatistician Tero Vahlberg, and surgeon Riitta Aaltonen who have kindly contributed to this research with their knowledge and time.

I wish to acknowledge the important advice and comments provided by all the reviewers of the original publications, and especially the valuable improvements suggested by Docent Päivi Auvinen from Kuopio University Hospital and University of Eastern Finland, and Docent Teijo Kuopio from University of Jyväskylä and Central Finland Health Care District.

I would like to thank all colleagues and co-workers at Turku University Hospital or elsewhere, who have participated in my daily and lifelong learning path. Practical assistance provided by Tiina Jonkka at the Department of Oncology was greatly appreciated.

My warmest gratitude goes to my parents Tuija and Kimmo for their love and support on my path to become a medical professional and an academic. I am thankful for my sisters Sohvi and Roosa, and my closest friends for being there to lighten up any troublesome moments.

I am forever grateful for Jonni for his love and companionship that has enabled me to pursue this direction in life. And our precious daughter Silja who has provided me with never-ending smiles even on the hardest days of this journey.

This research was made possible by the grants from Cancer Society of South-West Finland, the Finnish Medical Foundation, Turku University Foundation, and

the state research funding of the Turku University Hospital expert responsibility area. I would like to thank Auria Clinical Informatics for the assistance with the collection of my data.

Turku, April 2021

Vilma Madekivi

References

Ahn SK, Kim MK, Kim J, Lee E, Yoo TK, Lee HB, Kang YJ, Kim J, Moon HG, Chang JM, Cho N, Moon WK, Park IA, Noh DY, Han W. Can we skip intraoperative evaluation of sentinel lymph nodes? Nomogram predicting involvement of three or more axillary lymph nodes before breast cancer surgery. *Cancer Res Treat*. 2017; 49:1088–1096.

Ai X, Wang M, Li J, Hu Y, Hou L, Zheng X, Yan Y, Pan Q, Jin Y, Liu W, Tan X, Tian Y, Zhang Y, Tang P, Jiang J. Supraclavicular lymph node dissection with radiotherapy versus radiotherapy alone for operable breast cancer with synchronous ipsilateral supraclavicular lymph node metastases: a real-world cohort study. *Gland Surg*. 2020; 9:329–341.

Alabousi M, Zha N, Salameh JP, Samoilov L, Sharifabadi AD, Pozdnyakov A, Sadeghirad B, Freitas V, McInnes MDF, Alabousi A. Digital breast tomosynthesis for breast cancer detection: a diagnostic test accuracy systematic review and meta-analysis. *Eur Radiol*. 2020; 30:2058–2071.

Allison KH, Hammond MEH, Dowsett M, McKernin SE, Carey LA, Fitzgibbons PL, Hayes DF, Lakhani SR, Chavez-MacGregor M, Perlmutter J, Perou CM, Regan MM, Rimm DL, Symmans WF, Torlakovic EE, Varella L, Viale G, Weisberg TF, McShane LM, Wolff AC. Estrogen and Progesterone Receptor Testing in Breast Cancer: ASCO/CAP Guideline Update. *J Clin Oncol*. 2020; 38:1346–1366.

Anders CK, Hsu DS, Broadwater G, Acharya CR, Foekens JA, Zhang Y, Wang Y, Marcom PK, Marks JR, Febbo PG, Nevins JR, Potti A, Blackwell KL. Young age at diagnosis correlates with worse prognosis and defines a subset of breast cancers with shared patterns of gene expression. *J Clin Oncol*. 2008; 26:3324–3330.

Alpaydin, Ethem. *Introduction to Machine Learning*. Fourth edition. Cambridge, Massachusetts: The MIT Press, 2020. Print.

Aromaa-Häyhä A, Auvinen P, Sarkeala T, Heinävaara S, Lamminmäki A, Malila N, Kataja V. Improved diagnostics and change of tumour characteristics in breast cancer: a retrospective study over two decades. *Acta Oncol.* 2018; 57:1331–1338.

Association of Breast Surgery at Baso 2009. Surgical guidelines for the management of breast cancer. *Eur J Surg Oncol.* 2009;35 Suppl 1:1–22.

Ayala de la Peña F, Andrés R, Garcia-Sáenz JA, Manso L, Margelí M, Dalmau E, Pernas S, Prat A, Servitja S, Ciruelos E. SEOM clinical guidelines in early stage breast cancer (2018). *Clin Transl Oncol.* 2019; 21:18–30.

Bakker MF, de Lange SV, Pijnappel RM, Mann RM, Peeters PHM, Monninkhof EM, Emaus MJ, Loo CE, Bisschops RHC, Lobbes MBI, de Jong MDF, Duvivier KM, Veltman J, Karssemeijer N, de Koning HJ, van Diest PJ, Mali WPTM, van den Bosch MAAJ, Veldhuis WB, van Gils CH; DENSE Trial Study Group. Supplemental MRI screening for women with extremely dense breast tissue. *N Engl J Med.* 2019; 381:2091–2102.

Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: More than meets the eye. *The Lancet Oncology*. 2015;16:e17-e180.

Balasubramanian I, Fleming CA, Corrigan MA, Redmond HP, Kerin MJ, Lowery AJ. Meta-analysis of the diagnostic accuracy of ultrasound-guided fine-needle aspiration and core needle biopsy in diagnosing axillary lymph node metastasis. *Br J Surg*. 2018; 105:1244–1253.

Basu P, Ponti A, Anttila A, Ronco G, Senore C, Vale DB, Segnan N, Tomatis M, Soerjomataram I, Primic Žakelj M, Dillner J, Elfström KM, Lönnberg S, Sankaranarayanan R. Status of implementation and organization of cancer screening in The European Union Member States-Summary results from the second European screening report. *Int J Cancer*. 2018; 142:44–56.

Bazemore AW, Smucker DR. Lymphadenopathy and malignancy. *Am Fam Physician*. 2002; 66:2103–2110.

Berg WA, Gutierrez L, NessAiver MS, Carter WB, Bhargavan M, Lewis RS, Ioffe OB. Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer. *Radiology*. 2004; 233:830–849.

Bevilacqua JL, Kattan MW, Fey JV, Cody HS 3rd, Borgen PI, Van Zee KJ. Doctor, what are my chances of having a positive sentinel node? A validated nomogram for risk estimation. *J Clin Oncol.* 2007; 25:3670–3679.

Biganzoli L, Cardoso F, Beishon M, Cameron D, Cataliotti L, Coles CE, Delgado Bolton RC, Trill MD, Erdem S, Fjell M, Geiss R, Goossens M, Kuhl C, Marotti L, Naredi P, Oberst S, Palussière J, Ponti A, Rosselli Del Turco M, Rubio IT, Sapino A, Senkus-Konefka E, Skelin M, Sousa B, Saarto T, Costa A, Poortmans P. The requirements of a specialist breast centre. *Breast*. 2020; 51:65–84.

Bleicher RJ, Ciocca RM, Egleston BL, Sesa L, Evers K, Sigurdson ER, Morrow M. Assocation of routine pretreatment MRI with time to surgery, mastectomy rate, and margin status. *J Am Coll Surg.* 2009; 209:180–295.

Bonsang-Kitzis H, Mouttet-Boizat D, Guillot E, Feron JG, Fourchotte V, Alran S, Pierga JY, Cottu P, Lerebours F, Stevens D, Vincent-Salomon A, Sigal-Zafrani B, Campana F, Rouzier R, Reyal F. Medico-economic impact of MSKCC non-sentinel node prediction nomogram for ER-positive HER2-negative breast cancers. *PLoS One*. 2017;12:e0169962.

Bosch AM, Kessels AG, Beets GL, Rupa JD, Koster D, van Engelshoven JM, von Meyenfeldt MF. Preoperative estimation of the pathological breast tumour size by physical examination, mammography and ultrasound: a prospective study on 105 invasive tumours. *Eur J Radiol*. 2003; 48:285–292.

Boutros C, Mazouni C, Lerebours F, Stevens D, Lei X, Gonzalez-Angulo AM, Delaloge S. A preoperative nomogram to predict the risk of synchronous distant metastases at diagnosis of primary breast cancer. *Br J Cancer*. 2015; 112:992–997.

Brennan M, Spillane A, and Houssami N. The role of breast MRI in clinical practice. *Aust Fam Physician.* 2009; 38:513–519.

Brennan ME, Turner RM, Ciatto S, Marinovich ML, French JR, Macaskill P, Houssami N. Ductal carcinoma in situ at core-needle biopsy: meta-analysis of underestimation and predictors of invasive breast cancer. *Radiology*. 2011; 260:119–128.

Bromham N, Schmidt-Hansen M, Astin M, Hasler E, Reed MW. Axillary treatment for operable primary breast cancer. *Cochrane Database of Systematic Reviews.* 2017. Issue 1. Art. No.: CD004561.

Brück N, Koskivuo I, Boström P, Saunavaara J, Aaltonen R, Parkkola R. Preoperative Magnetic Resonance Imaging in Patients With Stage I Invasive Ductal Breast Cancer: A Prospective Randomized Study. *Scand J Surg*. 2018; 107:14–22.

Buchholz TA, Somerfield MR, Griggs JJ, El-Eid S, Hammond MEH, Lyman GH, Mason G, and Newman LA. Margins for Breast-Conserving Surgery With Whole-Breast Irradiation in Stage I and II Invasive Breast Cancer: American Society of Clinical Oncology Endorsement of the Society of Surgical Oncology/American Society for Radiation Oncology Consensus Guideline. *Journal of Clinical Oncology*. 2014; 14:1502–1506

Burstein HJ, Curigliano G, Loibl S, Dubsky P, Gnant M, Poortmans P, Colleoni M, Denkert C, Piccart-Gebhart M, Regan M, Senn HJ, Winer EP, Thurlimann B; Members of the St. Gallen International Consensus Panel on the Primary Therapy of Early Breast Cancer 2019. Estimating the benefits of therapy for early-stage breast cancer: the St. Gallen International Consensus Guidelines for the primary therapy of early breast cancer 2019. *Ann Oncol*. 2019; 30:1541–1557.

Cady B, Stone MD, Schuler JG, Thakur R, Wanner MA, Lavin PT. The New Era in Breast Cancer: Invasion, Size, and Nodal Involvement Dramatically Decreasing as a Result of Mammographic Screening. *Arch Surg*. 1996; 131:301–308.

Cardoso F, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rubio IT, Zackrisson S, Senkus E, on behalf of the ESMO Guidelines Committee. Early breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2019; 30:1194–1220.

Cardoso F, Spence D, Mertz S, Corneliussen-James D, Sabelko K, Gralow J, Cardoso MJ, Peccatori F, Paonessa D, Benares A, Sakurai N, Beishon M, Barker SJ, Mayer M. Global analysis of advanced/metastatic breast cancer: Decade report (2005-2015). *Breast*. 2018; 39:131–138.

Castelo M, Hu SY, Dossa F, Acuna SA, Scheer AS. Comparing Observation, Axillary Radiotherapy, and Completion Axillary Lymph Node Dissection for Management of Axilla in Breast Cancer in Patients with Positive Sentinel Nodes: A Systematic Review. *Ann Surg Oncol*. 2020; 27:2664–2676.

Caudle AS, Cupp JA, Kuerer HM. Management of axillary disease. *Surg Oncol Clin N Am.* 2014; 23:473–486.

Causer PA, Jong RA, Warner E, Hill K, Wong JW, Curpen BN, Plewes DB. Breast cancers detected with imaging screening in the BRCA population: emphasis on MR imaging with histopathologic correlation. *Radiographics*. 2007; 27:S165–82.

Chang JM, Leung JWT, Moy L, Ha SM, Moon WK. Axillary Nodal Evaluation in Breast Cancer: State of the Art. *Radiology*. 2020; 295:500–515.

Chen JJ, Wu J. Management strategy of early-stage breast cancer patients with a positive sentinel lymph node: With or without axillary lymph node dissection. *Crit Rev Oncol Hematol*. 2011; 79:293–301.

Chen K, Jin L, Zhu L, Shan Q, Su F. Which nomograms may be the best for predicting nonsentinel lymph node metastasis in breast cancer patients: A meta-analysis. *Cancer Res*. 2012;72:P-09.

Chen K, Liu J, Li S, Jacobs L. Development of nomograms to predict axillary lymph node status in breast cancer patients. *BMC Cancer*. 2017; 17:561

Chen X, Yuan Y, Gu Z, Shen K. Accuracy of estrogen receptor, progesterone receptor, and HER2 status between core needle and open excision biopsy in breast cancer: a meta-analysis. *Breast Cancer Res Treat.* 2012; 134:957–967.

Cho N, Moon WK, Han W, Park IA, Cho J, Noh DY. Preoperative sonographic classification of axillary lymph nodes in patients with breast cancer: node-to-node correlation with surgical histology and sentinel node biopsy results. *AJR Am J Roentgenol*. 2009; 193:1731–1737.

Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019; 110:12–22.

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 Group (EBCTCG). 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.

Coburn NG, Chung MA, Fulton J, Cady B. Decreased breast cancer tumor size, stage, and mortality in Rhode Island: an example of a well-screened population. *Cancer Control*. 2004; 11:222–230.

Collaborative Group on Hormonal Factors in Breast Cancer. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol*. 2012; 13:1141–1151.

Cserni G, Bori R, Maráz R, Leidenius MH, Meretoja TJ, Heikkila PS, Regitnig P, Luschin-Ebengreuth G, Zgajnar J, Perhavec A, Gazic B, Lázár G, Takács T, Vörös A, Audisio RA. Multi-institutional comparison of non-sentinel lymph node predictive tools in breast cancer patients with high predicted risk of further axillary metastasis. *Pathol Oncol Res*. 2013; 19:95–101.

Cserni G, Chmielik E, Cserni B, Tot T. The new TNM-based staging of breast cancer. *Virchows Arch*. 2018; 472:697–703.

Cserni G, Gregori D, Merletti F, Sapino A, Mano M, Ponti A, Sandrucci S, Baltás B, Bussolati G. Meta-analysis of non-sentinel node metastases associated with micrometastatic sentinel nodes in breast cancer. *Br J Surg*. 2004; 91:1245–1252.

Cvetković J. Breast Cancer Patients' Depression Prediction by Machine Learning Approach. *Cancer Invest*. 2017; 35:569–572.

Danckert B, Ferlay J, Engholm G , Hansen HL, Johannesen TB, Khan S, Køtlum JE, Ólafsdóttir E, Schmidt LKH, Virtanen A and Storm HH. NORDCAN: Cancer Incidence, Mortality, Prevalence and Survival in the Nordic Countries, Version 8.2 (26.03.2019). Association of the Nordic Cancer Registries. Danish Cancer Society. Available from: <http://www.ancr.nu> [Accessed 28.12.2020]

Dafni U, Tsourti Z, Alatsathianos I. Breast Cancer Statistics in the European Union: Incidence and Survival across European Countries. *Breast Care (Basel).* 2019; 14:344-353.

Davidson A, Chia S, Olson R, Nichol A, Speers C, Coldman AJ, Bajdik C, Woods R, Tyldesley S. Stage, treatment and outcomes for patients with breast cancer in British Columbia in 2002: a population-based cohort study. *CMAJ Open*. 2013; 1:E134–141.

De Lorenzi F, Hubner G, Rotmensz N, Bagnardi V, Loschi P, Maisonneuve P, Venturino M, Orecchia R, Galimberti V, Veronesi P, Rietjens M. Oncological results of oncoplastic breast-conserving surgery: Long term follow-up of a large series at a single institution: A matched-cohort analysis. *Eur J Surg Oncol*. 2016; 42:71–77.

Desai AA, Hoskin TL, Day CN, Habermann EB, Boughey JC. Effect of primary breast tumor location on axillary nodal positivity. *Ann Surg Oncol*. 2018; 25:3011-3018.

Diepstraten SC, Sever AR, Buckens CF, Veldhuis WB, van Dalen T, van den Bosch MA, Mali WP, Verkooijen HM. Value of preoperative ultrasound-guided axillary lymph node biopsy for preventing completion axillary lymph node dissection in breast cancer: a systematic review and meta-analysis. *Ann Surg Oncol*. 2014; 21:51–59.

Ding X, Xie S, Chen J, Mo W, Yang H. A support vector machine model for predicting non-sentinel lymph node status in patients with sentinel lymph node positive breast cancer. *Tumor Biol*. 2013; 34:1547–1552.

Donker M, van Tienhoven G, Straver ME, Meijnen P, van de Velde CJ, Mansel RE, Cataliotti L, Westenberg AH, Klinkenbijl JH, Orzalesi L, Bouma WH, van der Mijle HC, Nieuwenhuijzen GA, Veltkamp SC, Slaets L, Duez NJ, de Graaf PW, van Dalen T, Marinelli A, Rijna H, Snoj M, Bundred NJ, Merkus JW, Belkacemi Y, Petignat P, Schinagl DA, Coens C, Messina CG, Bogaerts J, Rutgers EJ. Radiotherapy or surgery of the axilla after a positive sentinel node in breast cancer (EORTC 10981-22023 AMAROS): a randomised, multicentre, open-label, phase 3 non-inferiority trial. *Lancet Oncol.* 2014; 15:1303–1310.

Dyrstad SW, Yan Y, Fowler AM, Colditz GA. Breast cancer risk associated with benign breast disease: systematic review and meta-analysis. *Breast Cancer Res Treat*. 2015; 149:569–575.

Elias SG, Adams A, Wisner DJ, Esserman LJ, van't Veer LJ, Mali WP, Gilhuijs KG, Hylton NM. Imaging features of HER2 overexpression in breast cancer: a systematic review and meta-analysis. *Cancer Epidemiol Biomarkers Prev*. 2014; 23:1464–1483.

Ellis O, Al-Sam S, Anderson N, Carder P, Deb R, Girling A, Hales S, Hanby A, Ibrahim M, Lee AHS, Liebmann R, Mallon E, Pinder SE, Provenzano E, Quinn C, Rakha E, Rowlands D, Stephenson T, Wells CA. Pathology reporting of breast disease in surgical excision specimens incorporating the dataset for histological reporting of breast cancer. Published by the Royal College of Pathologists. 2016. Available from <https://www.rcpath.org/uploads/assets/693db661-0592-4d7e-9644357fbfa00a76/G148_BreastDataset-lowres-Jun16.pdf> [Accessed 24.11.2020]

Ebner F, Wöckel A, Schwentner L, Blettner M, Janni W, Kreienberg R, Wischnewsky M. Does the number of removed axillary lymphnodes in high risk breast cancer patients influence the survival? *BMC Cancer*. 2019; 19: 90.

Ecanow JS, Abe H, Newstead GM, Ecanow DB, Jeske JM. Axillary staging of breast Cancer: what the radiologist should know. *RadioGraphics*. 2013 33:6, 1589-1612.

Field AS, Raymond WA, Rickard M, Arnold L, Brachtel EF, Chaiwun B, Chen L, Di Bonito L, Kurtycz DFI, Lee AHS, Lim E, Ljung BM, Michelow P, Osamura RY, Pinamonti M, Sauer T, Segara D, Tse G, Vielh P, Chong PY, Schmitt F. The International Academy of Cytology Yokohama System for Reporting Breast Fine-Needle Aspiration Biopsy Cytopathology. *Acta Cytol*. 2019; 63: 257–273.

Finnish Breast Cancer Group. *Rintasyövän valtakunnallinen diagnostiikka- ja hoitosuositus*. FICAN South, HYKS, HUS. 2019. Available from: https://rintasyoparyhma.yhdistysavain.fi/hoitosuositus/ [Accessed 30.12.2020]

Finnish Breast Cancer Group. *Rintasyövän valtakunnallinen diagnostiikka- ja hoitosuositus.* 2013.

Finnish Cancer Registry. Available from: <https://cancerregistry.fi/statistics/cancer-statistics/> [Accessed 9.11.2020] [a]

Finnish Cancer Registry. Available from: <https://cancerregistry.fi/screening/breast-cancer-screening/> [Accessed 11.12.2020] [b]

Fu MR, Wang Y, Li C, Qiu Z, Axelrod D, Guth AA, Scagliola J, Conley Y, Aouizerat BE, Qiu JM, Yu G, Van Cleave JH, Haber J, Cheung YK. Machine learning for detection of lymphedema among breast cancer survivors. *Mhealth*. 2018; 4:17.

Fusco R, Sansone M, Filice S, Carone G, Amato DM, Sansone C, Petrillo A. Pattern Recognition Approaches for Breast Cancer DCE-MRI Classification: A Systematic Review. *J Med Biol Eng*. 2016; 36:449–459.

Galimberti V, Cole BF, Viale G, Veronesi P, Vicini E, Intra M, Mazzarol G, Massarut S, Zgajnar J, Taffurelli M, Littlejohn D, Knauer M, Tondini C, Di Leo A, Colleoni M, Regan MM, Coates AS, Gelber RD, Goldhirsch A; International Breast Cancer Study Group Trial 23-01. Axillary dissection versus no axillary dissection in patients with breast cancer and sentinel-node micrometastases (IBCSG 23-01): 10-year follow-up of a randomised, controlled phase 3 trial. *The Lancet Oncology.* 2018; 19:1385–1393.

Garg PK, Prakash G. Current definition of locally advanced breast cancer. *Curr Oncol*. 2015; 22:e409–410.

Gentilini O, Botteri E, Dadda P, Sangalli C, Boccardo C, Peradze N, Ghisini R, Galimberti V, Veronesi P, Luini A, Cassano E, Viale G, Veronesi U. Physical function of the upper limb after breast cancer surgery. Results from the SOUND (Sentinel node vs. Observation after axillary Ultra-souND) trial. *Eur J Surg Oncol*. 2016; 42:685–689.

Gentilini O, Veronesi U. Abandoning sentinel lymph node biopsy in early breast cancer? A new trial in progress at the european institute of oncology of milan (SOUND: Sentinel node vs observation after axillary UltraSouND). *The Breast*. 2012; 21:678–681.

Gipponi M, Bassetti C, Canavese G, Catturich A, Di Somma C, Vecchio C, Nicolò G, Schenone F, Tomei D, Cafiero F. Sentinel lymph node as a new marker for therapeutic planning in breast cancer patients. *J Surg Oncol.* 2004; 85:102–111.

Giuliano AE, Ballman KV, McCall L, Beitsch PD, Brennan MB, Kelemen PR, Ollila DW, Hansen NM, Whitworth PW, Blumencranz PW, Leitch AM, Saha S, Hunt KK, Morrow M. Effect of axillary dissection vs no axillary dissection on 10-year overall survival among women with invasive breast cancer and sentinel node metastasis: The ACOSOG Z0011 (alliance) randomized clinical trial. *JAMA.* 2017; 318:918–926. [b],

Giuliano AE, Connolly JL, Edge SB, Mittendorf EA, Rugo HS, Solin LJ, Weaver DL, Winchester DJ, Hortobagyi GN. Breast Cancer-Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA Cancer J Clin*. 2017; 67:290–303. [a]

Giuliano AE, Edge SB, Hortobagyi GN. Eighth Edition of the AJCC Cancer Staging Manual: Breast Cancer. *Ann Surg Oncol*. 2018; 25:1783–1785.

Global Cancer Observatory. International Agency for Research on Cancer 2020. Available from <https://gco.iarc.fr/> [Accessed 28.12.2020]

Goethals A, Rose J. Mastectomy. StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK538212/> [Accessed 19.11.2020]

Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ; Panel members. 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; 22:1736–1747.

Gong J, Yu Y, Wu G, Lin C, Tu X. Should internal mammary lymph node sentinel biopsy be performed in breast cancer: a systematic review and meta-analysis. *World J Surg Oncol*. 2019; 17:135.

Goyal A, Dodwell D. POSNOC: A Randomised Trial Looking at Axillary Treatment in Women with One or Two Sentinel Nodes with Macrometastases. *Clin Oncol (R Coll Radiol)*. 2015; 27:692-695.

Gøtzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. *Cochrane Database of Systematic Reviews*. 2013; 2013:CD001877.

Gray RJ, Pockaj BA, Garvey E, Blair S. Intraoperative Margin Management in Breast-Conserving Surgery: A Systematic Review of the Literature. *Ann Surg Oncol*. 2018; 25:18–27.

Guo Q, Schmidt MK, Kraft P, Canisius S, Chen C, Khan S, Tyrer J, Bolla MK, Wang Q, Dennis J, Michailidou K, Lush M, Kar S, Beesley J, Dunning AM, Shah M, Czene K, Darabi H, Eriksson M, Lambrechts D, Weltens C, Leunen K, Bojesen SE, Nordestgaard BG, Nielsen SF, Flyger H, Chang-Claude J, Rudolph A, Seibold P, Flesch-Janys D, Blomqvist C, Aittomäki K, Fagerholm R, Muranen TA, Couch FJ, Olson JE, Vachon C, Andrulis IL, Knight JA, Glendon G, Mulligan AM, Broeks A, Hogervorst FB, Haiman CA, Henderson BE, Schumacher F, Le Marchand L, Hopper JL, Tsimiklis H, Apicella C, Southey MC, Cox A, Cross SS, Reed MW, Giles GG, Milne RL, McLean C, Winqvist R, Pylkäs K, Jukkola-Vuorinen A, Grip M, Hooning MJ, Hollestelle A, Martens JW, van den Ouweland AM, Marme F, Schneeweiss A, Yang R, Burwinkel B, Figueroa J, Chanock SJ, Lissowska J, Sawyer EJ, Tomlinson I, Kerin MJ, Miller N, Brenner H, Dieffenbach AK, Arndt V, Holleczek B, Mannermaa A, Kataja V, Kosma VM, Hartikainen JM, Li J, Brand JS, Humphreys K, Devilee P, Tollenaar RA, Seynaeve C, Radice P, Peterlongo P, Bonanni B, Mariani P, Fasching PA, Beckmann MW, Hein A, Ekici AB, Chenevix-Trench G, Balleine R; kConFab Investigators, Phillips KA, Benitez J, Zamora MP, Arias Perez JI, Menéndez P, Jakubowska A, Lubinski J, Jaworska-Bieniek K, Durda K, Hamann U, Kabisch M, Ulmer HU, Rüdiger T, Margolin S, Kristensen V, Nord S, Evans DG, Abraham JE, Earl HM, Hiller L, Dunn JA, Bowden S, Berg C, Campa D, Diver WR, Gapstur SM, Gaudet MM, Hankinson SE, Hoover RN, Hüsing A, Kaaks R, Machiela MJ, Willett W, Barrdahl M, Canzian F, Chin SF, Caldas C, Hunter DJ, Lindstrom S, García-Closas M, Hall P, Easton DF, Eccles DM, Rahman N, Nevanlinna H, Pharoah PD. Identification of novel genetic markers of breast cancer survival. *J Natl Cancer Inst.* 2015; 107:djv081.

Han L, Zhu Y, Liu Z, Yu T, He C, Jiang W, Kan Y, Dong D, Tian J, Luo Y. Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer. *Eur Radiol*. 2019; 29:3820–3829.

Haraldsdóttir KH, Jónsson Þ, Halldórsdóttir AB, Tranberg KG, Ásgeirsson KS. Tumor size of invasive breast cancer on magnetic resonance imaging and conventional imaging (mammogram/ultrasound): comparison with pathological size and clinical implications. *Scand J Surg.* 2017; 106:68–73.

Harrison BT, Brock JE. Contemporary evaluation of breast lymph nodes in anatomic pathology. *Am J Clin Pathol.* 2018; 150:4–17.

Hata T, Takahashi H, Watanabe K, Takahashi M, Taguchi K, Itoh T, Todo S. Magnetic resonance imaging for preoperative evaluation of breast cancer: a comparative study with mammography and ultrasonography. *J Am Coll Surg*. 2004; 198:190–197.

Hauge IH, Pedersen K, Olerud HM, Hole EO, Hofvind S. The risk of radiation-induced breast cancers due to biennial mammographic screening in women aged 50-69 years is minimal. *Acta Radiol.* 2014; 55:1174–1179.

Heinävaara S, Sarkeala T, Anttila A. Impact of organised mammography screening on breast cancer mortality in a case-control and cohort study. *Br J Cancer*. 2016; 114:1038–1044.

Heuts EM, van der Ent FW, von Meyenfeldt MF, Voogd AC. Internal mammary lymph drainage and sentinel node biopsy in breast cancer - A study on 1008 patients. *Eur J Surg Oncol*. 2009; 35:252–257.

Hieken TJ, Harrison J, Herreros J, Velasco JM. Correlating sonography, mammography, and pathology in the assessment of breast cancer size. *Am J Surg* 2001; 182:351–354.

Holli-Helenius K, Salminen A, Rinta-Kiikka I, Koskivuo I, Brück N, Boström P, Parkkola R. MRI texture analysis in differentiating luminal A and luminal B breast cancer molecular subtypes - a feasibility study. *BMC Med Imaging*. 2017; 17:69.

Houssami N, Macaskill P, Marinovich ML, Morrow M. The association of surgical margins and local recurrence in women with early-stage invasive breast cancer treated with breast-conserving therapy: a meta-analysis. *Ann Surg Oncol*. 2014; 21:717–730.

Houssami N, Sainsbury R. Breast cancer: multidisciplinary care and clinical outcomes. *Eur J Cancer.* 2006; 42:2480–2491.

Houssami N, Turner RM, Morrow M. Meta-analysis of pre-operative magnetic resonance imaging (MRI) and surgical treatment for breast cancer. *Breast Cancer Res Treat*. 2017; 165:273–283.

Houvenaeghel G, Lambaudie E, Classe JM, Mazouni C, Giard S, Cohen M, Faure C, Charitansky H, Rouzier R, Daraï E, Hudry D, Azuar P, Villet R, Gimbergues P, Tunon de Lara C, Martino M, Fraisse J, Dravet F, Chauvet MP, Boher JM. Lymph node positivity in different early breast carcinoma phenotypes: A predictive model. *BMC Cancer*. 2019; 19:45.

Hwang RF, Krishnamurthy S, Hunt KK, Mirza N, Ames FC, Feig B, Kuerer HM, Singletary SE, Babiera G, Meric F, Akins JS, Neely J, Ross MI. Clinicopathologic factors predicting involvement of nonsentinel axillary nodes in women with breast cancer. *Ann Surg Oncol*. 2003; 10:248–254.

Izci H, Tambuyzer T, Tuand K, Depoorter V, Laenen A, Wildiers H, Vergote I, Van Eycken L, De Schutter H, Verdoodt F, Neven P. A Systematic Review of Estimating Breast Cancer Recurrence at the Population Level With Administrative Data. *J Natl Cancer Inst*. 2020; 112:979–988.

Joensuu H, Asola R, Holli K, Kumpulainen E, Nikkanen V, Parvinen LM. Delayed diagnosis and large size of breast cancer after a false negative mammogram. *Eur J Cancer*. 1994; 30A:1299–1302.

Jones, J. Breast. *Underwood's Pathology: a Clinical Approach* (editors Cross S and Underwood JCE). Seventh edition. Edinburgh: Churchill Livingstone/Elsevier, 2019. pp. 416–437.

Jorns JM, Kidwell KM. Sentinel Lymph Node Frozen-Section Utilization Declines After Publication of American College of Surgeons Oncology Group Z0011 Trial Results With No Change in Subsequent Surgery for Axillary Lymph Node Dissection. *Am J Clin Pathol*. 2016; 146:57–66.

Karlsson A, Gonzalez V, Jaraj SJ, Bottai M, Sandelin K, Arver B, Eriksson S. The accuracy of incremental pre-operative breast MRI findings - Concordance with histopathology in the Swedish randomized multicenter POMB trial. *Eur J Radiol.* 2019; 114:185–191.

Katz MS, McCall L, Ballman K, Jagsi R, Haffty BG, Giuliano AE. Nomogram-based estimate of axillary nodal involvement in ACOSOG Z0011 (Alliance): validation and association with radiation protocol variations. *Breast Cancer Res Treat*. 2020; 180:429–436.

Katz B, Raker C, Edmonson D, Gass J, Stuckey A, Rizack T. Predicting Breast Tumor Size for Pre-operative Planning: Which Imaging Modality is Best? *Breast J*. 2017; 23:52–58.

Keam B, Im SA, Park S, Nam BH, Han SW, Oh DY, Kim JH, Lee SH, Han W, Kim DW, Kim TY, Park IA, Noh DY, Heo DS, Bang YJ. Nomogram predicting clinical outcomes in breast cancer patients treated with neoadjuvant chemotherapy. *J Cancer Res Clin Oncol*. 2011; 137:1301–1308.

Kesson EM, Allardice GM, George WD, Burns HJ, Morrison DS. Effects of multidisciplinary team working on breast cancer survival: retrospective, comparative, interventional cohort study of 13 722 women. *BMJ*. 2012; 344:e2718.

Kim MK, Han W, Moon HG, Ahn SK, Kim J, Lee JW, Kim JY, Kim T, Lee KH, Kim TY, Han SW, Im SA, Kim TY, Park IA, Noh DY. Nomogram for predicting breast conservation after neoadjuvant chemotherapy. *Cancer Res Treat*. 2015; 47:197–207.

Kim I, Ryu JM, Kim JM, Choi HJ, Lee SK, Yu JH, Lee JE, Kim SW, Nam SJ. Development of a nomogram to predict N2 or N3 stage in T1–2 invasive breast cancer patients with no palpable lymphadenopathy. *Journal of breast cancer*. 2017; 20:270–278.

Knuttel FM, Menezes GL, van Diest PJ, Witkamp AJ, van den Bosch MA, Verkooijen HM. Meta-analysis of the concordance of histological grade of breast cancer between core needle biopsy and surgical excision specimen. *Br J Surg*. 2016; 103:644–655.

Kondo T, Hayashi N, Ohde S, Suzuki K, Yoshida A, Yagata H, Niikura N, Iwamoto T, Kida K, Murai M, Takahashi Y, Tsunoda H, Nakamura S, Yamauchi H. A model to predict upstaging to invasive carcinoma in patients preoperatively diagnosed with ductal carcinoma in situ of the breast. *J Surg Oncol.* 2015; 112:476–480.

Krag DN, Anderson SJ, Julian TB, Brown AM, Harlow SP, Costantino JP, Ashikaga T, Weaver DL, Mamounas EP, Jalovec LM, Frazier TG, Noyes RD, Robidoux A, Scarth HM, Wolmark N. Sentinel-lymph-node resection compared with conventional axillary-lymph-node dissection in clinically node-negative patients with breast cancer: overall survival findings from the NSABP B-32 randomised phase 3 trial. *Lancet Oncol*. 2010; 11:927–933.

Kronqvist P, Kuopio T, Nykänen M, Helenius H, Anttinen J, Klemi P. Predicting aggressive outcome in T1N0M0 breast cancer. *Br J Cancer.* 2004; 91:277–281.

Kuerer HM, Smith BD, Chavez-MacGregor M, Albarracin C, Barcenas CH, Santiago L, Edgerton ME, Rauch GM, Giordano SH, Sahin A, Krishnamurthy S, Woodward W, Tripathy D, Yang WT, Hunt KK. DCIS Margins and Breast Conservation: MD Anderson Cancer Center Multidisciplinary Practice Guidelines and Outcomes. *J Cancer.* 2017; 8:2653–2662.

Kösters JP, Gøtzsche PC. Regular self-examination or clinical examination for early detection of breast cancer. *Cochrane Database Syst Rev*. 2003; 2003(2):CD003373.

Lai HW, Chen CJ, Lin YJ, Chen SL, Wu HK, Wu YT, Kuo SJ, Chen ST, Chen DR. Does breast magnetic resonance imaging combined with conventional imaging modalities decrease the rates of surgical margin involvement and reoperation?: A case-control comparative analysis. *Medicine*. 2016; 95:e3810.

Lakhani, Sunil R. *WHO Classification of Tumours of the Breast.* 4th ed. International Agency for Research on Cancer, (I A R C) (UN), Lyon, France. 2012. Available from: ProQuest Ebook Central.

Latosinsky S, Berrang TS, Cutter CS, George R, Olivotto I, Julian TB, Hayashi A, Baliski C, Croshaw RL, Erb KM, Chen J. CAGS and ACS Evidence Based Reviews in Surgery. 40. Axillary dissection versus no axillary dissection in women with invasive breast cancer and sentinel node metastasis. *Can J Surg.* 2012; 55:66–69.

Layfield DM, Agrawal A, Roche H, Cutress RI. Intraoperative assessment of sentinel lymph nodes in breast cancer. *Br J Surg*. 2011; 98:4-17.

Le EPV, Wang Y, Huang Y, Hickman S, Gilbert FJ. Artificial intelligence in breast imaging. *Clin Radiol.* 2019; 74:357–366.

Lee MC and Sabel MS. Axillary Management. *Early Diagnosis and Treatment of Cancer Series: Breast Cancer.* Philadelphia: Elsevier/Saunders. 2011: 217–239.

Lehman CD, Gatsonis C, Romanoff J, Khan SA, Carlos R, Solin LJ, Badve S, McCaskill-Stevens W, Corsetti RL, Rahbar H, Spell DW, Blankstein KB, Han LK, Sabol JL, Bumberry JR, Gareen I, Snyder BS, Wagner LI, Miller KD, Sparano JA, Comstock C. Association of Magnetic Resonance Imaging and a 12-Gene Expression Assay With Breast Ductal Carcinoma In Situ Treatment. *JAMA Oncol*. 2019; 5:1036–1042.

Li S, Yang X, Zhang Y, Fan L, Zhang F, Chen L, Zhou Y, Chen X, Jiang J. Assessment accuracy of core needle biopsy for hormone receptors in breast cancer: a meta-analysis. *Breast Cancer Res Treat.* 2012; 135:325–334.

Liu LC, Lang JE, Lu Y, Roe D, Hwang SE, Ewing CA, Esserman LJ, Morita E, Treseler P, Leong SP. Intraoperative frozen section analysis of sentinel lymph nodes in breast cancer patients: a meta-analysis and single-institution experience. *Cancer*. 2011; 117:250–258.

Lombardi A, Nigri G, Maggi S, Stanzani G, Vitale V, Vecchione A, Nania A, Amanti C. Role of frozen section in sentinel lymph node biopsy for breast cancer in the era of the ACOSOG Z0011 and IBCSG 23-10 trials. *Surgeon*. 2018; 16:232–236.

Lovrics PJ, Cornacchi SD, Vora R, Goldsmith CH, Kahnamoui K. Systematic review of radioguided surgery for non-palpable breast cancer. *Eur J Surg Oncol*. 2011; 37:388–397.

Lyman GH, Somerfield MR, Bosserman LD, Perkins CL, Weaver DL, Giuliano AE. Sentinel lymph node biopsy for patients with early-stage breast cancer: American society of clinical oncology clinical practice guideline update. *JCO*. 2016; 35:561–564.

Mann RM, Hoogeveen YL, Blickman JG, Boetes C. MRI compared to conventional diagnostic work-up in the detection and evaluation of invasive lobular carcinoma of the breast: a review of existing literature. *Breast Cancer Res Treat*. 2008; 107:1–14.

Mansel RE, Fallowfield L, Kissin M, Goyal A, Newcombe RG, Dixon JM, Yiangou C, Horgan K, Bundred N, Monypenny I, England D, Sibbering M, Abdullah TI, Barr L, Chetty U, Sinnett DH, Fleissig A, Clarke D, Ell PJ. Randomized multicenter trial of sentinel node biopsy versus standard axillary treatment in operable breast cancer: the ALMANAC Trial. *J Natl Cancer Inst*. 2006; 98:599–609.

Marinovich ML, Macaskill P, Irwig L, Sardanelli F, Mamounas E, von Minckwitz G, Guarneri V, Partridge SC, Wright FC, Choi JH, Bhattacharyya M, Martincich L, Yeh E, Londero V, Houssami N. Agreement between MRI and pathologic breast tumor size after neoadjuvant chemotherapy, and comparison with alternative tests: individual patient data meta-analysis. BMC Cancer. 2015; 15:662.

Mariscotti G, Durando M, Tagliafico A, Campanino PP, Bosco D, Casella C, Bussone R, Ala A, Castellano I, Sapino A, Bergamasco L, Fonio P, Houssami N. Preoperative breast cancer staging with multi-modality imaging and surgical outcomes. *Eur J Radiol*. 2019; 122:108766.

Mathioudakis AG, Salakari M, Pylkkanen L, Saz-Parkinson Z, Bramesfeld A, Deandrea S, Lerda D, Neamtiu L, Pardo-Hernandez H, Solà I, Alonso-Coello P. Systematic review on women's values and preferences concerning breast cancer screening and diagnostic services. *Psychooncology*. 2019; 28:939–947.

Matsen C, Van Zee KJ. Use of established nomograms to predict non-sentinel lymph node metastasis. *Curr Breast Cancer Rep*. 2014; 6:24–31.

Mazzei MA, Di Giacomo L, Fausto A, Gentili F, Mazzei FG, Volterrani L. Efficacy of Second-Look Ultrasound with MR Coregistration for Evaluating Additional Enhancing Lesions of the Breast: Review of the Literature. *Biomed Res Int*. 2018; 2018:3896946.

McCart Reed AE, Kalita-De Croft P, Kutasovic JR, Saunus JM, Lakhani SR. Recent advances in breast cancer research impacting clinical diagnostic practice. *J Pathol*. 2019; 247:552–562.

McTiernan A. Behavioral risk factors in breast cancer: can risk be modified? *Oncologist*. 2003; 8:326–334.

Mejdahl MK, Andersen KG, Gärtner R, Kroman N, Kehlet H. Persistent pain and sensory disturbances after treatment for breast cancer: six year nationwide follow-up study. *BMJ*. 2013; 346:f1865.

Meretoja TJ, Audisio RA, Heikkilä PS, Bori R, Sejben I, Regitnig P, Luschin-Ebengreuth G, Zgajnar J, Perhavec A, Gazic B, Lázár G, Takács T, Kővári B, Saidan ZA, Nadeem RM, Castellano I, Sapino A, Bianchi S, Vezzosi V, Barranger E, Lousquy R, Arisio R, Foschini MP, Imoto S, Kamma H, Tvedskov TF, Jensen MB, Cserni G, Leidenius MH. International multicenter tool to predict the risk of four or more tumor-positive axillary lymph nodes in breast cancer patients with sentinel node macrometastases. *Breast Cancer Res Treat*. 2013; 138:817–827.

Meretoja TJ, Heikkilä PS, Mansfield AS, Cserni G, Ambrozay E, Boross G, Zgajnar J, Perhavec A, Gazic B, Arisio R, Tvedskov TF, Jensen MB, Leidenius MH. A predictive tool to estimate the risk of axillary metastases in breast cancer patients with negative axillary ultrasound. *Ann Surg Oncol*. 2014; 21:2229–2236.

Meretoja TJ, Leidenius MH, Heikkilä PS, Boross G, Sejben I, Regitnig P, Luschin-Ebengreuth G, Žgajnar J, Perhavec A, Gazic B, Lázár G, Takács T, Vörös A, Saidan ZA, Nadeem RM, Castellano

I, Sapino A, Bianchi S, Vezzosi V, Barranger E, Lousquy R, Arisio R, Foschini MP, Imoto S, Kamma H, Tvedskov TF, Kroman N, Jensen MB, Audisio RA, Cserni G. International multicenter tool to predict the risk of nonsentinel node metastases in breast cancer. *J Natl Cancer Inst.* 2012; 104:1888–1896.

Meretoja TJ, Strien L, Heikkilä PS, Leidenius MH. A simple nomogram to evaluate the risk of nonsentinel node metastases in breast cancer patients with minimal sentinel node involvement. *Ann Surg Oncol.* 2011; 19:567–576.

Merkel DE, Osborne CK. Prognostic factors in breast cancer. *Hematol Oncol Clin North Am*. 1989; 3:641–652.

Middleton LP, Vlastos G, Mirza NQ, Eva S, Sahin AA. Multicentric mammary carcinoma: evidence of monoclonal proliferation. *Cancer*. 2002; 94:1910–1916.

Montazeri M, Montazeri M, Montazeri M, Beigzadeh A. Machine learning models in breast cancer survival prediction. *Technol Health Care*. 2016; 24:31-42.

Monticciolo DL. Magnetic resonance imaging of the breast for cancer diagnosis and staging. *Semin Ultrasound CT MR*. 2011; 32:319–330.

Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ.* 2009; 338:b375.

Morrow M. Management of the Node-Positive Axilla in Breast Cancer in 2017: Selecting the Right Option. *JAMA Oncol*. 2018; 4:250–251.

Myers ER, Moorman P, Gierisch JM, Havrilesky LJ, Grimm LJ, Ghate S, Davidson B, Mongtomery RC, Crowley MJ, McCrory DC, Kendrick A, Sanders GD. Benefits and Harms of Breast Cancer Screening: A Systematic Review. *JAMA*. 2015; 314:1615–1634.

Myller S, Ipatti P, Jääskeläinen A, Haapasaari KM, Jukkola A, Karihtala P. Early progression of breast cancer during neoadjuvant chemotherapy may predict poorer prognoses. *Acta Oncol*. 2020; 59: 1036–1042.

National Cancer Institute [a]. NCI Dictionary of cancer terms. Available from https://www.cancer.gov/publications/dictionaries/cancer-terms [Accessed 19.3.2021]

National Cancer Institute [b]. Surveillance, Epidemiology, and End Results (SEER) Program. Cancer Stat Facts: Female Breast Cancer. Available from https://seer.cancer.gov/statfacts/html/breast.html [Accessed 19.3.2021]

Nahabedian MY. Managing the opposite breast: contralateral symmetry procedures. *Cancer* J. 2008; 14:258–263.

NCCN Clinical Practice Guidelines in Oncology. Version 6.2020. National Comprehensive Cancer Network. 2020. Available from <www.nccn.org> [Accessed 15.12.2020]

Nguyen PL, Taghian AG, Katz MS, Niemierko A, Abi Raad RF, Boon WL, Bellon JR, Wong JS, Smith BL, Harris JR. Breast cancer subtype approximated by estrogen receptor, progesterone receptor, and HER-2 is associated with local and distant recurrence after breast-conserving therapy. *J Clin Oncol*. 2008; 26:2373–2378.

Ngô C, Mouttet D, De Rycke Y, Reyal F, Fourchotte V, Hugonnet F, Falcou MC, Bidard FC, Vincent-Salomon A, Fourquet A, Alran S. Validation over time of a nomogram including HER2 status to predict the sentinel node positivity in early breast carcinoma. *Eur J Surg Oncol*. 2012; 38:1211–1217.

Nielsen Moody A, Bull J, Culpan AM, Munyombwe T, Sharma N, Whitaker M, Wolstenhulme S. Preoperative sentinel lymph node identification, biopsy and localisation using contrast enhanced ultrasound (CEUS) in patients with breast cancer: a systematic review and meta-analysis. *Clin Radiol*. 2017; 72:959–971.

Nieweg OE, Tanis PJ, Kroon BB. The definition of a sentinel node. *Ann Surg Oncol*. 2001; 8:538–541.

Nindrea RD, Aryandono T, Lazuardi L, Dwiprahasto I. Diagnostic Accuracy of Different Machine Learning Algorithms for Breast Cancer Risk Calculation: a Meta-Analysis. *Asian Pac J Cancer Prev.* 2018; 19:1747–1752.

Nowikiewicz T, Wnuk P, Srutek E, Kowalewski J, Zegarski W. Validation of new predictive tool of non-sentinel lymph node involvement in breast cancer patients with positive sentinel lymph node biopsies. *European Journal of Cancer.* 2014; 50:S12–S159.

Nowikiewicz T, Wnuk P, Małkowski B, Kurylcio A, Kowalewski J, Zegarski W. Application of artificial neural networks for predicting presence of non-sentinel lymph node metastases in breast cancer patients with positive sentinel lymph node biopsies. *Archives of medical science*. 2017; 13:1399–1407.

Ofri A, Moore K. Occult breast cancer: Where are we at? *Breast.* 2020; 54:211-215.

Pal SK, Childs BH, Pegram M. Triple negative breast cancer: unmet medical needs. *Breast Cancer Res Treat*. 2011; 125:627–636.

Paluch-Shimon S, Cardoso F, Sessa C, Balmana J, Cardoso MJ, Gilbert F, Senkus E, on behalf of the ESMO Guidelines Committee. Prevention and screening in BRCAmutation carriers and other breast/ovarian hereditary cancer syndromes: ESMO Clinical Practice Guidelines for cancer preventionand screening. *Annals of Oncology*. 2016; 27: v103–v110.

Papanikolaou IG, Dimitrakakis C, Zagouri F, Marinopoulos S, Giannos A, Zografos E, Zografos CG, Kritikou D, Rodolakis A, Zografos GC, Loutradis D. Paving the way for changing perceptions in breast surgery: a systematic literature review focused on oncological and aesthetic outcomes of oncoplastic surgery for breast cancer. *Breast Cancer*. 2019; 26:416–427.

Phung MT, Tin Tin S, Elwood JM. Prognostic models for breast cancer: a systematic review. *BMC Cancer.* 2019; 19:230.

Pepels MJ, Vestjens JH, de Boer M, Bult P, Van Dijck JA, Menke-Pluijmers M, van Diest PJ, Borm G, Tjan-Heijnen VC. Models predicting non-sentinel node involvement also predict for regional recurrence in breast cancer patients without axillary treatment. *European Journal of Surgical Oncology.* 2013; 39:1351–1357.

Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L. *European guidelines for quality assurance in breast cancer screening and diagnosis*, fourth edition. European Communities. 2006.

Pinto AE, André S, Pereira T, Nóbrega S, Soares J. Prognostic comparative study of S-phase fraction and Ki-67 index in breast carcinoma. *Journal of Clinical Pathology*. 2001; 54:543–549.

Plichta JK, Campbell BM, Mittendorf EA, Hwang ES. Anatomy and Breast Cancer Staging: Is It Still Relevant? *Surg Oncol Clin N Am*. 2018; 27:51–67.

Poortmans PM, Collette S, Kirkove C, Van Limbergen E, Budach V, Struikmans H, Collette L, Fourquet A, Maingon P, Valli M, De Winter K, Marnitz S, Barillot I, Scandolaro L, Vonk E, Rodenhuis C, Marsiglia H, Weidner N, van Tienhoven G, Glanzmann C, Kuten A, Arriagada R, Bartelink H, Van den Bogaert W; EORTC Radiation Oncology and Breast Cancer Groups. Internal Mammary and Medial Supraclavicular Irradiation in Breast Cancer. *N Engl J Med*. 2015; 373:317–327.

Prat A, Pineda E, Adamo B, Galván P, Fernández A, Gaba L, Díez M, Viladot M, Arance A, Muñoz M. Clinical implications of the intrinsic molecular subtypes of breast cancer. Breast. 2015; 24:S26–35.

Pyo JS, Jung J, Lee SG, Kim NY, Kang DW. Diagnostic Accuracy of Fine-Needle Aspiration Cytology and Core-Needle Biopsy in the Assessment of the Axillary Lymph Nodes in Breast Cancer-A Meta-Analysis. *Diagnostics (Basel)*. 2020; 10:717.

Qiu X, Jiang Y, Zhao Q, Yan C, Huang M, Jiang T. Could Ultrasound-Based Radiomics Noninvasively Predict Axillary Lymph Node Metastasis in Breast Cancer? *J Ultrasound Med.* 2020; 39:1897–1905.

Rao AA, Feneis J, Lalonde C, Ojeda-Fournier H. A Pictorial Review of Changes in the BI-RADS Fifth Edition. *Radiographics.* 2016; 36:623–639.

Rao R, Euhus D, Mayo HG, Balch C. Axillary node interventions in breast cancer. A systematic review. *JAMA*. 2013; 310:1385–1394.

Reynolds C, Mick R, Donohue JH, Grant CS, Farley DR, Callans LS, Orel SG, Keeney GL, Lawton TJ, Czerniecki BJ. Sentinel lymph node biopsy with metastasis: can axillary dissection be avoided in some patients with breast cancer? *J Clin Oncol*. 1999; 17:1720–1726.

Rouzier R, Pusztai L, Garbay JR, Delaloge S, Hunt KK, Hortobagyi GN, Berry D, Kuerer HM. Development and validation of nomograms for predicting residual tumor size and the probability of successful conservative surgery with neoadjuvant chemotherapy for breast cancer. *Cancer*. 2006; 107:1459–1466.

Sacre RA. Clinical evaluation of axillar lymph nodes compared to surgical and pathological findings. *Eur J Surg Oncol*. 1986; 12:169–173.

Saha M, Mukherjee R, Chakraborty C. Computer-aided diagnosis of breast cancer using cytological images: A systematic review. *Tissue Cell.* 2016; 48:461–474.

Sakorafas GH, Safioleas M. Breast cancer surgery: an historical narrative. Part III. From the sunset of the 19th to the dawn of the 21st century. *Eur J Cancer Care (Engl).* 2010; 19:145–166.

Sardanelli F, Boetes C, Borisch B, Decker T, Federico M, Gilbert FJ, Helbich T, Heywang-Köbrunner SH, Kaiser WA, Kerin MJ, Mansel RE, Marotti L, Martincich L, Mauriac L, Meijers-Heijboer H, Orecchia R, Panizza P, Ponti A, Purushotham AD, Regitnig P, Del Turco MR, Thibault F, Wilson R. Magnetic resonance imaging of the breast: recommendations from the EUSOMA working group. Magnetic resonance imaging of the breast: recommendations from the EUSOMA working group. *Eur J Cancer*. 2010; 46:1296–1316.

Sarkeala T, Luostarinen T, Dyba T, Anttila A. Breast carcinoma detection modes and death in a female population in relation to population-based mammography screening. *Springerplus*. 2014; 3:348.

Sawaki M, Shien T, Iwata H. TNM classification of malignant tumors (Breast Cancer Study Group). *Jpn J Clin Oncol*. 2019; 49:228–231.

Shah R, Rosso K, Nathanson SD. Pathogenesis, prevention, diagnosis and treatment of breast cancer. *Worl J Clin Oncol*. 2014; 5:283–298.

Shao J, Rodrigues M, Corter AL, Baxter NN. Multidisciplinary care of breast cancer patients: a scoping review of multidisciplinary styles, processes, and outcomes. *Curr Oncol.* 2019; 26:e385–e397.

Siesling S, Hueting T, Tip B, Mentink R, Koffijberg E. Clinical risk prediction models for breast cancer: A review of models developed between 2010 and 2018. *Cancer Research*. 2019; 79:4 Supplement 1.

Singh D, Singh AK. Role of image thermography in early breast cancer detection- Past, present and future. *Comput Methods Programs Biomed*. 2020; 183:105074.

Stamatakos MD. Invasive breast cancer. *Early Diagnosis and Treatment of Cancer Series: Breast Cancer*. Philadelphia: Elsevier/Saunders. 2011. Pages 21-54.

Sugimoto M, Takada M, Toi M. Development of web tools to predict axillary lymph node metastasis and pathological response to neoadjuvant chemotherapy in breast cancer patients. *Int J Biol Markers.* 2014; 29:e372–379.

Tajima CC, de Sousa LLC, Venys GL, Guatelli CS, Bitencourt AGV, Marques EF. Magnetic resonance imaging of the breast: role in the evaluation of ductal carcinoma in situ. Magnetic resonance imaging of the breast: role in the evaluation of ductal carcinoma in situ. *Radiol Bras*. 2019; 52:43–47.

Tanis PJ, Nieweg OE, Valdés Olmos RA, Kroon BB. Anatomy and physiology of lymphatic drainage of the breast from the perspective of sentinel node biopsy. *J Am Coll Surg.* 2001; 192:399–409.

Tapia G, Ying V, Di Re A, Stellin A, Cai TY, Warrier S. Predicting non-sentinel lymph node metastasis in Australian breast cancer patients: are the nomograms still useful in the post-Z0011 era? *ANZ J Surg*. 2019; 89:712–717.

Turku University Hospital instructions. Institutional instructions for clinicians, Department of Oncology, Turku University Hospital. Updated 30.1.2020.

Turnbull L, Brown S, Harvey I, Olivier C, Drew P, Napp V, Hanby A, Brown J. Comparative effectiveness of MRI in breast cancer (COMICE) trial: a randomised controlled trial. *Lancet*. 2010; 375:563–571.

Turner RR, Chu KU, Qi K, Botnick LE, Hansen NM, Glass EC, Giuliano AE. Pathologic features associated with nonsentinel lymph node metastases in patients with metastatic breast carcinoma in a sentinel lymph node. *Cancer.* 2000; 89:574–581.

Tyks Cancer Centre webpage. Treatment paths for cancer patients at the Western Finland Cancer Centre. Treatment path of patient with breast cancer. Updated 23.10.2019. Available from <https://www.vsshp.fi/en/syopakeskus/potilaalle/hoitopolut/Pages/default.aspx.> [Accessed 29.12.2020]

Tvedskov TF, Meretoja TJ, Jensen MB, Leidenius M, Kroman N. Cross-validation of three predictive tools for non-sentinel node metastases in breast cancer patients with micrometastases or isolated tumor cells in the sentinel node. *Eur J Surg Oncol*. 2014; 40:435-441.

van den Ende C, Oordt-Speets AM, Vroling H, van Agt HME. Benefits and harms of breast cancer screening with mammography in women aged 40-49 years: A systematic review. *Int J Cancer.* 2017; 141:1295–1306.

van den Hoven I, Kuijt G, Roumen R, Voogd A, Steyerberg EW, Vergouwe Y. A head to head comparison of nine tools predicting non-sentinel lymph node status in sentinel node positive breast cancer women. *J Surg Oncol*. 2015; 112:133-8.

van der Ploeg IM, Nieweg OE, van Rijk MC, Valdés Olmos RA, Kroon BB. Axillary recurrence after a tumour-negative sentinel node biopsy in breast cancer patients: A systematic review and meta-analysis of the literature. *Eur J Surg Oncol*. 2008; 34:1277–1284.

van la Parra RFD, Peer PG, Ernst MF, Bosscha K. Meta-analysis of predictive factors for non-sentinel lymph node metastases in breast cancer patients with a positive SLN. *Eur J Surg Oncol.* 2011; 37:290–299.

van Maaren MC, de Munck L, de Bock GH, Jobsen JJ, van Dalen T, Linn SC, Poortmans P, Strobbe LJA, Siesling S. 10 year survival after breast-conserving surgery plus radiotherapy compared with mastectomy in early breast cancer in the Netherlands: a population-based study. *Lancet Oncol*. 2016; 17:1158–1170.

van Nijnatten TJA, Ploumen EH, Schipper RJ, Goorts B, Andriessen EH, Vanwetswinkel S, Schavemaker M, Nelemans P, de Vries B, Beets-Tan RGH, Smidt ML, Lobbes MBI. Routine use of standard breast MRI compared to axillary ultrasound for differentiating between no, limited and advanced axillary nodal disease in newly diagnosed breast cancer patients. *Eur J Radiol.* 2016; 85:2288–2294.

Van Zee KJ, Manasseh DM, Bevilacqua JL, Boolbol SK, Fey JV, Tan LK, Borgen PI, Cody HS 3rd, Kattan MW. A nomogram for predicting the likelihood of additional nodal metastases in breast cancer patients with a positive sentinel node biopsy. *Ann Surg Oncol*. 2003; 10:1140–1151.

Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol*. 2005; 58:475–483.

Veronesi U, Paganelli G, Viale G, Luini A, Zurrida S, Galimberti V, Intra M, Veronesi P, Robertson C, Maisonneuve P, Renne G, De Cicco C, De Lucia F, Gennari R. A randomized comparison of sentinel-node biopsy with routine axillary dissection in breast cancer. *N Engl J Med*. 2003; 349:546–553.

Veronesi U, Stafyla V, Luini A, Veronesi P. Breast cancer: from "maximum tolerable" to "minimum effective" treatment. *Front Oncol*. 2012; 2:125.

von Euler-Chelpin M, Lillholm M, Vejborg I, Nielsen M, Lynge E. Sensitivity of screening mammography by density and texture: a cohort study from a population-based screening program in Denmark. *Breast Cancer Res*. 2019; 21:111.

Waks AG, Winer EP. Breast Cancer Treatment: A Review. *JAMA*. 2019; 321:288–300.

Wallis MG, Kilburn-Toppin F, Taylor-Phillips S. Does preoperative axillary staging lead to overtreatment of women with screen-detected breast cancer? *Clin Radiol* 2018; 73:467–472.

Walsh R, Kornguth PJ, Soo MS, Bentley R, DeLong DM. Axillary lymph nodes: mammographic, pathologic, and clinical correlation. *AJR Am J Roentgenol*. 1997; 168:33–38.

Wang L, Cohen JC, Devasenapathy N, Hong BY, Kheyson S, Lu D, Oparin Y, Kennedy SA, Romerosa B, Arora N, Kwon HY, Jackson K, Prasad M, Jayasekera D, Li A, Guarna G, Natalwalla S, Couban RJ, Reid S, Khan JS, McGillion M, Busse JW. Prevalence and intensity of persistent post-surgical

pain following breast cancer surgery: a systematic review and meta-analysis of observational studies. *Br J Anaesth*. 2020; 125:346–357.

Wang M, He X, Chang Y, Sun G, Thabane L. A sensitivity and specificity comparison of fine needle aspiration cytology and core needle biopsy in evaluation of suspicious breast lesions: A systematic review and meta-analysis. *Breast*. 2017; 31:157–166.

Wang XW, Xiong YH, Zen XQ, Lin HB, Liu QY. Diagnostic accuracy of ultrasonograph guided fine-needle aspiration cytologic in staging of axillary lymph node metastasis in breast cancer patients: a meta-analysis. *Asian Pac J Cancer Prev*. 2012; 13:5517–5523.

Wang Z, Wu LC, Chen JQ. Sentinel lymph node biopsy compared with axillary lymph node dissection in early breast cancer: a meta-analysis. *Breast Cancer Res Treat*. 2011; 129:675–689.

Weiss A, Chavez-MacGregor M, Lichtensztajn DY, Yi M, Tadros A, Hortobagyi GN, Giordano SH, Hunt KK, Mittendorf EA. Validation Study of the American Joint Committee on Cancer Eighth Edition Prognostic Stage Compared With the Anatomic Stage in Breast Cancer. *JAMA Oncol.* 2018; 4:203–209.

Werkoff G, Lambaudie E, Fondrinier E, Levêque J, Marchal F, Uzan M, Barranger E, Guillemin F, Darai E, Uzan S, Houvenaeghel G, Rouzier R, Coutant C. Prospective multicenter comparison of models to predict four or more involved axillary lymph nodes in patients with breast cancer with one to three metastatic sentinel lymph nodes. *J Clin Oncol*. 2009; 27:5707–5712.

WHO Classification of Tumours Editorial Board. *Breast Tumours*. Fifth edition. International Agency for Research on Cancer. Lyon (France). 2019.

Willett AM, Michell MJ, Lee MJR. *Best practice diagnostic guidelines for patients presenting with breast symptoms*. Association of Breast Surgery, The National Institute for Health and Clinical Excellence (NICE). 2010. Available from https://associationofbreastsurgery.org.uk/media/1416/best-practice-diagnostic-guidelines-for-patients-presenting-with-breast-symptoms.pdf [Accessed 30.12.2020]

Wolff AC, Hammond MEH, Allison KH, Harvey BE, Mangu PB, Bartlett JMS, Bilous M, Ellis IO, Fitzgibbons P, Hanna W, Jenkins RB, Press MF, Spears PA, Vance GH, Viale G, McShane LM, Dowsett M. Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. *J Clin Oncol*. 2018; 36:2105–2122.

Wu P, Zhao K, Liang Y, Ye W, Liu Z, Liang C. Validation of Breast Cancer Models for Predicting the Nonsentinel Lymph Node Metastasis After a Positive Sentinel Lymph Node Biopsy in a Chinese Population. *Technol Cancer Res Treat*. 2018; 17:1533033818785032.

Yan S, Wang W, Zhu B, Pan X, Wu X, Tao W. Construction of Nomograms for Predicting Pathological Complete Response and Tumor Shrinkage Size in Breast Cancer. *Cancer Manag Res*. 2020; 12:8313–8323.

Yassin NIR, Omran S, El Houby EMF, Allam H. Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review. *Comput Methods Programs Biomed*. 2018; 156:25–45.

Yin L, Duan JJ, Bian XW, Yu SC. Triple-negative breast cancer molecular subtyping and treatment progress. *Breast Cancer Res*. 2020; 22:61.

Zhang M, Sun S, and Mesurolle B. The impact of pre-operative breast MRI on surgical waiting time. *PLoS One.* 2017;12:e0169756.

Zhu L, Jin L, Li S, Chen K, Jia W, Shan Q, Walter S, Song E, Su F. Which nomogram is best for predicting non-sentinel lymph node metastasis in breast cancer patients? A meta-analysis. *Breast Cancer Res Treat*. 2013; 137:783–795.

Zielonke, N., Gini, A., Jansen, E., Anttila, A., Segnan, N., Ponti, A., Veerus, P., de Koning, H. J., van Ravesteyn, N. T., Heijnsdijk, E., & EU-TOPIA consortium. Evidence for reducing cancer-specific mortality due to screening for breast cancer in Europe: A systematic review. *Eur J Cancer*. 2020; 127:191–206.

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama Oy, Turku, Finland 2021

ISBN 978-951-29-8440-4 (PRINT)
ISBN 978-951-29-8441-1 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)