TISSUE MICROARRAY TECHNOLOGY IN BREAST CANCER SUBTYPING

Anssi Nieminen

Master's Thesis

University of Jyväskylä

Biotechnology

August 2008

PREFACE

This study was conducted at the Central Hospital of Central Finland, Department of Pathology during autumn 2006 and spring 2007. The text was finalised at the University of Kuopio during the following year. The study is about tissue microarray technology, a relatively novel method in the field of histology. Equally much the study is about breast cancer, for it is the most common cancer among women, and the cause of too many deaths despite the improvements in both treatment and diagnostics. Dr. Kuopio from the Department of Pathology suggested this topic to me, and I was more than happy to accept it; it contained many interesting aspects and covered an important topic as well: how to further enhance methods in the field of breast cancer research? This work, on its part, aims to answer this question, but it also became evident that much remains for future researchers to study!

Acknowledgements

I wish to express my gratitude to the following persons: Teijo Kuopio, M.D., chief pathologist at the Central Hospital of Central Finland, Department of Pathology for supervision and guidance during this project, for professional views, and for making this study possible. Likewise, I want to thank Jorma Anttinen, M.D., Reino Pitkänen, M.Sc., and Marjukka Nykänen, M.Sc., for professional support in the fields of pathology and immunohistochemistry. Special thanks go to Juha Kononen, M.D., from Beecher Instruments, who kindly provided the tissue microarrayer used in this study and gave valuable specialist opinions regarding the project. Also, I wish to thank the staff of the Department of Pathology for help and guidance in various technical aspects. Finally, the text was revised and proofread by Milla Vilkman, M.A., which is greatly appreciated.

Anssi Nieminen
August 2008

University of Jyväskylä
Abstract of Master´s Thesis
Faculty of Mathematics and Science

Author: Anssi Nieminen
Title of thesis: Tissue Microarray Technology in Breast Cancer Subtyping
Finnish title: Tissue microarray -teknologia rintasyövän tyypityksessä
Date: 20.8.2008 **Pages:** 62
Department: Department of Biological and Environmental Science
Chair: Biotechnology
Supervisor(s): Teijo Kuopio

Abstract:

Breast carcinoma is the most common cancer in the world among women and it is an important cause of death. For decades, breast carcinoma research and diagnostics have largely relied on traditional histological methods, such as immunohistochemical staining of tissue samples. As a reliable and well-established set of methods, immunohistochemistry is the cornerstone of diagnostics, but it is often too laborious and time-consuming to satisfy the needs of high-throughput requirements of research.

Tissue microarray (TMA) technology is a relatively novel innovation aiming to considerably increase the amount of sample specimens researchers can produce from formalin-fixed, paraffin-embedded tissue samples. TMA technology uses special arrayers, which can place hundreds of histological sample specimens on a single paraffin block, thus allowing the simultaneous staining of these samples, requiring the time and reagents of just one histological staining. Furthermore, up to several dozens identical sections can be prepared from one TMA block, which can then be stained for various cancer markers or other proteins of interest. These characteristics together make TMA technology an intriguing tool for histological research. Finally, computer-assisted arrayers and automated optical readers can further enhance the technology in the near future.

The purpose of this study was to evaluate the characteristics of TMA technology and its suitability for the research purposes of a hospital pathology laboratory. Three TMA blocks consisting of 734 cancer samples from 367 patients were created by a Beecher Instruments Manual Tissue Arrayer. These samples were then stained for eight cancer marker proteins, and the results were compared to values found in the literature. In addition, the time and effort of constructing and analysing the TMA was considered. Finally, the results of this study can be used for multivariable studies in the future, when linked with the patient diagnostic and follow-up data, but this remained beyond the scope of this work.

Oestrogen receptor was found to be positive in 76.3% of cases. Progesterone receptor was positive in 65.7% and the other cancer markers were positive as follows: Her-2 18.8%, E-Cad 92.5%, cytokeratins 5 and 17 14.5%, BCL-2 80.5%, and CD 117 3.2%. The amount of patients tested ranged from 257 to 313 per marker, where from 56 to 94 patient samples, depending on the given marker, remained unrepresentative (no cancer cells in the sample area) and were therefore discarded from the analysis. On average, 22% of sample specimens were found to be unrepresentative. For one person, it required approximately four months of four-hour working days to construct and analyse the TMA.

The results obtained from this study are well in line with results found in the literature in majority of the markers analysed, indicating that TMA technology is a useful tool even for a research project with limited resources regarding available space, funds or work force. With some practice, a full-time worker is likely to be able to achieve a much higher sample throughput than that described in this work.

Keywords: tissue microarray technology, breast cancer, breast carcinoma, immunohistochemistry, cancer marker, marker protein

Jyväskylän yliopisto
Pro gradu –tutkielman tiivistelmä
Matemaattis-luonnontieteellinen tiedekunta

Tekijä: Anssi Nieminen
Tutkielman nimi: Tissue microarray -teknologia rintasyövän tyypityksessä
English title: Tissue Microarray Technology in Breast Cancer Subtyping
Päivämäärä: 20.8.2008 **Sivumäärä:** 62
Laitos: Bio- ja ympäristötieteiden laitos
Oppiaine: Biotekniikka
Tutkielman ohjaaja(t): Teijo Kuopio

Tiivistelmä:

Rintasyöpä on yleisin naisten syöpä maailmassa ja se on merkittävä kuolinsyy. Vuosikymmenten ajan rintasyöpätutkimuksessa ja -diagnostiikassa on laajalti käytetty perinteisiä histologisia menetelmiä, kuten kudosnäytteiden immunohistokemiallista värjäystä. Immunohistokemia on luotettava ja vakiintunut menetelmä, ja siksi se on diagnostiikan kulmakivi, mutta usein se on liian työlästä ja aikaa vievää täyttääkseen tutkimuksen suurten näytevolyymien vaatimukset.

Tissue microarray (TMA) -teknologia on melko uusi keksintö, jonka tavoitteena on lisätä merkittävästi niiden näytteiden määrää, jotka tutkijat voivat valmistaa formaliinikiinnitetyistä, parafiiniin valetuista kudosnäytteistä. TMA-teknologiassa käytetään erityistä laitteistoa, jonka avulla voidaan asettaa satoja histologisia näytteitä yhdelle parafiiniblokille, joka mahdollistaa näiden näytteiden samanaikaisen värjäyksen ja vaatii vain yhden histologisen värjäyksen verran aikaa ja reagensseja. Lisäksi jopa useita kymmeniä identtisiä leikkeitä voidaan valmistaa yhdestä TMA-blokista, joista voidaan värjätä useita syöpämarkkereita tai muita kiinnostuksen kohteena olevia proteiineja. Nämä piirteet tekevät yhdessä TMA-teknologiasta kiehtovan työvälineen histologiseen tutkimukseen. Lisäksi tietokoneavusteiset laitteistot ja automatisoidut optiset lukijat voivat vielä parantaa teknologiaa lähitulevaisuudessa.

Tämän tutkimuksen tarkoituksena oli arvioida TMA-teknologian ominaisuuksia ja sen sopivuutta sairaalan patologian laboratorion tutkimustarkoituksiin. Kolme TMA-blokkia, jotka käsittivät 734 syöpänäytettä 367 potilaalta luotiin Beecher Instruments Manual Tissue Arrayerillä. Näytteistä värjättiin kahdeksan syöpämarkkeriproteiinia, ja tuloksia verrattiin kirjallisuudessa esiintyviin arvoihin. Lisäksi arvioitiin TMA:n valmistamisen ja analysoinnin vaatimaa aikaa ja työpanosta. Tämän tutkimuksen tuloksia voidaan jatkossa käyttää monimuuttuja-analyyseihin, kun ne liitetään potilaiden diagnostisiin ja seurantatietoihin, mutta tässä työssä aihetta ei käsitelty.

Estrogeenireseptori oli positiivinen 76,3 %:ssa tapauksista. Progesteronireseptori oli positiivinen 65,7 %:ssa tapauksista ja muut syöpämarkkerit olivat positiivisia seuraavasti: Her-2 18,8 %, E-Cad 92,5 %, sytokeratiinit 5 ja 17 14,5 %, BCL-2 80,5 %, ja CD 117 3,2 %. Testattujen potilaiden määrä vaihteli 257:stä 313:een markkeria kohti, joista 56 - 94 potilasnäytettä markkerista riippuen olivat epäedustavia (näytteessä ei syöpäsoluja), ja siksi ne poistettiin analyysistä. Keskimäärin 22 % näytteistä oli epäedustavia. Yhdeltä ihmiseltä kului noin neljä tuntia päivässä neljän kuukauden ajan valmistaa ja analysoida tutkimuksen TMA:t.

Suuressa osassa analysoiduista markkereista tämän tutkimuksen tulokset ovat hyvin samanlaiset kuin kirjallisuudessa esiintyvät tulokset, mikä viittaa siihen, että TMA-teknologia on hyödyllinen työväline myös tutkimusprojekteissa, joissa on rajallisesti tilaa, rahoitusta tai henkilöstöä. Harjoittelemalla kokoaikainen työntekijä saavuttaa todennäköisesti suuremman näytevolyymin kuin tässä työssä saavutettiin.

Avainsanat: tissue microarray -teknologia, rintasyöpä, rintakarsinooma, immunohistokemia, syöpämarkkeri, markkeriproteiini

ABBREVIATIONS

CISH	chromogenic *in situ* hybridisation
CK	cytokeratine
DCIS	ductal carsinoma *in situ*
FISH	fluorescent *in situ* hybridisation
HE	haematoxylin-eosin
HR	hormone receptor
IHC	immunohistochemistry
LCIS	lobular carcinoma *in situ*
MTA	Manual Tissue Arrayer
MTTB	multitumour tissue block
PEFF	paraffin-embedded, formalin-fixed
PR	progesterone receptor
TMA	tissue microarray
VTM	virtual tissue matrix

1. INTRODUCTION

Immunohistochemical methods have been among the cornerstones of modern pathology for decades. Despite this, relatively little has changed regarding the standard immunohistological methods: sections prepared by microtome from formalin fixed, paraffin embedded tissue samples (PEFF) and stained with specific antibodies have formed one of the most commonly used tools for pathologists since the advent of monocolonal antibody production. Even today, the method remains as the basis of clinical pathology diagnostics. (Kumar et al., 2005) However, for research purposes, the process of making and staining tissue sections is laborious and time consuming, especially regarding high throughput screening of large volumes of sample specimens (Kononen et al., 1998).

Tissue microarray (TMA) technology is a recent invention aiming to alleviate the process of generating large amounts of representative immunohistological sample specimens from PEFF tissue blocks, and their subsequent screening for proteins of interest (most commonly disease-associated proteins). TMA technology uses novel sampling devices in order to construct an array of up to hundreds of tissue samples onto single microscope slide, which can then be stained in a similar manner as a standard immunohistological sample. Thus, a single object slide can contain a vast number of samples, allowing researchers to speed up the screening of marker proteins. In addition, multiple object slides can be constructed from a single array block, permitting the study of dozens of markers from a single set of tissue samples, while preserving most of the sample material for future use: single sample core extracted from a tissue block is typically less than one millimetre in diameter. Finally, new, computer-assisted image processing methods are being developed, aiming to further enhance the process and allow for faster and accurate analysis of large sample volumes. (Kononen et al., 1998; Kallioniemi et al., 2001)

In this study, the suitability of TMA technology for breast cancer research was evaluated in a medium-sized pathology laboratory environment. Breast cancer is an important condition that affects thousands of women annually especially in developed nations. Although breast cancer diagnostics and treatment have markedly improved during the past decades, it is still the cause of multiple deaths each year (see e.g. the Finnish Cancer Registry, 2007). In

pursuit for even more effective treatment and better understanding of the disease, vast numbers of cancer samples need be studied for new meaningful marker proteins and existing markers analysed for their effect on the outcome of the disease.

This kind of research can be conducted with relative ease with the aid of TMA technology. This permits even one researcher to construct and analyse a large number of samples. TMA technology ideally suits pathology laboratories, which commonly have comprehensive archives of PEFF tissue samples from cancer patients and which have well-established staining protocols and other equipments required for TMA studies, such as microtomes and automatic staining devices. The tissue microarrayer itself, of which many models are available for different needs, is the only acquisition such a laboratory has to perform in order to launch its own TMA project.

In this study, several well-established cancer markers were evaluated from a comprehensive set of breast carcinoma samples, and their prevalence was compared to values found in the literature in order to verify the suitability of TMA technology for clinical pathology laboratory.

2. TISSUE MICROARRAY TECHNOLOGY IN BREAST CANCER SUBTYPING

2.1 Breast cancer

2.1.1 General

Breast carcinoma is the most common cancer in the world among women. It is estimated that approximately one woman out of ten is affected by the disease. In Finland, around 4 000 cases are reported annually, and the numbers are increasing. In many Western countries the rates of incidence are fairly similar and also increasing, although variations exist – most likely due to genetic differences and environmental factors. Men are rarely affected by breast carcinoma: less than 1% of all breast carcinomas occur in men. (Finnish Cancer Registry, 2007)

Breast cancer incidence rates are very low among women below the age of 25. After this age the incidence rate starts to increase. For example, according to the Finnish Cancer Registry (2007), approximately one quarter of the breast cancers occur in the age group of 25 to 49 years, one quarter in women of 50 to 59 years of age, and the occurrence peaks in women over 60 years of age – circa half of the cases fall into this age category.

Although the number of cases is increasing, the overall mortality due to breast cancer has decreased. Mostly this is attributed to early detection and advanced treatment possibilities. In Finland, approximately 80% of the breast cancer patients are alive after a period of five years from the date of diagnosis. (Finnish Cancer Registry 2007; Rakha et al., 2007a) However, certain tumour subtypes have markedly poorer outcome, and currently no specific and effective treatment exists for some of these cancer subtypes.

The underlying causes of breast cancer are mostly yet to be elucidated. The origins of the disease are complex, but certain risk factors are known. For example, some genetically predisposed persons have an increased risk of developing a breast carcinoma; well-known examples of major genetic risks are inactivating mutations to the BRCA1 or BRCA2

genes, which lead to up to 80% life-time risk of breast or ovarian cancer (Schulz, 2005). In genetically predisposed families there are more breast cancer cases than in the average population, but genetic factors alone are not thought to be sufficient to develop the carcinoma (Antoniou et al., 2006; Cox et al., 2006). It is generally considered that cells must meet seven criteria before they turn into cancer cells, these criteria being: loss of apoptosis, genome instability, loss of growth inhibition, self-sufficient growth, angiogenesis, limitless replication, and tissue invasion (Kumar et al., 2005). Estimations of the significance of genetic background vary, but for example Schulz (2005) estimates that approximately 10 to 20% of breast cancer cases have a genetic family background. In these persons, mutations in certain genes already expose these persons to one or several of the required criteria and, therefore, less alterations in the cells are needed for cancer development (Kumar et al., 2005; Cox et al., 2006).

This notwithstanding, there are thought to be certain types of prevalent genetic polymorphism that cause smaller increase in the cancer risk to individual women, but are probably significant on the population level. Examples of these are variants of genes encoding steroid metabolising enzymes and some heterozygous mutations to genes that facilitate DNA repair (Schulz, 2005).

Since only a small proportion of breast cancers is distinctly tied to the family background, environmental factors also play a role. A wide range of environmental agents is regarded or suspected to increase the risk of developing breast cancer, some of which show a marked statistical correlation, and others that are not straightforward to prove. These include such factors as diet, environmental pollutants, hormones and hormonal treatment, obesity, and cigarette smoking. (Schulz, 2005; Claudio, 2006; Nagata et al., 2006; Tworoger et al., 2007)

Furthermore, there are no effective means of prophylaxis available, which makes early detection one of the key factors in fighting the disease – especially in women over 50 years of age the impact of mammography is profound (Finnish Cancer Registry 2007). Therefore, one common age group for mammographic screening is 50 to 60 year-olds, although the practices vary considerably both between and within countries. There are

differing opinions on the exact age group that should be screened by default, as well as on the other basis on which the screening should be performed (e.g. see Hakama et al., 1997; Immonen-Raiha et al., 2005; Trop and Deck, 2006)

Breast carcinomas are a heterogeneous group of tumours that are diverse in behaviour, outcome and response to therapeutic agents. Most breast cancers are adenocarcinomas, other types constitute less than 5% of the total (Kumar et al., 2005). Traditionally, anatomic and morphologic features have been the main ways to categorise breast cancers. There are 15 to 20 lobules in the breast, which are separated by connective tissue. The lobules are connected by ducts, which are lined with epithelial cells. The lobules and the ducts also form the two main categories of breast carcinoma; most cancers emerge from the ductal cells, and the cancer is termed ductal carcinoma. The lobular form is the other main group, and circa 10 to 15% of breast cancers are of lobular origin. Lobular cancers have somewhat bigger tendency to involve both breasts, when compared to other breast cancer types. (Weinberg, 2006) Ductal cancers often display a preceding form, *ductal carcinoma in situ* (DCIS), which is a state where malignant growth occupies the lumen of the ducts, but has not penetrated the basal cell layer or invaded the surrounding tissues. Lobular carcinoma also has an *in situ* form (LCIS), but it is not regarded as a cancer itself, but as a risk factor for cancer instead. Approximately one out of four women that display LCIS develops invasive carcinoma within 25 years. (Weinberg, 2006)

Other forms of breast cancer are less common. These include types such as tubular, medullary, mucinous, cribriform, and papillary carcinoma. Apart from the medullary form, these cancer types tend to have less relapses than the common ductal and lobular carcinomas. Medullar carcinoma, however, has the same prognosis as the common forms. Breast cancer can also have an inflammatory manifestation, which is more difficult to treat than non-inflammatory forms. (Kumar et al., 2005; Weinberg, 2006)

In addition to the above mentioned, morphology-based categories, there are various cellular, genetic, and molecular subtypes of breast carcinoma, which are of diagnostic and especially prognostic importance. These cancer subtypes often require special immunohistochemical staining methods, fluorescence *in situ* hybridisation or mRNA

profiling in order to be identified. There are two major categories based on the epithelial cell type of the cancer, which have clinical impact: basal and luminal carcinomas (Rakha et al., 2007a). Breast cancers also have hundreds of different genetic and molecular profiles, and their impact on the cancer development and treatment is in many cases still to be elucidated. These topics will be discussed in more detail in Chapter 2.1.4.

2.1.2 Diagnosis and prognosis

As stated in the previous chapter, early diagnosis is of paramount importance in treating all kinds of cancers. Large-scale mammography screenings along with women's awareness of the condition of their breasts by palpation are by far the most important factors regarding the early detection of breast carcinoma. In certain cases, the family history of breast cancer is also an important piece of knowledge that allows e.g. the mammography to be performed on younger age in order to ensure the earliest possible detection of the onset of the carcinoma.

Not only needs the diagnosis be prompt, but it is also essential that the correct subtype of the cancer be defined. This is crucial, for the treatment of choice also depends on the molecular profile of the cancer. Other important factors regarding the treatment are e.g. the lymph node status and the tumour size.

After the initial suspicion of cancer based on e.g. mammography and/or palpation – such as calciferous regions or dense mass in the radiograph, or a palpable lump – biopsy is often extracted from the site of the finding; biopsy is the only definite way to determine if cancer is present. Along with this, personal and family medical histories are evaluated, as well as general signs of health. A lump in the breast is often benign. For example cysts fall into this category. If cancer cells are present in the biopsy, however, further measures need to be undertaken.

Pathologic examination is performed to determine the primary carcinoma and axillary lymph node status. The tumour along with any compromised lymph nodes is removed

surgically and sent for histological analysis. In some cases where the cancer has invaded surrounding tissues, a mastectomy may need to be performed; this is determined by various factors, such as the size of tumour, the size of the breast, the patient's wishes, and the multifocality of the tumour.

2.1.3 Major prognostic factors

Diagnosis is formed with pathological examination and histological results. Based on these results, prognostic factors are determined and the patient is informed. There are varying practices and guidelines for prognosis in different countries, but in general, they follow the same principles. For example, the American Joint Committee on Cancer (AJCC) has created a widely used staging system that lists predictive factors that help to determine the likelihood of response to a particular therapeutic method. According to Kumar et al. (2005), there are two categories listed: major and minor prognostic factors, which are summarised below:

Major prognostic factors:

1) *Invasive carcinoma or* in situ *disease*. By definition, *in situ* carcinoma is confined to the ductal system and has not metastasised. Breast cancer deaths linked with ductal carcinoma *in situ* (DCIS) are due to the subsequent development of invasive carcinoma or areas of invasion that were left undetected at the time of diagnosis. However, majority of women with properly treated DCIS are cured. On the contrary, at least half of invasive carcinomas have developed metastases at the time of diagnosis.
2) *Distant metastases*. Once the cancer manifests with distant metastases, cure is unlikely. However, long-lasting remission can be attained, especially for patients with hormonally responsive tumours. Frequent sites of dissemination include the lungs, bones, liver, adrenals, and brains. The genetic profile of the primary carcinoma can be used to evaluate the risk of metastasis. The profile may be different among cancers displaying distant metastases and those that metastasise to

axillary lymph nodes. If the cancer cells express chemokine receptors, they are potentially attracted to organs excreting corresponding chemokines.

3) *Lymph node metastases*. Axillary lymph node status is the single most important prognostic factor in the absence of distant metastases. Generally, clinical assessment of nodal status is inaccurate, with false positive findings (e.g. palpable reactive nodes) and false negative findings (e.g. nodes with miniscule metastatic deposits) common. Thereby, biopsy is required for accurate prognosis. The lymph node status strongly correlates with the outcome of the disease; the more there are lymph nodes with metastases, the poorer is the outcome. Breast carcinomas may disseminate from one or two *sentinel nodes*. These are the first nodes in the close proximity of the primary cancer, and can be traced by coloured dye, radiotracer or both. Sentinel node biopsy can spare women from complete axillary node dissection and from the increased morbidity associated with the procedure. Macrometastases (> 0.2 cm) are of proven prognostic significance, and also novel methods of detecting micrometastases have given reason to initiate clinical trials in order to elucidate their prognostic impact.
4) *Tumour size*. The size of the primary carcinoma is the second most important factor concerning prognosis. It is independent of the lymph node status, although the risk of axillary lymph node metastasis increases with the size of the tumour. Women without nodal metastases and with the primary carcinoma under 1 cm in diameter have prognosis close to that of women without breast cancer. The 10-year survival rate for such women even without treatment is approximately 90%. However, over half of the women with tumour size over 2 cm in diameter have lymph node metastases, and many of these patients succumb to the disease.
5) *Locally advanced disease*. Breast cancers invading into skin or skeletal muscle are often connected with subsequent or concurrent distant disease. With improved detection methods of breast cancer, however, these cases have markedly decreased in frequency and are uncommon at initial presentation of the carcinoma.
6) *Inflammatory carcinoma*. Women that have breast swelling and skin thickening as a clinical manifestation have a particularly poor prognosis. The 3-year survival rate of these patients is merely from 3 to 10%.

The major prognostic factors mentioned above are used by the AJCC to divide breast carcinomas into following clinical stages (Kumar et al., 2005):

- Stage 0. Ductal carsinoma *in situ* or lobular carcinoma *in situ*. These patients have the 5-year survival rate of 92%.
- Stage I. Invasive carcinoma that has the tumour size of less than 2 cm in diameter (including carcinoma *in situ* with microinvasion) without lymph node metastases (or only metastases < 0.02 cm in diameter). The 5-year survival rate for patients in this category is 87%.
- Stage II. Invasive carcinoma of 5 cm in diameter or less with up to three involved axillary lymph nodes or invasive carcinoma larger than 5 cm without nodal involvement. The 5-year survival rate is 75%.
- Stage III. This stage includes several categories: Invasive carcinoma 5 cm or less in diameter with four or more metastasised axillary nodes; invasive carcinoma with larger than 5 cm tumour with nodal involvement; invasive carcinoma with 10 or more involved axillary nodes; invasive carcinoma with metastases in the ipsilateral internal mammary lymph nodes; or invasive carcinoma that manifests in the skin (oedema, ulceration, or satellite skin nodules), chest wall fixation, or clinical inflammatory carcinoma. The 5-year survival rate for this stage is 46%.
- Stage IV. All breast carcinomas with distant metastases. The 5-year survival rate: 13%.

In addition to major prognostic factors, there are multiple lesser or less-well determined factors that can be used to specify the prognosis – these are called minor prognostic factors. Women with nodal metastases and/or tumours over 1 cm in diameter benefit from some form of systemic therapy, and for these patients, minor prognostic factors can be used in order to decide among chemotherapeutic options and/or hormonal therapies (Kumar et al., 2005). Furthermore, for women without nodal involvement and small cancers, minor prognostic factors are employed to identify the patients most likely to benefit from systemic therapy and those who might not need additional treatment. These are discussed in more detail in the following chapter.

2.1.4 Subtypes and minor prognostic factors

Several well-established cancer subtype markers exist and these give valuable information regarding the cancer type and required treatment. Perhaps the most traditional markers that are routinely immunohistochemically visualised are the hormone receptors for oestrogen and progesterone (ER and PR, respectively, or hormone receptors [HR] collectively). Cancers displaying these receptors typically require the corresponding hormone in order to grow, and therefore therapeutic agents targeting these receptors can be used to effectively fight the cancer. For the same reasons, HR negative cancers may pose problems regarding the treatment options, as certain therapeutics (such as anti-estrogens relying on ER expression) are of no use. (Weinberg 2006) HR negative cancers are also more often poorly differentiated, of higher histological grade, have higher recurrence rates, and a decreased overall survival (Rakha et al., 2007a).

Furthermore, certain other antibodies for cancer markers are widely employed in the cancer typing process. Examples of markers that are often targeted in immunological staining are HER-2 (also known as C-erb-B2), androgen receptor, epidermal growth factor receptor (EGFR), cadherins P and E, transcription factor p53, as well as basal cytokeratine kinases (CK) 5, 6 and 14 (Rakha et al., 2007a).

Finally, chromosomal analysis can be used in order to categorize cancer types. For example, chromogenic *in situ* hybridisation (CISH) can be employed to detect chromosomal aberrations. One such DNA-level target is the C-erb-B2-gene, which encodes the protein HER-2 (Vocaturo et al., 2006). Duplication of this gene causes elevated levels of HER-2 expression, which in turn is a sign of an aggressive carcinoma (Rakha et al., 2007a).

The American Joint Committee on Cancer (AJCC) has determined various minor prognostic factors that can be used in order to determine the optimal treatment for each patient, which include the following, summarised from Kumar et al. (2005):

1. *Histologic subtypes*. The 30-year survival rate of women with special types of invasive cancers (such as tubular, mucinous, medullary, lobular, and papillary) is greater than 60%, as opposed to less than 20% for women with carcinomas of no special type.
2. *Tumour grade*. The most frequently used grading system to evaluate the degree of cancer differentiation (*Scarff Bloom Richardson*) combines nuclear grade, tubule formation, and mitotic rate. The effect of differentiation, or lack of it, is profound: 85% of women with well-differentiated grade I carcinoma, 60% of women with moderately differentiated grade II carcinoma, and 15% of women with poorly differentiated grade III carcinoma survive for 10 years.
3. *Oestrogen and progesterone receptors*. Currently immunohistochemistry is employed to visualise these receptors in the cancer cell nucleus. Approximately 50 - 85% of breast carcinomas express oestrogen receptors (ER), with some variation among different age groups: ER positive tumours are more common in post-menopausal women. In general, women with hormone receptor positive carcinomas have a somewhat better prognosis than women with hormone receptor negative cancers. The assessing of hormone receptor status is important when predicting the response of the cancer to therapy. Eighty per cent of ER and progesterone receptor (PR) positive carcinomas respond to hormonal therapeutics, whereas only 40% of either ER+/PR- or ER-/PR+ cancers respond to these kinds of treatments. Finally, ER-/PR- tumours have less than 10% likelihood of positive response.
4. *HER-2/neu*. HER-2, (human epidermal growth factor receptor 2, also known as *C-erb-B2* and *neu*) a transmembrane glycoprotein that has a role in controlling cell growth. It is thought to act as a co-receptor for various growth factors. Approximately 20% to 30% of breast cancers over-express HER-2, and of these cases, an estimated 90% are due to an amplification of the gene in 17q21, which can be detected either by evaluating protein content with immunohistochemistry or by determining gene copy number by chromosomal analyses, such as FISH. Many studies indicate that HER-2 is associated with poor prognosis, and some recent studies also address the possibility of these tumours responding differently to certain therapeutic agents, such anthracycline chemotherapeutics or hormonal treatment. However, the assessment of the HER-2 status is most important in order

to define the likely response to therapy targeted specifically to this protein, e.g. with monoclonal antibody trastuzumab (trade name Herceptin®).

5. *Lymphovascular invasion (LVI)*. Cancer cells can be discovered from the vascular spaces, i.e. lymphatics or small capillaries, surrounding tumours. This finding strongly correlates with lymph node metastases, and therefore indicates poor prognosis, even despite possible negative lymph node findings. Furthermore, tumour cells in lymphatics of the dermis are strongly associated with the clinical manifestation of inflammatory cancer, and thus bodes a very poor prognosis.
6. *Proliferation rate*. There are several ways to measure the proliferation of cancer cells, such as flow cytometry, mitotic counts, thymidine labelling index or immunohistochemical detection of cellular proteins produced during the cell cycle. Of these, mitotic counts are included as part of the standard grading system. The presence of certain proteins of the cell cycle, such as both of the cyclin E isoforms expressed concurrently, may indicate elevated chances of survival. Generally, fast proliferation rates are often associated with worse prognosis.
7. *DNA content*. The amount of DNA per cancer cell can be discovered by image analysis of tissue sections or by flow sytometric measures. Carcinomas that display a DNA index of 1 have the similar total amount of DNA as normal diploid cells. This does not, however, exclude the possibility of marked karyotypic changes. Cancers with abnormal DNA indices are called aneuploid, and they have a slightly worse prognosis.

In addition to these factors that are already in use, there are several promising new options, such as gene expression profiling (Kumar et al., 2005; Gruvberger-Saal et al., 2006). For example, a set of 70 marker genes with known prognosis signatures, either good or poor, can be analysed for changes. Their effect on the prognosis is then calculated and an overall prognosis is determined. This method has given more precise prognosis than conventional methods, but it requires rapidly frozen tissues and therefore is not suitable for assessing most breast cancers. Notwithstanding, the method shows considerable promise, and techniques allowing the use of paraffin-embedded tissue samples are being developed. (Kumar et al., 2005)

Furthermore, mRNA studies can indicate the expression of cancer-specific proteins. mRNA studies are usually done with the aid of reverse transcriptase (RT) PCR. RT-PCR - amplified DNA sequences are then subjected to electrophoresis, which allows medical personnel to resolve the expression of certain marker proteins. There are also some more experimental methods of subtyping breast cancer, e.g. the analysis of DNA methylation status and assessment of telomere and telomerase activity. In brief, the former method indicates the presence of specific changes to the DNA structure or genomic modifications: hypermethylation within the promoter regions of selected genes is frequently associated with inactivation of involved gene(s) and may be an early event in certain cancer subtypes. The latter method, on the other hand, is used for detection of the maintenance of telomere length and elevated telomerase enzyme activity, which are though to be required from cancer cells in order to maintain their ability to divide indefinitely. (Boultwood and Fidler, 2002)

Finally, tissue microarray (TMA) technology can be employed to study large volumes of cancer samples in order to conduct population-level studies on breast cancer subtypes and their role as different prognostic factors. The method allows the use of archival paraffin-embedded, formalin-fixed tissue samples for protein-level or nucleic acid analyses, and these results, in turn, can be compared to patient diagnostic and follow-up data. The features and uses of TMA technology are discussed in more detail in Chapter 2.2.

2.1.5 Current therapeutics and treatment possibilities

Modern-day breast cancer treatment aims for the complete cure of the patient. However, in some cases with very poor prognosis, a positive outcome of the therapy is unlikely, and in these cases a long-lasting remission is often the best result that can be achieved. Regardless of these unfortunate situations, most breast cancer patients can be cured. For example, in Finland approximately 80% of breast cancer patients are alive after a 5-year follow-up period (Finnish Cancer Registry, 2007). Therefore, a diagnosis of breast cancer is not a "death sentence". Most cancers have been present for several years prior to detection, and even for disseminated cases the survival time can be counted in years (Kumar et al., 2005).

According to Kumar et al. (2005) the treatment(s) of choice are based on initial clinical findings, as well as on the major and minor prognostic factors listed above. Current therapeutic approaches include local and regional control, using surgery (breast conservation or mastectomy) and post-operative radiotherapy in combination. In addition, systemic control is conducted with hormonal therapeutic agents or chemotherapy. Examination of sentinel lymph nodes or dissection of axillary nodes is often performed for prognostic uses, but the axilla can also be treated by radiation alone. Furthermore, negative sentinel lymph node status strongly correlates with negative axillary nodes, therefore reduces the need for removal of the majority of axillary contents, and thus lowers the overall morbidity associated with the treatment (Mansfield et al., 2007).

Newer therapeutic methods include e.g. inhibition of membrane-bound growth factor receptors (such as HER-2/neu), stromal proteases, and angiogenesis (Kumar et al., 2005). Pharmacologic agents or specific antibodies are used for these purposes. For example, trastuzumab (trade name Herceptin®) is a rather well-known therapeutic agent composed of humanized monocolonal antibody, which binds to the HER-2 protein and, therefore, inhibits the proliferation of carcinoma cells (Glick and Pasternak, 2003; Schulz, 2005).

There are also various other novel therapeutic agents that utilise the knowledge of molecular profiling of the cancer. Drugs that act as chemopreventive agents are divided into two major groups: those that prevent oestrogen receptor (ER) positive breast cancers and those that target ER negative cancers. According to Colozza et al. (2006) and Macaskill et al. (2006), the former group consists of such therapeutic agents as selective oestrogen receptor modulators (SERM), aromatase inhibitors, gonadotrophin-releasing hormone (GnRH) agonists, and phytoestrogens, whereas the latter group includes therapeutics such as cyclo-oxygenase-2 (COX-2) inhibitors, retinoids, statins, telomerase inhibitors and the above-mentioned HER-2 antibodies. Of the therapeutic types listed above, the aromatase inhibitor letrozole (trade name Femera®) warrants mentioning: this new, third-generation therapeutic agent has proved to be superior to tamoxifen, a popular therapy of choice for patients unable to undergo surgery (Macaskill et al., 2006). However, some studies point out that the optimal modality of administration is still unclear for some of these modern aromatase inhibitors (Colozza et al., 2006).

As a conclusion, it can be said that traditional methods of cancer therapy are still widely used and often effective, but there are various scenarios where they are inadequate – especially in cases of aggressive, disseminated carcinomas that have metastasised to other organs. Even today, these cases usually remain incurable, although long-lasting remissions can sometimes be achieved. This notwithstanding, the increased knowledge of molecular basis of cancer has caused the emergence of novel therapeutic agents that allow more specific and effective ways to treat various cancer types. To effectively utilise the possibilities of these new therapeutics, the understanding of various breast cancer subtypes and their effect on prognosis and the cancer behaviour is even more essential than before.

2.2 Tissue microarray technology

2.2.1 General

Large-scale, genome-level analysis techniques, such as cDNA microarrays, serial analysis of gene expression (SAGE) and proteomics surveys have made it possible to analyse the expression levels of thousands of genes and proteins at a time. Nowadays, biomedical research widely uses these types of high-throughput screening methods in both basic research, and industrial-scale efforts in order to identify targets for new diagnostic and therapeutic products. This requires the ability to select the optimal targets from the population of tens of thousands of candidate molecules. Analysis of these molecules *in situ* – their expression across all the tissues and their clinical features – would provide useful insights regarding their potential as future diagnostics or therapeutics. (Kononen et al., 1998; Kallioniemi et al., 2001)

However, when compared to the high-throughput methods of proteomics and genetics, conventional tissue-based molecular methods tend to be slow and laborious. In addition, only about 300 five micrometer sections can be prepared from an average-sized paraffin-embedded tissue sample for e.g. immunohistochemical staining or fluorescence *in situ* hybridisation. This amount of samples represents less than 1.5% of all the estimated ~20 000 to 25 000 protein-coding genes in the human genome (IHGSC, 2004). In this light,

genome-scale analyses are not possible with traditional histopathological methods. Furthermore, large volumes of histological samples are difficult and time consuming to process even, if there was a small number of molecules of interest. They also consume substantial amounts of reagents, of which some are very costly (e.g. monoclonal antibodies for immunohistochemistry). (Kallioniemi et al., 2001). These facts together have brought forth the need for faster and more cost-effective ways to analyse large volumes of histopathological samples.

Tissue microarray (TMA) technology is a method where up to one thousand separate tissue cores from typically paraffin-embedded tissue samples (donor blocks) can be arranged as an array to a recipient paraffin block. This allows simultaneous analysis of a vast number of proteins or nucleic acids of interest: a single TMA section can give information of up to one thousand tissue specimens with one staining or hybridisation procedure. One TMA paraffin block can yield up to three hundred of such sections, allowing e.g. three hundred different antigens to be stained, each from up to a thousand samples. Furthermore, multiple replicate blocks can be constructed by taking several samples from all donor blocks, and positioning them identically in all TMAs. However, in practice, the number of tissue cores on a single TMA block is often smaller, ranging from approximately 400 samples to 800, depending on the experience and purposes of the user, as well as on the size of the block used and diameter of the core extracted. (Kallioniemi et al., 2001)

A predecessor for TMAs was a multitumour tissue block (MTTB) technology, first introduced by H. Battifora in 1986. The method allowed the analysis of up to one hundred neoplasm or other tumour samples with immunohistochemical methods. The current technique was developed by J. Kononen et al. in 1998, and it used a novel and precise approach in construction of arrays with multiple tissue samples. (Battifora, 1986; Kononen et al., 1998)

2.2.2 Constructing a tissue microarray

TMA technology uses a hollow needle to extract tissue cores typically from formalin-fixed paraffin-embedded (FFPE) tissue samples, although acetone-fixed frozen (AFF) or glycol methacrylate resin embedded (GMRE) tissue samples can also be used. Generally, FFPE offers good morphology and allows the use of archival samples, whereas AFF provides good antigenicity. Morover, GMRE is a recent method that combines these two advantages. (Howat et al., 2005). The most widely used core size is 0.6 mm in diameter, but core sizes up to 2 mm have been used. Naturally, the increase in core size decreases the amount of samples that can be placed on a single TMA, and generally it is regarded that larger sample amounts offer more benefits along with better statistical possibilities than what increased sample size would bring with improved representativeness of each sample. Therefore, core sizes as small as 0.4 mm have been mentioned. (Kallioniemi et al., 2001) The tissue cores from selected sample material are transferred to a recipient paraffin block, where they are arranged next to one another in an array-like manner. For example approximately 300 to 350 of 0.6 mm samples with 1 mm spacing can be placed on a single 20 x 25 mm paraffin block. In practice, the amount is often slightly smaller due to need of increased readability of the TMAs (see Image 1 for an example). TMAs can be constructed by using manual arrayers, such as the Beecher Instruments Inc. (Sun Prairie, USA) MTA series (Image 2), or automatic tissue arrayers.

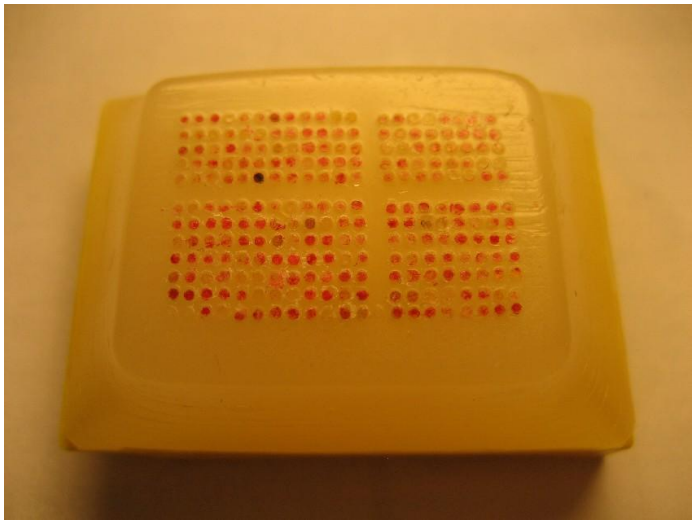


Image 1: an FFPE block containing approximately 250 embedded sample cores. An empty row and a column have been left on the block for improved readability of and navigation on the subsequent TMA object slides created from the block. The red colour of the sample cores is from the marker used to indicate the location of cancer tissue, and is removed during the ensuing microtome trimming before the object slides are constructed.

Image 2: a manual tissue arrayer (Beecher Instruments MTA II, picture courtesy of Beecher Instruments Inc.)

However, most of the time and effort in the construction of a TMA is spent in the review of the pathological samples, as well as in organisation and processing of the samples that

are selected for the array. The representativeness of the samples is verified by making a fresh haematoxylin and eosin (HE) slide from each of the sample blocks. The HE slides are used to locate the area of interest in the donor block. In addition to histological representativeness of the sample by terms of sufficient numbers of target cells, the actual size of the sample is also important: the donor block should be 1 mm thick at minimum, and preferably 3 to 4 mm. Archival samples dating back 20 to 40 years are normally sufficient if they have been fixed in buffered 4% formalin. These samples are valid e.g. for immunohistochemistry and DNA fluorescence *in situ* hybridisation (FISH). However, mRNA *in situ* hybridisation may prove to be challenging due to the degradation and cross-linking of RNA molecules by formalin fixation. (Kallioniemi et al., 2001; Simon et al., 2005)

Tissue microarrays are then subjected to microtome sectioning; five micrometer sections that are commonly used in immunohistochemistry are cut from the TMA block. The resulting sections are transferred on an object slide by common histochemical protocols. Further manipulation of the slides is performed similarly to other standard histochemical procedures. After appropriate pre-treatment, the TMA slides are ready e.g. for HE or immunological staining (see Image 3), or fluorescence *in situ* hybridisation, where they provide an insight into hundreds of samples, while consuming the reagents approximately equal to only one sample preparation by conventional methods. Similarly, there is no need to handle and analyse hundreds of separate slides, because all the samples are available in a single slide. TMAs also provide information on the cellular origin and location of the target molecules, thus supplementing the information available from gene expression studies. If multiple microtome sections are to be made from the TMA block, it is advisable to stain a new HE slide approximately after every 50 sections in order to eliminate errors caused by the possible change of morphology within the tissue sample. (Kallioniemi et al., 2001; Simon et al., 2005)

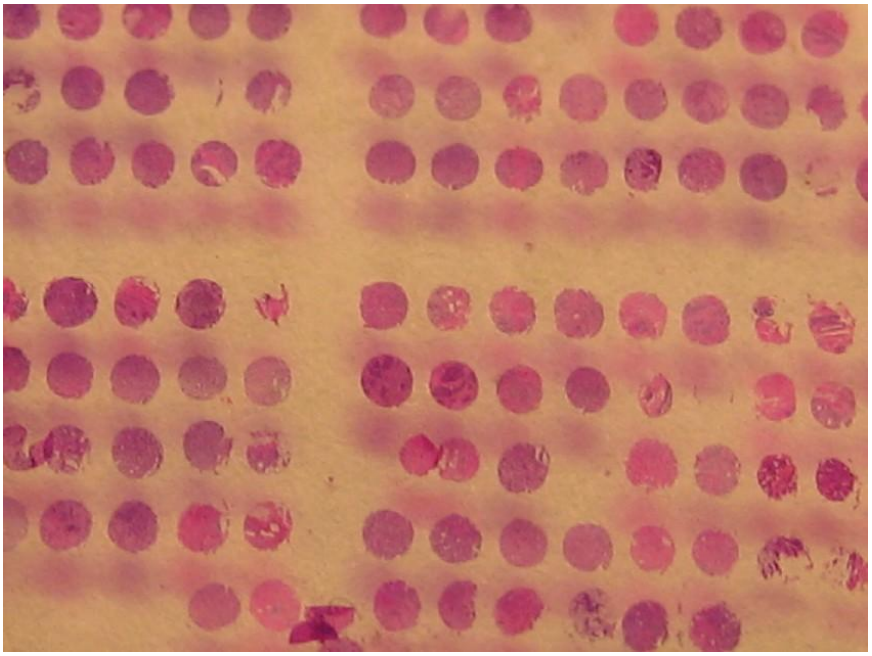


Image 3: H&E-stained TMA object slide. The image displays a portion of the paraffin block seen in the Image XX, now attached on an object slide and stained with H&E. Each spot covers an individual sample specimen.

2.2.3 Molecular analysis methods

Molecular detection methods developed for whole tissue sections are generally adaptable for TMA sections, although the nature of the tissue fixation may be a limiting factor. The most common application of TMA technology is the immunohistochemical (IHC) detection of cellular antigens. IHC TMA slides can be used to construct e.g. a panel of different tissues, tumours, or cell lines, which in turn offer tools to evaluate and optimise pre-treatment conditions, antibody concentrations, and detection systems. Such TMA slides are also suitable for actual molecular studies. Because all the target tissues are on the same slide, the reliability and reproducibility of the staining procedure are improved – every sample has been subjected to similar staining conditions (i.e. pre-treatment methods, incubation times, stain/antibody concentrations etc.). Uniform staining conditions help to minimise any variations caused by the staining protocol. (Kallioniemi et al., 2001)

The screening method of choice depends on various factors, such as the sample amount and available equipment. Often, TMA screening is performed by a scientist or a

pathologist manually by using a microscope. The results, such as positive antibody binding or fluorescent signal, are scored and recorded for statistical analysis. For large sample volumes that often result from TMA experiments, automated screening instruments are under development. These rely on fluorescent signal detection, or on sophisticated computer algorithms that take into account the staining intensity and cellular morphology (such as the Beecher Instruments TMAx analysis software). However, like the TMA construction automation, also the image analysis software is currently under intense development, and improved and more efficient applications are expected from several manufacturers (Conway et al., 2006). Furthermore, a sample tracking system – which can also be used along manual TMA methods – is available. It utilises the tissue core coordinates, which can be linked to the corresponding patient in a database containing patient demographic, diagnostic, and survival-rate entries, allowing the comparison of molecule-level results with clinical information. (Kallioniemi et al., 2001)

2.2.4 Tissue microarray technology in cancer profiling

The first study utilising the modern TMA technology by Kononen et al. (1998) targeted gene amplifications, as well as cellular tumour antigen p53 (a regulatory transcription factor) and oestrogen receptor in breast cancer samples. Six gene amplifications were detected in this study, along with the p53 and oestrogen receptor expression. From thereon, TMA technology has strongly established itself as a widely-used and reliable method in cancer research along with other high-throughput technologies, such as cDNA microarrays. Whereas cDNA microarrays allow for determining expression levels of different genes and their likely relationship to healthy and cancerous tissue, and thus make it possible to categorise cancer subtypes (Kapp et al., 2006), they do not permit the accurate analysis of cellular locations of the proteins the target genes encode (Kallioniemi et al., 2001). TMA technology facilitates a powerful tool for studying different proteins of interest on the cellular level. Immunohistochemical staining can be effectively used to pinpoint the cellular location of either known carcinoma markers, or new candidate proteins. Furthermore, with TMA technology it is possible to identify subsets of cells and cell

populations within a tissue specimen, thus giving an insight to features that are not morphologically distinguishable.

One of these features is the origin of cancer cells. From a histological sample, it is not always possible to determine the origin of neoplastic cells. However, the origin of the tumour is often of prognostic importance, for carcinomas emerging from different tissues may have different treatment possibilities and prognostic factors (DePalma et al., 2006). Furthermore, cancer cells from the same tissue can also have distinctly dissimilar properties between patients. A good example of this is the receptor status of breast carcinoma: oestrogen and progesterone receptor (ER and PR, respectively) status determines whether the cancer responds to certain receptor-specific therapeutics. Negative ER and PR status also correlates to higher mortality rate, partially due to the lack of response to these receptor-linked therapeutic agents. (Carey et al., 2006; Dunnwald et al., 2007; Rakha et al., 2007a)

If an ER and PR negative cancer also lacks the HER-2 (member of the epidermal growth factor receptor family), it is termed "triple-negative" (Rakha et al., 2007a), which further reduces the responsiveness of the cancer to common therapeutic agents – even though HER-2 negativity in itself indicates a better outcome when compared to HER-2 positive disease. Finally, the ethnic background, age, and hormonal status of the patient have an effect on the typical cancer profile in a given population; e.g. the basal-like breast cancer subtype is distinctly more prevalent among pre-menopausal African American women than post-menopausal women in the same ethnic group (Carey et al., 2006).

The distinct heterogeneity of breast cancer subtypes has caused the emergence of various ways to categorize these tumours, which are discussed in Chapter 2.1. In clinical practice, the conventional methods of histology and immunological staining are well-established and standard practice. A pathologist's evaluation of a traditional microscopic slide made from a tissue section is needed for a reliable diagnosis. However, when studies are conducted on the population level and sample numbers are much higher, tissue microarray technology offers a rapid subtyping method with a possibility of extensive statistical analysis.

In immunohistochemical studies, ER/PR status along with the HER-2 are typically included in many TMA staining, so as to give an overall insight to the sample population, and to act as a means of quality control. The actual study can then target proteins of interest, whose expression can be compared with existing tumour type and patient data, e.g. comparison with patient survival rate or tumour grade can be made against the expression levels of the target protein. Similarly, genetic amplifications can be analysed from a large sample population by targeting the DNA, and the results are then compared to statistics. (Kallioniemi et al., 2001)

2.2.5 The future of tissue microarrays

The most time-consuming processes in traditional tissue microarray techniques are the screening of the initial sample specimens for cancer locations, the assembly of the TMA paraffin block prior to the sectioning on the object slides and the subsequent manual microscopy in order to score the stained microscopy slides. Whilst the sectioning can be done with a relative ease with a microtome, and the staining is readily automated in most pathology laboratories, these three stages have this far been fairly tedious and time-consuming, although some automatic arrayers and sample tracking systems already exist (Conway et al., 2006)

Several new innovations have now emerged to hasten these steps. The first major area of development is computer-assisted TMA automation, which allows construction of multiple TMA blocks in a fast, accurate and reliable manner. With automatic arrayers, up to 180 sample cores can be transferred per hour in up to 27 simultaneous blocks; a marked increase in output if compared to an estimated 30 to 70 cores per hours placed manually in up to 4 replicate blocks (Beecher Instruments homepage, 2008). Automatic arrayers also eliminate the need of hand-written notes for sample locations and specimen tracking. Furthermore, the selected sample regions can be marked on-screen, instead of manually labelling each paraffin block. Finally, the quality of the TMA is more even, for the differences between individuals can be largely negated by automation.

The second important future concept is image analysis software. With carefully designed computer algorithms linked to sophisticated imaging tools, analysis software has the potential to greatly reduce the time required for screening the object slides. An example of such a system is the Virtual Tissue Matrix (VTM) by Conway et al. (2006). Currently, several manufacturers also provide imaging software for various needs. For example, TMAx analysis software, a Beecher Instruments application built on object-oriented image engine from Definiens AG (Munich, Germany), provides screening tools for various IHC-stained TMA slides, recognising e.g. malignant cells from a heterogeneous tissue sample. For protein expression analysis, Ariol TMA module from Applied Imaging Corporation (San Jose, CA, USA) provides high-throughput screening options for researchers. In addition to the examples mentioned above, the increasing markets of image analysis software attract increasing numbers of other companies and research facilities. (Conway et al., 2006; AI Corporation homepage 2008; Beecher Instruments homepage 2008; Definiens Corporation homepage 2008)

The constant evolution of the arrayer machines and automatic image processing is likely to markedly enhance the efficiency of the TMA technology and make it possible to analyse larger sample amounts and more comprehensive cancer marker panels. Naturally, other research areas utilising immunohistochemistry can take advantage of the technology.

2.2.6 Limitations of tissue microarray technology

Tissue microarray technology has many advantages over conventional histological methods, such as faster sample throughput, lessened manual interaction, and reduced need of chemicals and reagents. However, there are certain limitations that concern this technology, which reduce its usefulness in some scenarios. One area in particular, where traditional histology still surpasses most TMA applications, is patient diagnostics. The small sample area renders TMA technology prone to false negative results, because e.g. tumour biopsies are often heterogeneous, and there is a chance that despite careful selection of the sample region, the sample core may still be devoid of cancer cells.

Therefore, even though statistical analysis of large sample populations can easily overcome isolated cases of negative results, i.e. negative samples can be completely excluded from the sample population and only positive samples are analysed, patient diagnostics cannot afford false negative findings. For the same reason, TMA technology is rarely used to discover tumour prevalence in an unknown sample population, but is instead used to evaluate a large number of samples that are already known to have carcinoma. This, in turn, allows researchers to discard any negative results, and concentrate on the properties of the cancer itself. (Kallioniemi et al., 2001; Gancberg et al., 2002)

When constructing large arrays, a certain portion of the samples are usually defective, *i.e.* they are devoid of the desired cell type. This ratio of unrepresentative samples mostly depends on the experience of the array constructor and the size of the sampling needle. Also the type of tissue and abundance of the target cell type have an effect. Even among experienced persons, there are invalid samples in the array, usually around 20 to 30% of the total sample volume (Hager et al., 2007). This should be kept in mind when considering the suitable amount of samples in the study. There is some contradiction about this, however. For example, Gancberg et al. (2007) found TMA technology to be a reliable method in FISH studies, discovering only 3.4% unrepresentative samples.

Use of duplicate samples can help to reduce the risk of having some tumour being excluded from the study due to lack of desired cells, but there are differing opinions whether this approach introduces more positive or negative aspects to the overall results, as taking two sample cores from each tumour increases the workload and lessens the overall amount of tumours represented in a given TMA. (see e.g. Kallioniemi et al., 2001; Hager et al., 2007)

Furthermore, certain tissue types tend to be more challenging to arrange properly to a TMA. Examples of these are certain tissues that typically contain a considerable amount of adipose tissue, such as the breast. Also calciferous or collagenous tissues may prove to be difficult to process due to their hard or tenacious nature. The problems concerning adipose tissue usually relate to its poor attachment to the object slide, as well as to its tendency to dissolve during fixation and other treatments. Hard materials, on the other hand, may prove

to be challenging to extract from the donor tissue sample, and may be difficult to section with a microtome. Finally, necrotic areas in the sample specimen may also yield negative results. (Gancberg et al., 2002; Hager et al., 2007)

Finally, it is generally regarded that RNA studies are more challenging to conduct than those that target proteins or the DNA. RNA is an easily degraded macromolecule, and therefore common methods of fixation and preservation of tissue samples have a tendency to denature and cross-link RNA. Therefore cryosections or alternative fixation methods may need to be used for RNA studies instead of common formalin-fixed paraffin-embedded (FFPE) tissue samples. As a more stable molecule, DNA can also be targeted in FFPE samples. Similarly, proteins are well preserved in these types of samples, and archival specimens dating back to several decades can be usually used without problems. (Kallioniemi et al., 2001)

3. AIM OF THE STUDY

This study is part of breast cancer research conducted in the Central Hospital of Central Finland, Department of Pathology led by Dr. Teijo Kuopio. All the research was conducted under the supervision of Dr. Kuopio, in association with other medical personnel. The tissue microarrayer was kindly provided by Beecher Instruments. Dr. Juha Kononen from Beecher Instuments also provided specialist support regarding the arrayer and its use.

The Department of Pathology has a comprehensive archive of breast cancer samples, linked with detailed patient diagnostic and follow-up data. Therefore, it is important to devise a functional and cost-effective way to conduct research on these samples, in order to improve our understanding in cancer and its treatment. The Department of Pathology required a pilot study on the function of the arrayer and its suitability for the medical laboratory environment. Furthermore, the results from the study, if sufficient, are to be used to discover diagnostic or prognostic patterns when compared to the patient data.

RESEARCH MISSION:

The purpose of this study is to determine the suitability of tissue microarray technology for a middle-sized pathology laboratory, or a similar research group in an academic organization. In addition, the ensuing results are to be used to discover meaningful links between the prevalence of various cancer markers in breast cancer samples, and their respective patient data.

- A comprehensive tissue microarray will be created with a manual arrayer, using archived formalin-fixed, paraffin-embedded breast cancer samples.

- The array is stained with several important cancer markers, and the results are compared to the recent results published by well-established research groups in order to evaluate the method as it compares to other histopathological technologies, such as traditional immunohistochemistry, and its suitability for use in a pathological

laboratory for research purposes, when limited personnel resources can be allocated for the study.

- Finally, should the results be adequate and in line with results obtained by conventional methods, they are compared to patient diagnostic and follow-up data in order to discover possible links between expression of these markers, and outcome of the cancer.

4. IMPLEMENTATION OF THE STUDY

The permission for the use of patient samples was granted by the Ethical Committee of Jyväskylä Central Hospital. In this study, no laboratory animals were used.

4.1 Materials and methods

4.1.1 Research methodology

This study is based on quantitative research methodology, comprising a large sample repertoire characteristic of this research method. Furthermore, the objective view of the study, as well as the possibility to reliably repeat the experiment, support the quantitative approach. On the other hand, however, there are several qualitative aspects in the study, especially in stages where evaluation of cancer regions, and their respective analysis after immunohistochemical staining and tissue microarray construction took place; in these stages each sample was analysed individually for factors such as cellular morphology and staining patterns and intensity.

In addition, statistics play a role in this study, which is also typical for quantitative research methodology. High sample numbers and statistical analyses not only make the study itself more reliable, but increase the validity as well. (Creswell, 2003) Furthermore, the results are readily comparable to possible future studies as long as the sample population is kept similar. This notwithstanding, different sample populations (such as dissimilar age groups) can also be compared, and differences between them can be evaluated.

4.1.2 Ethical aspects

Ethics play a major part in all scientific and medical research. Ethical issues cover various topics, such as patient rights, use of patient samples, use and handling of laboratory animals, truthful reporting of results, and objective data analysis (Gregory, 2003).

This study was conducted according to the guidelines of research conducted with patient data, including the regulations concerning the use of patient samples. Results of the study remain archived and can be reviewed later if necessary. All patient data was handled anonymously and confidently. The study did not include experiments with laboratory animals, and therefore ethical topics regarding this are not subjected to further discussion.

An important factor to consider when conducting medical research is what the promises and expectations concerning the study are; there should not be false hopes and impressions about the study among the possible subjects of potential future treatments. It should be stressed that the way from any base-level research to practical pharmaceuticals or treatments takes years to develop, and not all projects ever give the desired results (Gregory 2003).

4.1.3 Sample specimens

Three hundred and sixty-nine patient samples were used in duplicate in this study. The samples were acquired from previously dissected cancer tumours from women that had undergone cancer treatment during the eighties and the nineties. The paraffin-embedded, formalin-fixed samples were chosen by the following criteria: 1) the patient was diagnosed with invasive breast carcinoma, 2) there was an existing haematoxylin-eosin stained histological section, allowing for the precise location of cancer cells to be found, 3) the patient follow-up data existed, and 4) the patient, during the time of diagnosis, belonged to the screening age group of Central Finland Healthcare District, e.g. ages from 50 to 60 years. The samples were collected between the years 1987 - 1997.

Later during the study, during the assembly of tissue microarrays, factors such as the overall amount of the malignant, invasive cells, the depth of the embedded tissue, and the overall quality of the tissue block were also noted. Samples that were unlikely to yield a representative sample were excluded in this stage in order to acquire as many valid samples as possible, despite the fact that tissue microarray technology is normally

associated with a certain percentage of samples being devoid of desired cell type nonetheless (Kallioniemi et al., 2001).

4.1.4 Rationale of the selection of sample specimens

The patient cohort consists of post-menopausal women (aged between 50 and 60 years), who are an important group among cancer patients for several reasons. Women in this category are, for example, in the age group where cancer incidence rate starts to increase markedly. Furthermore, when viewed on the population level, they are able to benefit from the treatment, and the outcome is generally good. Most of the women are still within active work force and, therefore, successful treatment of cancer is especially beneficial also on the economical level. This notwithstanding, from the humane perspective, it is equally important to study cancer in as wide age range as possible in order to help all cancer patients regardless of their age or any other such factor. However, focusing the study on this particular age group not only enables the optimal use of limited resources, but also offers valid knowledge concerning many cancer cases spanning outside the age group of the patients in this study.

Finally, the availability of suitable sample material played a distinctive role in what comes to the selection of sample specimens: women of this age group are routinely screened, and therefore the amount of cancer findings, and as a result, the number of available sample specimens is increased, thus allowing the conduction of a quantitative study.

The samples were collected between the years 1987 - 1997 and, therefore, they were old enough to have no clinical value regarding the treatment of the patient. However, the tissue microarray technology is a conservative method concerning the sample amounts, and thus allows a vast majority of the sample specimen to be preserved.

4.1.5 Quality control

The quality control procedures for this study were the same ones used in clinical practice, and thus fulfilled a very strict set of criteria, including external quality control specimens, protocol and equipment validation and audition, and daily quality control runs. Medically qualified professional personnel performed or supervised the performing of all phases of the study from sample selection to immunohistological staining and data analysis.

4.2 Tissue microarray

4.2.1 Tissue microarray construction

The tissue microarrays (TMAs) were constructed with Beecher Instruments MTA-1 manual arrayer (Beecher Instruments Inc., Sun Prairie, USA) on empty 20 x 25 mm paraffin blocks. A total of three TMAs were constructed, with the following sample amounts:

- Array I: 121 samples in duplicate (242 sample cores)
- Array II: 118 samples in duplicate (236 sample cores)
- Array III: 128 samples in duplicate (256 sample cores)

- Total number of 367 samples in duplicate (734 sample cores)

The sample positions were recorded on a Microsoft Excel table as well as manually on paper sheet. Each TMA had an empty line running horizontally and vertically through the array, dividing every construct in four separate, asymmetric sections. This was done in order to ensure that the correct alignment of the array was always possible to notice, as well as to make the screening of the arrays less demanding due to easier navigation along the lines and columns as the presence of empty lines permits the easier tracking of line and column numbers. Furthermore, each TMA had an additional asymmetric sample position to mark the first sample of the array in the lower left-hand corner of each TMA.

After the embedding of the sample cores, TMAs were subjected to one-hour incubation in +60 °C so as to allow the samples to properly merge into the blocks. Incubated TMAs were then ready for microtome sectioning. Standard 5-micrometer sections were cut with a sliding microtome equipped with a steel blade and mounted on glass slides. From each of the TMAs, up to ten sections were cut and the best were chosen for staining.

Finally, the TMAs, now mounted on object slides, were stained with either haematoxyline-eosin (HE) or with one of the specific antibodies against cancer markers (see Chapters 2.1.2 - 2.1.5). An Autostainer 480 automatic staining device (LabVision Corporation, USA) was employed for this purpose. The staining protocols for each antibody were the same as used for clinical samples.

4.2.2 Specifications of the cancer markers used in the study

The following chapter lists the main specifications of the antibodies used in this study, as provided by the manufacturers. The complete datasheet of a given antibody can be obtained from the Internet from their respective manufacturer's home page.

Oestrogen receptor

1) Monocolonal mouse anti-human estrogen receptor α, clone 1D5, isotype IgG1 kappa (DakoCytomation Denmark A/S, Glostrup, Denmark). Specificity: Antibody specifically reacts with oestrogen receptor α –positive cells, staining an epitope located in the N-terminal region of Erα. Pre-treatment: incubation in +100 °C for 10 minutes in 1 mM EDTA, 10 mM Tris / HCl, pH 8,5. Antibody dilution 1:50 in Large Volume UltrAb Diluent TA-125-UD (Lab Vision Corporation, Fremont, USA).
2) Monocolonal rabbit anti-human estrogen receptor α, clone SP1, isotype IgG (NeoMarkers Inc., Fremont, USA). Specificity: Antibody recognises a 67 kDa protein, which is identified as oestrogen receptor alpha. Pre-treatment: incubation

in +100 °C for 15 minutes in pH 6,0 citrate buffer (Dako Real Target Retrieval Solution S2031) (Dako Denmark A/S, Glostrup, Denmark). Antibody dilution 1:100 in Large Volume UltrAb Diluent (Lab Vision Corporation).

Progesterone receptor

Monoclonal mouse anti-human progesterone receptor, clone PgR 636, isotype IgG1 kappa (Dako Corporation, Carpinteria, USA). Specificity: Monoclonal anti-human progesterone receptor from clone 636 (anti-progesterone receptor) specifically recognises the A and B forms of the receptor. Pre-treatment: Incubation in +100 °C for 15 minutes in citrate buffer, pH 6.0 (Dako Denmark A/S). Total protein concentration 14.4 mg/ml, total mouse IgG concentration 238 µl/ml. Antibody dilution 1:400 in Large Volume UltrAb Diluent (Lab Vision Corporation).

C-erbB-2 / HER-2 / neu

One hundred and eighty-five kDa rabbit monoclonal antibody against human *c-erb*B-2 protein, clone SP3, isotype IgG. Specificity: Human *c-erb*B-2 receptor tyrosine kinase, which is closely related in structure to the epidermal growth factor receptor. Pre-treatment: Incubation in +100 °C for 15 minutes in pH 6.0 citrate buffer (Dako Denmark A/S). Antibody dilution 1:50 in Large Volume UltrAb Diluent (Lab Vision Corporation).

Cytokeratin 5

Lyophilized mouse monoclonal antibody against human cytokeratin 5, clone XM26, isotype IgG1 kappa (Novocastra Laboratories Ltd., Newcastle upon Tyne, UK). Specificity: Human cytokeratin 5 intermediate filament protein. Pre-treatment: Incubation in +100 °C for 10 minutes in buffer containing 1 mM EDTA, 10 mM Tris/HCl (pH 8,5). Total protein concentration 1.0 – 8.0 g/l, total antibody concentration greater than or equal

to 45.0 mg/l. Antibody dilution 1:200 in Large Volume UltrAb Diluent (Lab Vision Corporation).

Cytokeratin 14

Lyophilized mouse monocolonal antibody against human cytokeratin 14, clone LL002, isotype IgG3 (Novocastra Laboratories Ltd.). Specificity: Human cytokeratin 14 intermediate filament protein. Pre-treatment: Incubation in +100 °C for 10 minutes in buffer containing 1 mM EDTA, 10 mM Tris/HCl (pH 8.5). Total protein concentration: 1.0 – 8.0 g/l, total antibody concentration greater than or equal to 32.4 mg/l. Antibody dilution 1:200 in Large Volume UltrAb Diluent (Lab Vision Corporation).

Bcl-2α

Mouse monoclonal antibody against human Bcl-2α, clone 8C8, isotype IgG1 (NeoMarkers Inc.). Specificity: Human Bcl-2α oncoprotein that inhibits programmed cell death (apoptosis). Pre-treatment: Incubation in +100 °C for 10 minutes in pH 6.0 citrate buffer (Dako Denmark A/S). Antibody dilution 1:500 in Large Volume UltrAb Diluent (Lab Vision Corporation).

E-cadherin

One hundred and twenty kDa mouse monoclonal antibody against human E-cadherin, clone NCH-38, isotype 1gK1 kappa (NeoMarkers Inc.). Specificity: Human E-cadherin cell adhesion molecule of epithelial tissues. Pre-treatment: Incubation in +100 °C for 10 minutes in 1 mM EDTA, 10 mM Tris/HCl (pH 8.5) buffer. Antibody dilution 1:50 in Large Volume UltrAb Diluent (Lab Vision Corporation).

CD 117

Polyclonal rabbit anti-human CD 117, c-kit antibody (Dako Denmark A/S). The antibody labels the 145 kDa transmembrane tyrosine kinase receptor CD 117 / c-kit located, haematopoietic stem cells, melanocytes, mast cells, Cajal cells, basal cells of skin, and mammary ductal epithelia. Pre-treatment: Incubation in +100 °C for 10 minutes in 1 mM EDTA, 10 mM Tris/HCl (pH 8.5) buffer. Antibody dilution 1:500 in Large Volume UltrAb Diluent (Lab Vision Corporation).

5. RESULTS

A total of 2 123 representative sample cores (i.e. specimens that contained cancer cells) were analysed with immunological stains for seven important cancer markers: *oestrogen receptor, progesterone receptor, HER-2, E cadherin, cytokeratin 5+17, BCL-2,* and *CD 117.* Each sample core was scored as either negative (cancer cells in the sample did not express the given antigen) or positive (antigen expression indicated by the binding of its specific antibody). For further analysis, the staining intensity was also scored as +, ++ or +++, according to the intensity of the expression and the amount of positive cells when compared to the negative ones (adapted from Carey et al. 2006). Negative and positive controls were carried out according to the standard practices of a clinical pathology laboratory, as discussed in further detail in Chapter 4.1.5.

On average, 303 representative samples were observed (ranging from 275 to 313). On average, 66 unrepresentative sample cores were recorded per staining (ranging from 56 to 94), which is approximately 22% of the sample cores. The results of the staining, in absolute values, as well as in percentages, can be seen in Table 1.

TABLE 1

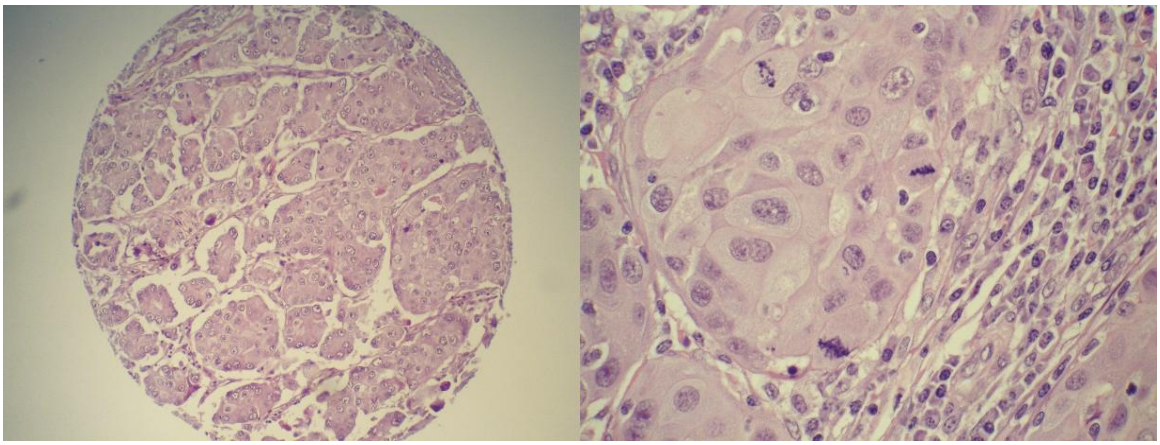
Marker:	ER	PR	HER-2	E-CAD	CYT 5+17	BCL-2	CD 117
x	61	57	56	64	94	67	61
0	73	107	254	23	235	59	298
1	44	58	13	61	22	54	9
2	108	90	29	149	6	111	1
3	83	57	17	72	12	78	0
Represent.	308	312	313	305	275	302	308
Negat.%	**23.7**	**34.3**	**81.2**	**7.5**	**85.5**	**19.5**	**96.8**
Posit. amnt.	235	205	59	282	40	243	10
Posit.%	**76.3**	**65.7**	**18.8**	**92.5**	**14.5**	**80.5**	**3.2**

Summary of the TMA scoring: x =number of unrepresentative samples, 0 = negative sample, 1 = low positive, 2 = average positive, 3 = strong positive. Represent. = number of representative samples in total, Negat.% = percentage of negative samples (no antigen binding), Posit. amnt. = number of positive samples in total, and Posit.% = percentage of positive samples (antibody binding observed).

In the following section, all the types of immunohistochemical staining used in this study are displayed with examples of the staining pattern, intensity and general cell morphology.

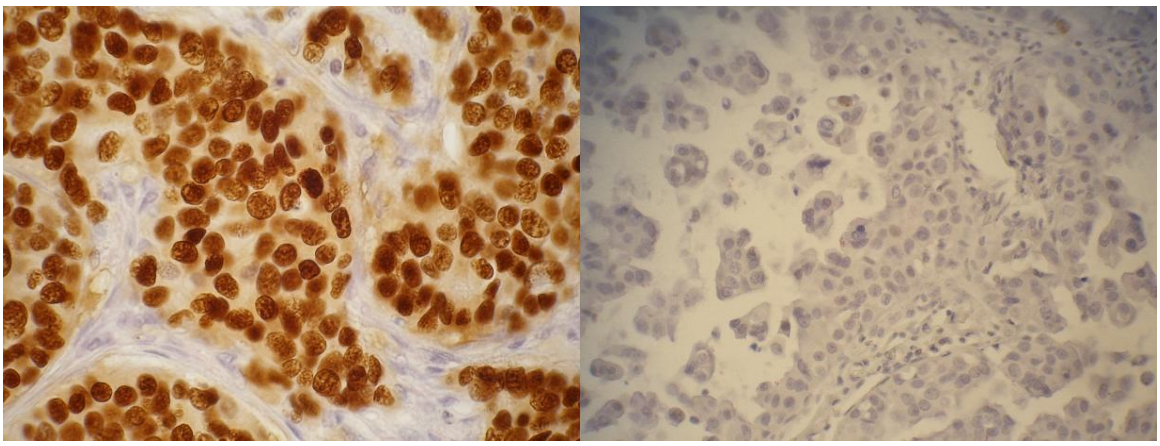
Each example includes positive and negative sample cores, as well as certain additional images where appropriate. In addition, an example of a standard HE image is given, as well as examples of typical unrepresentative sample cores.

Haematoxylin-eosin staining



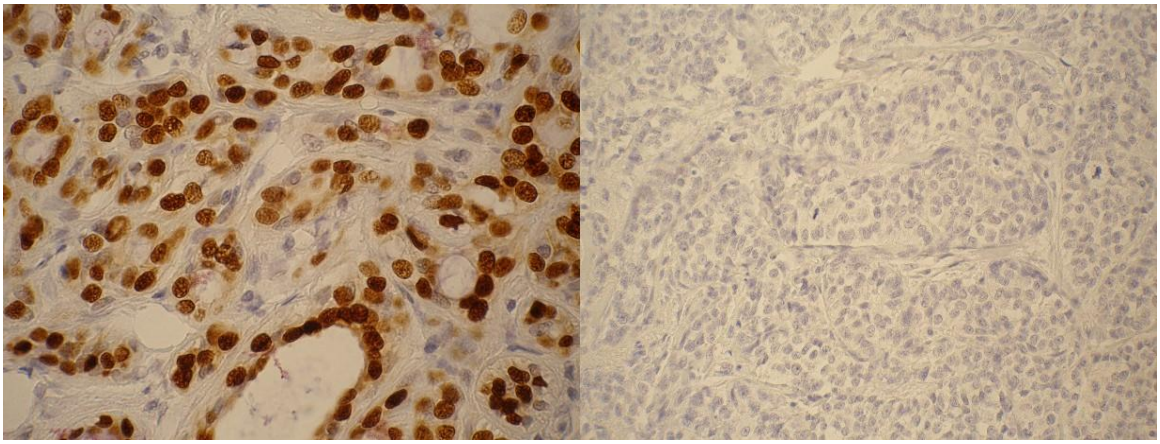
IMAGES 4a and 4b: 100x (a) and 400x (b) HE staining of breast carcinoma. 1a displays a TMA sample core fitting fully into the field of vision. Image 4b displays cancer cells, some of which with mitotic activity.

Oestrogen receptor staining



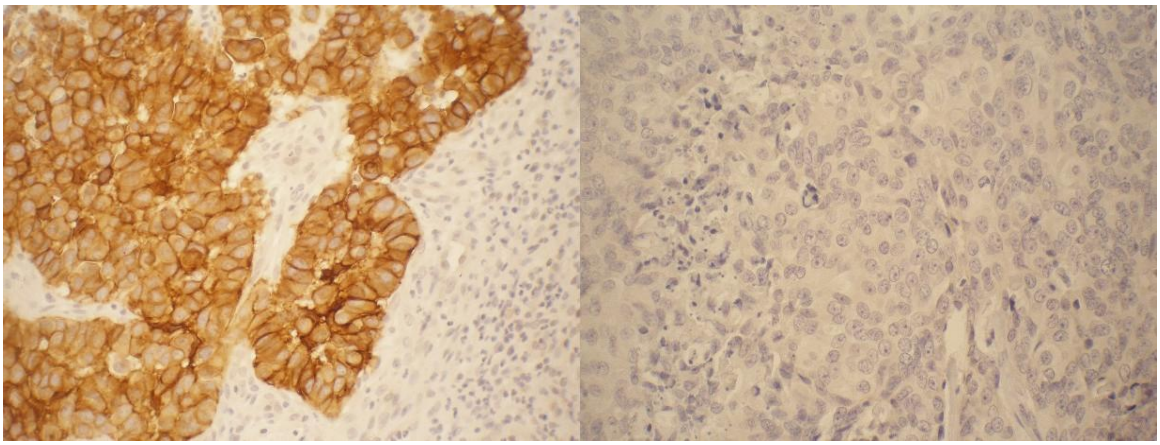
IMAGES 5a and 5b: 400x ER-positive (a) and 200x ER-negative breast cancer (b).

Progesterone receptor staining

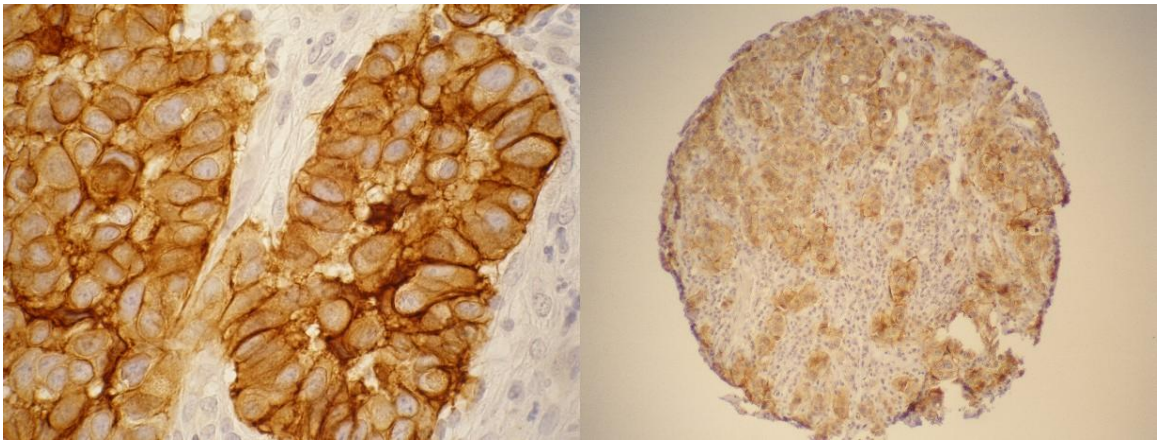


IMAGES 6a and 6b: 400x PR-positive (a) and 200x PR-negative (b) cancer. The nuclear expression of the progesterone receptor is clearly visible in 6a.

HER-2 staining

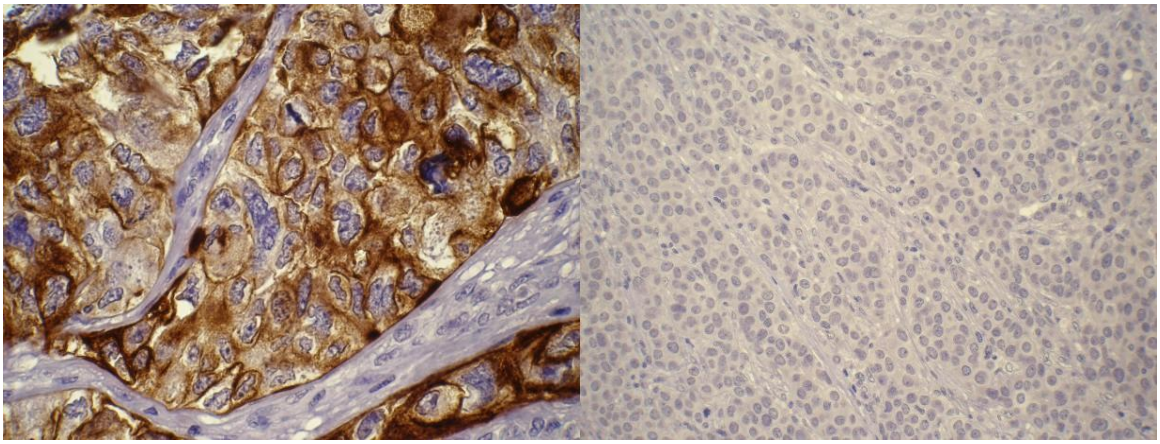


IMAGES 7a and 7b: 200x HER-2 -positive (a) and HER-2 -negative (b) cancer.



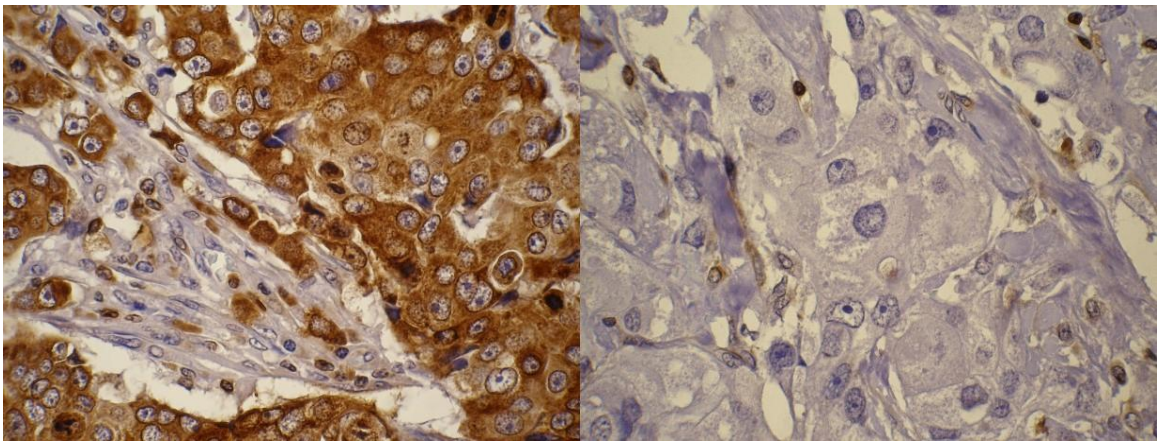
IMAGES 7c and 7d: 400x HER-2 -positive (c) and 100x HER-2- positive (d) cancer. 4c provides a closer look at the receptor expression, which is localized at the plasma membrane. 7d, in turn, displays a markedly lower intensity of staining; indicating lower expression rates of HER-2 receptor in this sample.

Cytokeratins 5 and 14



Images 8a and 8b: 400x cytokeratin 5 and 14 -positive (a) and 200x CK5+14 -negative (b) sample.

Bcl-2α



Images 9a and 9b: 400x Bcl-2α -positive (a) and -negative (b) cancer.

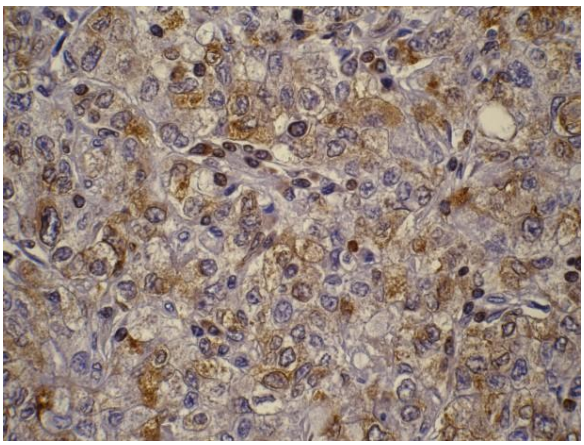
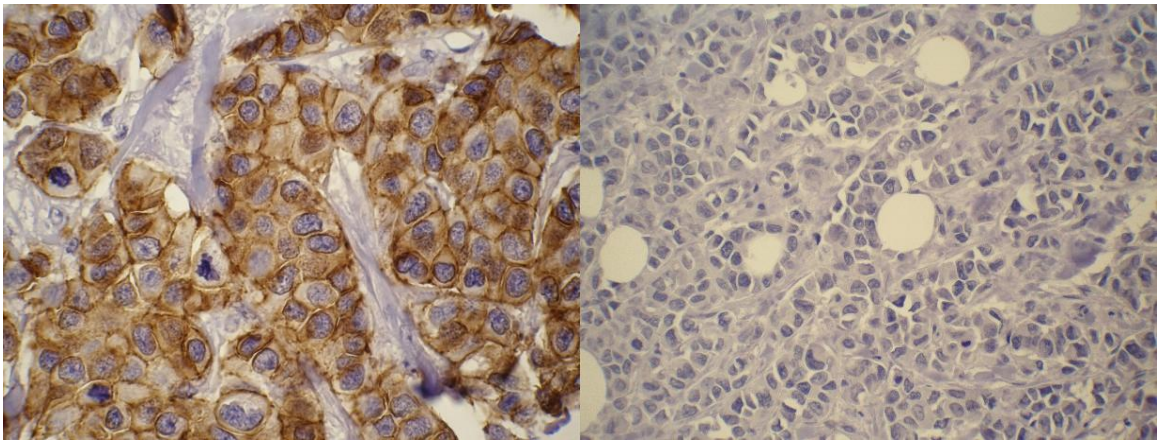


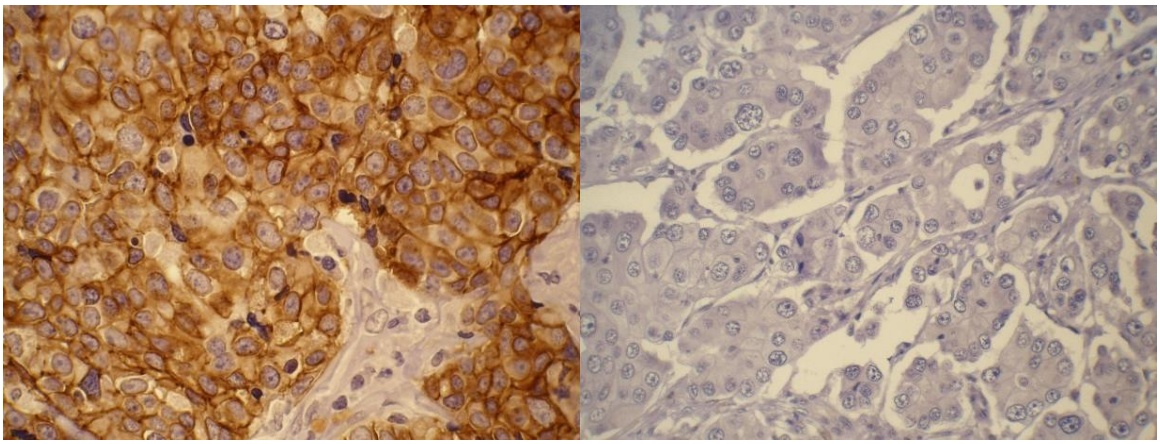
Image 9c: 400x Bcl-2α -positive sample. Note the weak cytoplasmic staining, when compared to the intense staining pattern of the Image 6a.

E-cadherin



Images 10a and 10b: 400x E-cadherin -positive (a) and 200x -negative (b) carcinoma. Several mitoses can be seen in the image 10a.

CD 117



Images 11a and 11b: 400x CD 117 -positive and -negative cancer.

Typical unrepresentative samples

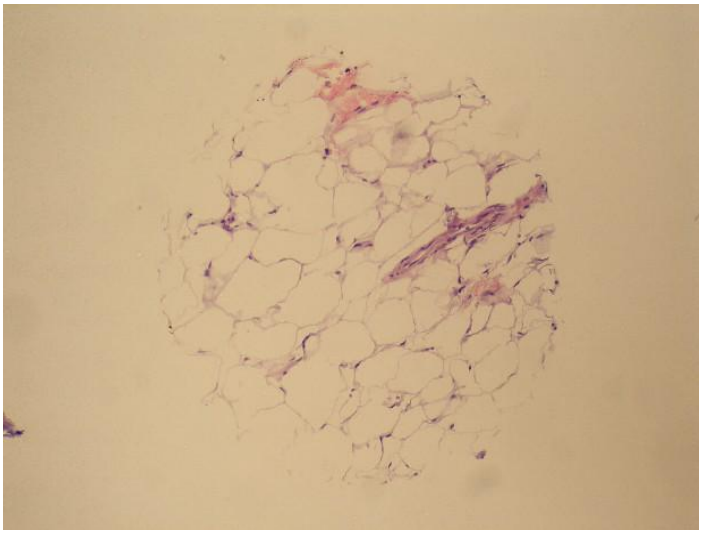


Image 12: 100x HE staining of breast tissue. The sample core has been extracted from a site of adipose tissue and contains no cancer cells.

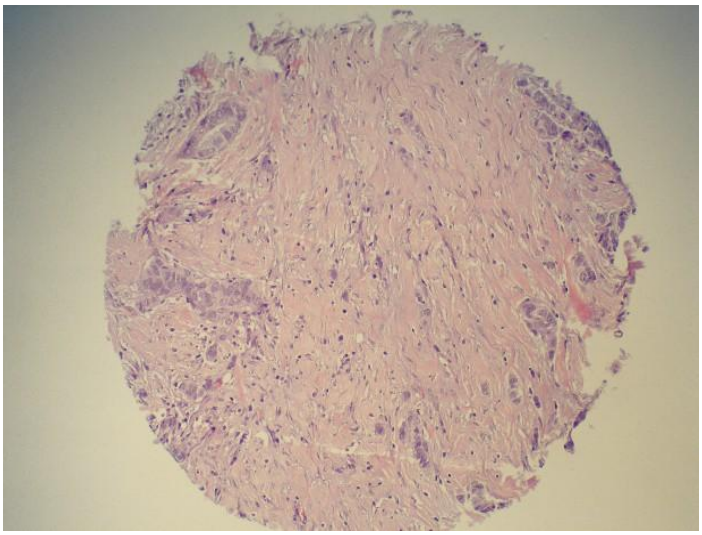


Image 13: 100x HE staining of breast tissue. The sampling needle ("punch") has hit a fibrous site in the tissue, thus extracting a core consisting mostly of connective tissue. Only limited amounts of cancer cells can be seen.

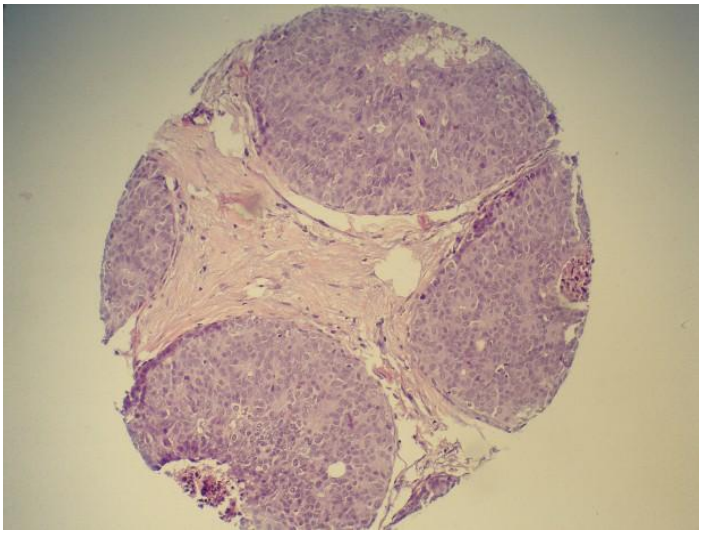


Image 14: 100x HE staining of breast tissue. The sample core contains prominent regions of ductal structures with morphological features suggesting ductal carcinoma in situ *(DCIS).*

6. DISCUSSION

Comparison of cancer marker prevalence with values in the literature

The prevalence of various cancer markers, such as those used in this study, has been extensively studied ever since the emergence of monoclonal antibody technology and subsequent immunological methods. This vast data allows in-depth evaluation of the tissue microarray data obtained from this study, and permits its assessment as a tool for cancer research in a small or middle-sized laboratory, where traditional histological methods would require too much manpower to conduct. It is important to understand the nature of the TMA methodology, and its similarities and differences with traditional means.

Due to the nature of TMA technology, a certain percentage of the sample cores are typically devoid of cancer cells, and are thus excluded from data analysis. This percentage is strongly linked with the abundance of cancer tissue and its heterogeneity in a given sample specimen, as well as with experience of the person constructing the array (Kallioniemi et al., 2001). In this study, on average 66 unrepresentative sample cores were recorded per staining (ranging from 56 to 94), which is circa 22% of all sample cores. According to the verbal statement from an experienced researcher of the field, these numbers are in line with reports from other research groups employing TMA technology. However, published studies about the actual percentages could not be found, for this kind of data is often left out from scientific papers as they probably are not part of the research problem. This notwithstanding, it is likely that the amount of unrepresentative samples can be significantly reduced by practice.

Apart from the higher number of unrepresentative samples, TMA constructs are very similar to the conventional histological methods: positive or negative results can be scored microscopically, although the method better suits measuring relative, rather than absolute, frequencies of cancer markers. Furthermore, the relative amount of cancer cells, or the intensity of the specific signal can be evaluated similarly to the traditional histology. Some investigators have preferred to use larger sample cores (e.g. 2 - 4 mm in diameter), but increasing the representativeness of heterogeneous samples is reported to be enhanced by

scoring several 0.6 mm sample cores from each tissue specimen - a method that has also been used in this study. (Kallioniemi et al., 2001)

However, much greater throughput of samples with profoundly reduced reagent consumption can be achieved with TMA technology, even when producing more than one sample core from each tissue specimen. The following chapter will separately consider each of the cancer markers used in this study, and compare the results with those found in the literature.

Oestrogen receptor

Out of 308 representative breast cancer samples, 76.3% were identified as oestrogen receptor (ER) positive. Each individual sample core was scored as stated in Chapter 5 and each patient sample (i.e. two sample cores from each patient) was deemed positive when a positive staining was detected in both of the sample cores taken from a tissue specimen. When the other sample core did not have cancer cells and was thus excluded from the analysis as an unrepresentative sample, one ER positive sample core was scored as a positive result. 23.7% of the ER stained samples were negative. A patient sample was deemed negative when both of the sample cores were negative, or in the event of only one representative sample core, it was scored as a negative result. Any ambiguous results (i.e. one sample core is positive and the other negative), would have been excluded from the analysis, but this did not take place. The scoring principles outlined above were applied to all subsequent cancer marker analysis.

Typically breast cancers have been reported to be ER positive in approximately 50 - 80% of cases when evaluated by traditional histochemical means or with such molecular biology methods as RT-PCR (e.g. Kumar et al., 2005 and Lynq et al., 2008). Invasive breast carcinomas, such as those used in this study, are often regarded to express the ER closer to the 75 - 80% end of the scale, although the exact values vary somewhat between different studies, due to differences in the age profile, ethnicity and other variable factors regarding the patient group in each study (Benz et al., 2005, Carey et al., 2006 and Lynq et

al., 2008). For example, ER expression levels are higher among post-menopausal women (Kumar et al., 2005). Since the age profile of the patients in this study is between 50 to 60 years of age, values close to the 80% end can be expected.

The percentage of ER positive samples, 76.3, found in this study is well in line with the values found in literature, indicating that the results are valid and represent the ER status of the studied population well.

Progesterone receptor

Three hundred and twelve representative samples were analysed for the progesterone receptor (PR) expression. Of these, 65.7% were positive and 34.3% negative, when analysed according to the outlines stated in the previous section.

It is estimated that approximately 55 to 60% of invasive breast cancers are progesterone receptor positive. (Weinberg 2006; Rakha et al., 2007b) However, since PR positive cells are very rarely ER negative (from 2.5 to 3.5%, according to De Maeyer et al. [2007] and Rakha et al. [2007b]), the PR expression is often compared to that of the ER. Literature reports indicate, that circa 65 - 85% of ER positive breast carcinomas are also PR positive (Weinberg 2006; Rakha et al., 2007b). In this study, only one PR positive cancer out of 205 (0.49%) samples was ER negative. Of 231 ER positive cancers, which were also representative for PR, 198 (85.7%) were also PR positive, which is very close to the values reported by Rakha et al. (2007b), 84.4%. However, the amount of ER-/PR+ cases appears to be somewhat lower than expected.

C-erbB-2 / HER-2 / neu

Fifty-nine (18.8%) out of 313 meaningful samples were positive for the HER-2 receptor, as evaluated by the principles listed above in the section *Oestrogen receptor*. HER-2 receptor is usually over-expressed due to amplification in the chromosome 17q21, and it is

estimated that 20 to 30% of invasive breast carcinomas over-express HER-2 (Kumar et al., 2005). In this light, findings of this study represent the general statistics relatively well, albeit being slightly below the low end of the scale.

Although HER-2 over-expression is generally regarded as a sign of an aggressive carcinoma type, therapeutic agents such as trastuzumab (trade name Herceptin®) have been successful in treating these types of cancer. However, certain HER-2 negative cancers are linked to the so-called "triple-negative" cancers, which in addition to being HER-2 negative are ER and PR negative. Patients with this kind of carcinoma therefore lack the advantages of trastuzumab and hormonal treatment. (Rakha et al., 2007a)

In this study, the amount of triple negative cancers was also evaluated. Of the 309 samples that were informative for all the three markers, 73 (23.6%) were ER-/PR-/HER-2- ("triple negative"). Rakha et al. (2007a) report in their comprehensive study that they found 16.3% of their breast cancer samples being triple negative. However, in their study, they report that the median age of their patient cohort is 49.9 years (ranging from 25 to 70 years versus 50 to 60 years in this study), which may account for some of the 7.3% difference between these two studies. Other variables, such as ethnicity may similarly influence the results.

Cytokeratins 5 and 14

For cytokeratins 5 and 14 (CK 5/14), a total of 275 meaningful samples were obtained. Out of these, 40 samples (14.5%) were positive and 235 (85.5%) negative. CK 5/14, which is linked to basal origins of a cancer, have been studied e.g. by Laakso et al. (2005), who report that 9% out of 288 ductal invasive breast carcinomas were positive for these proteins. Furthermore, they report an inverse relationship between CK 5/14 and HER-2 expression. Similar trend is also seen in this study, where only four (10%) weakly CK 5/14 positive samples also displayed HER-2 expression, a lower percentage when compared to the general level of HER-2 expression (18.8% in this study). Finally, it is worth noting that 22 (55%) of CK 5/14 positive samples were only weakly positive ("+"), which suggests that some of these might be negative under more precise scrutiny. All HER-2 + CK 5/14

positive samples were in this category. If only moderate ("++") and strong ("+++") staining results (N = 18) are included in the analysis, 6.5% of samples were clearly positive for CK 5/14.

In another, more recent study, Jumppanen et al. (2007) report that 13% of 375 stage II breast cancer tumours were CK 5/14 positive. Whilst results from our TMA analysis lie in close proximity of these reported values, more precise criteria need be established in order to increase the reliability of the analysis.

Bcl-2

Bcl-2, an apoptosis-related oncoprotein, is linked to many cancer types. It is over-expressed in ca. 40 to 80% of breast cancers, depending on the age, hormonal status, and ethnicity of the patient, as well as on the cancer subtype (Aqqarwal et al. 2007). In this study, anti-Bcl-2α (alpha chain) antibodies were used to identify this protein. Fifty-nine (19.5%) samples out of 302 were Bcl-2 negative and 243 (80.5%) positive. The results of our study are at the high-end of the reported values in the literature, but closer evaluation of the results would require age, hormonal status, and tumour grade -matched comparison group. At the time of writing, these kinds of results were not available, thus leaving the analysis of Bcl-2 results somewhat open.

E-cadherin

Two-hundred and eighty-two samples out of 305 meaningful results were positive for E-cadherin (92.5%). E-cadherin is an epithelial adhesion molecule that is located on cell surface. Its loss of expression or decrease in function is linked to decreased strength of cell adhesion, and subsequently to a higher risk of cells being able to infiltrate through the basement membrane and to the surrounding tissues. Thus it also increases the risk of metastasis formation. (Kumar et al. 2005)

Therefore, simple negative versus positive classification is not sufficient regarding this marker, as lessened expression can already indicate some degree of loss of function. In this study, 72 samples (23.6%) had strong ("+++") and 149 (48.9%) moderate ("++") E-cadherin staining. In 84 samples (27.5%) the expression levels were markedly reduced or absent ("+" and "-", respectively). However, the evaluation of the exact amount of reduction in the staining intensity that indicates loss of E-cadherin function is beyond the scope of this study.

It appears that E-cadherin expression in breast cancer has been studied relatively scarcely during recent years. However, certain papers can be found that compare the amount of E-cadherin expression, and evaluate whether it is normal or down-regulated. For example, Kowalski et al. (2003) report that in 55% of primary invasive ductal breast carcinomas (N = 22) the E-cadherin expression remained normal, and in 45% it was reduced or absent. In this study, the small number of samples makes the comparison of results somewhat troublesome. However, Zhang et al. (2007) evaluated 90 invasive breast carcinomas, in which they reported a 60% reduction (54 samples) in E-cadherin expression. Neither of these results is well in line with results obtained in this study, but as mentioned above, a more precise analysis between these studies requires more accurate ways to define "normal" and "reduced" E-cadherin expression in our study.

CD 117

The last cancer marker of the analysed set of proteins was CD 117 (also known as c-Kit and KIT), is a type III tyrosine kinase receptor, whose mutated forms are linked to several types of cancer, including breast carcinoma. It appears to be rarely expressed in invasive breast cancer: in our study only one sample (0.3%) was moderately positive ("++") and ten (3.2%) showed slight staining ("+"). These results are backed, for example, by the study of Dabiri et al. (2004), who studied a tissue microarray of 348 invasive breast cancer samples without finding any CD 117 positive samples. However, a large TMA study of 4 444 samples by Rajput et al. (2008) indicates, that despite the low prevalence of CD 117

expression in most cancers, the marker does have prognostic relevance when evaluating certain rarer forms of breast cancers, such as those that involve stromal mast cells.

Conclusion

The tissue microarray block and its subsequent analysis were performed by the author, with the possibility to consult pathologists and clinical biologists during the process. The cancer tissue on paraffin blocks was marked based on their corresponding haematoxylin-eosin stained object slides. The process of screening the slides, constructing the tissue microarray and scoring the results required approximately three months' time when working 4-hour days on average. However, the project was stalled on several occasions due to a brittle set of sampling needles ("punches") and ordering new needles from the United States. In this light, it is probable, that a full-time worker could construct and analyse a similar-sized TMA (734 sample cores) within a month's time, and analyse a panel of five to ten markers within a week or two. In this light, the method is a viable option for a small research group or laboratory, or when the resources are limited.

Furthermore, the whole process can easily be divided among several workers, when constructing bigger arrays or constructing arrays more rapidly becomes possible. On one hand, this allows large-scale screening projects to be conducted within a very reasonable time scale. On the other hand, even one person can construct a comprehensive set of analysis during a longer time span, since the materials and methods used in the process are not affected by time (the initial samples, the array itself and the subsequent stained object slides lasting for years).

When suitable samples are available, little else is required for the process, and it can be carried out in multiple locations if required. A tissue microarrayer is required along with a standard light microscope, and equipment for immunohistochemical staining need to be available. Tools for complex statistical analysis are also essential if multivariable analysis based on the patient diagnosis and/or outcome of the disease are conducted. Finally, certain

other types of staining can be performed on the TMA slides, such as fluorescent *in situ* hybridisation (FISH).

In this study, seven cancer markers were scored, and the results were compared to those found in the literature. Due to inevitable variations within the patient groups between the studies, some degree of uncertainty remains as long as age-, hormonal status- and ethnicity-matched studies are not readily available. For the most common markers, such as the oestrogen and progesterone receptors, plentiful data is available, which increases the reliability of the analysis. For some rarer markers, e.g. the E-cadherin, the research papers used for comparison had more flexibility regarding the variables between the patient groups.

This kept in mind, it can be said that the results of this study are reasonably well in line with the results published by other research groups. The construction of the array can be learnt with relative ease, and the stained samples can be analysed by a pathologist or a cell biologist, or other laboratory personnel under the supervision of the aforementioned.

Patient data analysis was also planned to be carried out based on the results of this study and the patient records of the Central Hospital of Central Finland. Unfortunately, however, the schedule of this study did not permit such a study be performed, due to the problems encountered with the sampling needles along the array construction that hindered the process. In-depth patient data analysis is planned to be performed at a later point of time as a follow-up for this work and published as an independent study.

7. FURTHER RESEARCH TOPICS

The breast cancer database created by tissue microarray technology in this study offers a vast material for further research. When linked with the patient diagnostic and follow-up data it provides means for a comprehensive multivariable study that can involve complex statistics. The effect of a given cancer type on the outcome of the cancer can give valuable information on the efficiency of the current treatment and use of therapeutics. Furthermore, linkages and patterns between the different cancer markers can be discovered, providing insights to the complex patterns of gene expression and their role in the subsequent cancer phenotype.

Additional staining can be performed as well from the existing blocks. For example, confocal laser scanning microscopy (CLSM) offers ways to analyse co-expression on cancer markers by using fluorescent dyes with different emission spectra. This allows the study of the expression of multiple cancer markers simultaneously from any given sample, something that is challenging to achieve with normal light microscopy when standard immunohistological samples are used. (Johnstone and Turner, 1997; Robertson et al., 2008)

However, the possibility to stain subsequent sections from the tissue microarray for different markers allows a fairly good degree of co-localisation studies even with light microscopy. Thus, the usefulness of CLSM studies needs to be evaluated further.

With the emergence of new and improved TMA arrayers, together with novel computerised image processing techniques, vast databases of archival cancer samples of various types can be created. The Central Hospital of Central Finland possesses a comprehensive collection of suitable paraffin-embedded tissue samples that are compatible with the method, and can be readily linked with patient data. In the future, more age groups or completely new cancer types can be added to the material of this study, and the distinct markers of these types can be compared to those of the breast carcinoma.

8. REFERENCES

Antoniou, A.C., A. Shenton, E.R. Maher, E. Watson, E. Woodward, F. Lalloo, D.F. Easton, and D.G. Evans. 2006. Parity and breast cancer risk among brca1 and brca2 mutation carriers. *Breast Cancer Res.* 8:R72.

Aqqarwal, H., P.S. Lubana, D.K. Jain, and R.K. Mathur. 2007. Estimation of BCL-2 protein in carcinoma of the breast and its clinical correlation in locally advanced breast cancer. *J Cancer Res Ther.* 3:207 - 210.

Battifora, H. 1986. The multitumor (sausage) tissue block: Novel method for immunohistochemical antibody testing. *Lab Invest.* 55:244 - 248.

Boultwood, J., and C. Fidler. 2002. Molecular analysis of cancer. *In* Methods in molecular medicine. Humana Press. 2 - 29.

Carey, L.A., C.M. Perou, C.A. Livasy, L.G. Dressler, D. Cowan, K. Conway, G. Karaca, M.A. Troester, C.K. Tse, S. Edmiston, S.L. Deming, J. Geradts, M.C.U. Cheang, T.O. Nielsen, P.G. Moorman, S. Earp, and R.C. Millikan. 2006. Race, breast cancer subtypes, and survival in the carolina breast cancer study. *JAMA.* 295:2492 - 2502.

Claudio, L. 2006. Making progress on breast cancer. *Environ Health Perspect.* 114:A98 - A99.

Colozza, M., E. de Azambuja, F. Cardoso, C. Bernard, and M. Piccart. 2006. Breast cancer: Achievements in adjuvant systemic therapies in the pre-genomic era. *Oncologist.* 11:111 - 125.

Conway, C.M., D. O'Shea, D. K. Lawler, G. D. Dodrill, A. O'Grady, H. Barrett, C. Culmann, L. O'Driscoll, W. M. Gallagher, E. W. Kay, and D. G. O'Shea. 2006. The development and validation of the Virtual Tissue Matrix, a software application that facilitates the review of tissue microarrays on line. *BMC Bioinformatics.* 7:256.

Cox, D.G., R.M. Tamimi, and D.J. Hunter. 2006. Gene x gene interaction between mnsod and gpx-1 and breast cancer risk: A nested case-control study. *BMC Cancer.* 6:217.

Creswell, J.W. 2003. Research design: qualitative, quantitative and mixed methods approaches. Sage Publications, Inc. 101 - 131.

Dabiri, S., D. Huntsman, N. Makretsov, M. Cheang, B. Gilks, C. Badjik, K. Gelmon, S. Chia, and M. Hayes. 2004. The presence of stromal mast cells identifies a subset of invasive breast cancers with a favourable prognosis. *Modern Pathology.* 17:690 - 695.

De Maeyer, L., E. Van Limbergen, K. De Nys, P. Moerman, N. Pochet, W. Hendrickx, H. Wilders, R. Paridaens, A. Smeets, M-R. Christiaens, I. Vergote, K. Leunen, F. Amant, and P. Neven. 2008. Does estrogen receptor negative/progesterone receptor positive breast carcinoma exist? *J Clin Oncol.* 26:335 - 6.

DePalma, G.D., S. Masone, M. Rega, I. Simeoli, M. Donisi, P. Addeo, L. Iannone, V. Pilone, and G. Persico. 2006. Metastatic tumors to the stomach: Clinical and endoscopic features. *World J Gastroenterol.* 12:7326 - 7328.

Dunnwald, L.K., M.A. Rossing, and C.I. Li. 2007. Hormone receptor status, tumor characteristic and prognosis – a prospective cohort of breast cancer patients. *Breast Cancer Res.* 9.

Gancberg, D., A. Di Leo, G. Rouas, T. Järvinen, A. Verhest, J. Isola, M.J. Piccart, and D. Larsimont. 2002. Reliability of the tissue microarray based FISH for evaluation of the HER-2 oncogene in breast carcinoma. *J Clin Pathol.* 55:315 - 317.

Glick, B.R., and J.J. Pasternak. 2003. Molecular biotechnology: Principles and applications of recombinant DNA. ASM Press, Washington D.C. 268 - 269.

Gregory, I. 2003. Ethics in research. Continuum International Publishing Group, Inc. 30 - 45, 88 - 92.

Gruvberger-Saal, S.K., H.E. Cunliffe, K.M. Carr, and I.A. Hedenfalk. 2006. Microarrays in breast cancer research and clinical practice--the future lies ahead. *Endocr Relat Cancer*. 13:1017 - 1031.

Hager, M., C. Kolbitsch, W. Tiefenthaler, H. Haufe, R. Kemmerling, and P.L. Moser. 2007. Tissue microarrays from renal cell tumors: exclusion criteria and rate of exclusion. *Scand J Urol Nephrol.* 16:465 - 489.

Hakama, M., E. Pukkala, M. Heikkela, and M. Kallio. 1997. Effectiveness of the public health policy for breast cancer screening in finland: Population based cohort study. *BMJ*. 314:864 - 867.

Howat, W.J., A. Watford, J.N. Mitchell, K.F. Clarke, J.S. Conquer, and J. McCafferty. 2005. Resin tissue microarrays : A universal format for immunohistochemistry. *Histochem Cytochem*. 53:1189 - 1197.

International Human Genome Sequencing Consortium, The (IHGSC). 2004. Finishing the euchromatic sequence of the human genome. *Nature*. 431:931 - 945.

Immonen-Raiha, P., L. Kauhava, I. Parvinen, K. Holli, P. Kronqvist, L. Pylkkanen, H. Helenius, A. Kaljonen, O. Rasanen, and P.J. Klemi. 2005. Mammographic screening reduces risk of breast carcinoma recurrence. *Cancer*. 103:474 - 482.

Johnstone, A.P., and M. W. Turner. 1997. Immunochemistry 2 - A practical approach. Oxfod university Press. 73 - 76.

Jumppanen, M., S. Gruvberger-Saal, P. Kauraniemi, M. Tanner, P.O. Bendahl, M. Lundin, M. Kroqh, P. Kataja, A. Borg, M. Fernö, and J. Isola. 2007. Basal-like phenotype is not associated with patient survival in estrogen-receptor-negative breast cancers. *Breast Cancer Res.* 9:R16.

Kallioniemi, O.-P., U. Wagner, J. Kononen, and G. Sauter. 2001. Tissue microarray technology for high-throughput molecular profiling of cancer. *Hum Mol Genet*. 10:657 - 662.

Kapp, A.V., S.S. Jeffrey, A. Langerød, A.-L. Børresen-Dale, W. Han, D.-Y. Noh, I.R.K. Bukholm, M. Nicolau, P.O. Brown, and R. Tibshirani. 2006. Discovery and validation of breast cancer subtypes. *BMC Genomics*. 7.

Kononen, J., L. Bubendorf, A. Kallioniemi, M. Barlund, P. Schraml, S. Leighton, J. Torhorst, M.J. Mihatsch, G. Sauter, and O.P. Kallioniemi. 1998. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med.* 4:844 - 847.

Kowalski, P.J., M.A. Rubin, and C.G. Kleer. 2003. E-cadherin expression in primary carcinomas of the breast and its distant metastases. *Breast Cancer Res.* 5:217 - 222.

Kumar, V., A.K. Abbas, and N. Fausto. 2005. Robbins and Cotran pathologic basis of disease. Elsevier Saunders. 1138 - 1146.

Laakso, M., N. Loman, A. Borg, and J. Isola. 2005. Cytokeratin 5/14-positive breast cancer: true basal phenotype confined to BRCA1 tumors. *Mod Pathol.* 18:1321 - 1328.

Lyng, M.B., A-V. Lankholm, N. Pallisgaard, and H.J. Ditzel. 2008. Identification of genes for normalization of real-time RT-PCR data in breast carcinomas. *BMC Cancer*. 8:20 - 41.

Macaskill, E.J., L. Renshaw, and J.M. Dixon. 2006. Neoadjuvant use of hormonal therapy in elderly patients with early or locally advanced hormone receptor-positive breast cancer. *Oncologist.* 11:1081 - 1088.

Mansfield, L., I. Sosa, R. Dionello, A. Subramanian, H. Devalia, and K. Mokbel. 2007. Current management of the axilla in patients with clinically node-negative breast cancer: A nationwide survey of united kingdom breast surgeons. *Int Semin Surg Oncol*:Epub ahead of print.

Nagata, C., T. Mizoue, I. Tsuji, K. Wakai, M. Inoue, and S. Tsugane. 2006. Tobacco smoking and breast cancer risk: An evaluation based on a systematic review of epidemiological evidence among the japanese population. *Jpn J Clin Oncol.* 36:387 - 394.

Rajput, A.B., D.A. Turbin, M.C.U. Cheang, D.K. Voduc, S. Leung, K.A. Gelmon, C.B. Gilks, and G. Huntsman. 2007. Stromal mast cells in invasive breast cancer are a marker of favourable prognosis: a study of 4,444 cases. *Breast Cancer Res Treat.* 107:249 - 257.

Rakha, E.A., M.E. El-Sayed, A.R. Green, A.H.S. Lee, J.F. Robertson, and I.O. Ellis. 2007 (a). Prognostic markers in triple-negative breast cancer. *Cancer*. 109:25 - 32.

Rakha, E.A., M.E. El-Sayed, A.R. Green, E.C. Paish, D.G. Powe, J. Gee, R.I. Nicholson, A.H.S. Lee, J.F.R. Robertson, and I.O. Ellis. 2007 (b). Biologic and clinical characteristics of breast cancer with single hormone receptor-positive phenotype. *J Clin Oncol.* 25:4772 - 4778.

Robertson, D, K. Savage, J.S. Reis-Filho, and C.M. Isacke. 2008. Multiple immunofluorescence labelling of formalin-fixed paraffin-embedded (FFPE) tissue. *BMC Cell Biol.* Epub ahead of print.

Schulz, W.A. 2005. Molecular biology of human cancer - an advanced student's textbook. Springer Science + Business Media, Inc.

Simon, R., M. Mirlacher, and G. Sauter. 2005. Tissue microarrays. *Methods Mol Med.* 114:257 - 268.

Trop, I., and W. Deck. 2006. Should women 40 to 49 years of age be offered mammographic screening? *Can Fam Physician.* 52:1050 - 1052.

Tworoger, S.S., A.H. Eliassen, T. Kelesidis, G.A. Colditz, W.C. Willett, C. Mantzoros, and S.E. Hankinson. 2007. Plasma adiponectin concentrations and risk of incident breast cancer. *J Clin Endocrinol Metab.* In print.

Vocaturo, A., F. Novelli, M. Benevolo, F. Marandino, A.M. Cianciulli, R. Merola, R.P. Donnorso, I. Sperduti, S. Buglioni, and M. Mottolese. 2006. Chromogenic in situ hybridization to detect her-2/neu gene amplification in histological and thinprep-processed breast cancer fine-needle aspirates: A sensitive and practical method in the trastuzumab era. *Oncologist.* 11:878 - 886.

Weinberg, R.A. 2006. The Biology of Cancer. Garland Science. 410-440.

Zhang, Y., J. Du, X. Tian, Y. Zhong, and W. Fang. 2007. Expression of E-cadherin, beta-catenin, cathepsin D, gelatinases and their inhibitors in ductal breast carcinomas. *Chin Med J.* 120:1597 - 1605.

References in the Internet

http://www.aicorp.com/products/02TMA.htm 111207 (retrieved on 18.2.2008)

http://www.beecherinstruments.com/index.php?option=com_content&task=section&id=4&Itemid=27 (retrieved on 28.3.2008)

http://www.beecherinstruments.com/index.php?option=com_content&task=section&id=10&Itemid=58 (retrieved on 28.3.2008)

http://www.definiens.com/news_dtls.php?id=43&link_id=27&sublink_id=30 (retrieved on 11.1.2008)

http://www.cancerregistry.fi/tutkimus/ (retrieved on 13.11.2007)