Università degli Studi di Trieste

Graduate School in MOLECULAR BIOMEDICINE

PhD Thesis

BREAST CANCER BIOMARKERS: FROM IDENTIFICATION TO APPLICATION TO FFPE TISSUES

DANAE PRACELLA

XXV ciclo – Anno Accademico 2012

UNIVERSITÀ' DEGLI STUDI DI TRIESTE

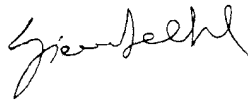
XXV CICLO DELLA SCUOLA DI DOTTORATO DI RICERCA IN BIOMEDICINA MOLECOLARE

BREAST CANCER BIOMARKERS: FROM IDENTIFICATION TO APPLICATION TO FFPE TISSUES

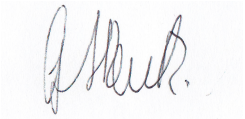
Settore scientifico-disciplinare: MED/08

DOTTORANDA
DANAE PRACELLA

DIRETTORE DELLA SCUOLA
PROF. GIANNINO DEL SAL



SUPERVISORE DI TESI
PROF. GIORGIO STANTA



TUTORE
DOTT. SSA SERENA BONIN

ANNO ACCADEMICO 2011 / 2012

ABSTRACT

Background: Breast carcinoma (BC) is the most common form of malignancy in women and the leading cause of cancer-related mortality among females internationally. BC includes a series of heterogeneous tumours with a great variability at histological level, biological level and clinical evolution. Because its complexity, BC treatment recommendations are continually changing with the new advances in this field. However there are still a significant number of patients with similar features that show distinct outcome. In order to detect the existing molecular differences and address patients to more personalized treatment, there is a need to find out new cancer biomarkers.

Aim: The main goal of my PhD project is to investigate on the possibility of combining traditional clinical and pathological features with new candidate biomarkers for the prognostication of breast cancer. The first part concerns the prognostic role of molecular classification of primary tumours, according to luminal A, luminal B, HER-2+ and basal-like subtypes. In the second part, the prognostic value of nine candidate genes was investigated at mRNA level. The genes of interest belong to the RB pathway (RB and CDK2), the RAS pathway (HER-2, PI3K, AKT1, AKT2, AKT3 and RAF1), and cellular differentiation mechanism (CK8), all involved in major cancer hallmarks, such as cell proliferation and survival.

Methods: This retrospective study comprises 305 women, which presented stage I-III invasive BC, age at diagnosis ≤55 years, and long follow-up period if alive. The cohort includes 151 lymph node negative (LN-) and 154 lymph node positive (LN+) patients. The molecular classification of primary tumours was performed by means of immunohistochemistry (IHC), using seven surrogate markers: ER, PR, HER-2, Ki67, CK8 and CK5/6, plus vimentin. Results of molecular classification were analysed with respect to morphologic and pathological features, and outcome. Moreover, the molecular characterization was also performed in a set of loco-regional metastatic lymph nodes, to compare the phenotype of primary tumour cells with their matched metastatic cells colonizing regional nodes. To the second purpose, gene's expression was investigated in the entire cohort of primary tumours by means of real-time PCR, using the TaqMan chemistry. The expression of the nine genes was investigated in connection with the clinical-pathological factors, molecular classification and BC specific patient's survival.

Results: Regarding molecular classification of primary tumours, luminal A, luminal B, HER-2+ and basal-like accounted for 46%, 34%, 8% and 12% respectively. Luminal A tumours were mainly LN-, well differentiated and stage I, while luminal B and HER2+ showed higher tumour grade, nodal metastases as well as higher proliferation status and stage. Luminal A exhibited better survival in comparison to the other subtypes ($p<0.001$). HER2+ and basal-like showed a poorer outcome, with a BC specific death occurring mostly within 5 and 10 years, respectively. Despite of the longer survival of patients with luminal tumours, they are the only one that underwent long-term recurrences, especially those patients characterized by high mitotic index and positive vimentin. The molecular classification at the level of loco-regional metastasis, revealed that HER-2, Ki67, CK8 and vimentin positivity was significantly decreased, whereas CK5/6 positivity was increased. No significant differences for ER and PR positivity between primary and metastatic lesions were found. The assessment of CK8 expression in the metastatic nodes was significantly associated to BC specific survival. Regarding gene's expression, we found a significant higher expression of RB, PI3K, AKT2 and AKT3 in LN- group, while CDK2, HER-2, AKT1 and CK8 were higher expressed in LN+ group. CDK2 high expression was significantly associated to the presence of later recurrences, in the entire cohort of patients. We also found specific pattern of genes' expression among molecular classes for: CDK2, HER-2, PI3K, AKT2, AKT3 and CK8. Survival analysis revealed that: i) high status of AKT1 and AKT2 seem to be related to worse survival in the entire cohort; ii) AKT3 and RB were associated to longer survival in LN-; iii) high AKT2 expression in luminal B seem to be associated to shorter survival; iv) high AKT3 status was associated to longer survival in luminal A tumours.

Conclusions: We believe that the use of the traditional biomarkers ER, PR, HER-2 and Ki67 is essential for BC characterisation and prognostication in association with clinical pathological features. Nevertheless, we identified new molecular markers that could better distribute patients into more homogeneous subgroups of BC. Finally, high AKT3 expression at mRNA level seems to be an independent indicator of good prognosis in patients showing ER+/HER2- and low proliferation index tumours.

TABLE OF CONTENTS

1 INTRODUCTION

Cancer represents a major public health problem in western societies, for this reason in the last decades many efforts aimed to defeat cancer have been made.

Standard diagnostic procedures for human tumours are worldwide performed in formalin fixed and paraffin embedded (FFPE) tissues, through histopathological examination and immunohistochemistry analysis combined to clinical data. The histopathological approach is aimed to define the nature, the tissue origin, the morphologic characteristics and the evolution of the disease (Comanescu et al., 2012). More recently, immunohistochemistry analysis allowed pathologists to complement morphology information to reach a better classification of tumours (Comanescu et al., 2012). Those data are relevant also for patient's prognosis and represent the rational basis for the choice of therapy. However, cancer is a heterogeneous and complex disease with differences also at the molecular level. Patients, with the same histopathological features could have very different clinical outcome. To improve this situation during recent years many efforts have been made to find molecular markers and gene signatures useful to predict patients' prognosis. All the experiments performed in this research were carried out on formalin fixed and paraffin embedded (FFPE) tissues, since they are greatly available and they could be selected with long follow-up period. Recently, most molecular analyses can be made in FFPE tissues (Stanta, 2011), consequently it is possible to gain new knowledge on cancer directly on human tissues, with a substantial reduction of the time needed for clinical application, compared to basic research.

1.1 Epidemiology

1.1.1 Breast Cancer Incidence and Mortality

Invasive breast cancer is the most common cancer among women. In 2008 a total of 1.38 million of breast cancer cases were diagnosed worldwide, accounting for 23% of all invasive cancer in female that year (Ferlay et al., 2010). To date, it is the most frequent cancer both in more developed and developing countries with an estimation of 690,000 new cases/year, but the rates of incidence are quite different after population size and age standardization: 71.7/100,000 for developed regions, and 29.3/100,000 for developing countries (Youlden et al., 2012). The incidence increases with age and the 89% of women with breast cancer are over 40 years (Ferlay et al., 2010). Highest incidence rates were recorded in Western Europe, Australia, Northern Europe and Northern America, where almost a 6% of women develop BC before 75 years age (Ellis et al., 2003) .Contrarily, the lowest incidence rates were observed in Eastern Africa, Middle Africa and Melanesia (Fig 1). Japan is the only developed country which showed a low incidence rate (Ellis et al., 2003). The discrepancy in incidence rates over the above-mentioned countries, is mainly due to social-cultural factors and lifestyle.

In the last decades, the achievement of an increasing quality of diagnostic and therapeutic technologies improved significantly the outcome of women with breast cancer in more developed countries.

Although the incidence of breast cancer is slightly increased, there is a significant decrement in the mortality rate, with an overall rate of mortality of 6-19/100,000. Almost 459,000 dead women for breast cancer were estimated worldwide in 2008 (Youlden et al., 2012). However breast cancer is the leading cause of cancer related mortality in women, representing almost 14% of all cancer deaths (Youlden et al., 2012). Although more breast cancer deaths occurred in less developed countries, the mortality rate was generally higher in developed countries (11.8 and 17.1 deaths per 100,000, respectively), as reported in Fig.1. Data from Italian register of tumours ranked BC first with an occurrence of 17% of all cancers death among females in the period 1998-2002 (http://www.registri-tumori.it/cms/?q=sede_mammella).

The rate of survival for women with breast cancer is higher than for other tumours, currently it is expected that about 20% of women with this disease die of it (Lester, 2010). In developed countries the five-year relative survival tended to be higher, between 95% and 90%, in North America, Australia, Japan and Northern Europe; while is lower in Eastern Europe and United Kingdom, between 75% and 80%. Limited data are available for Asia, Africa and South America (Youlden et al., 2012). In Italy, in 2007 was estimated a 5-year relative survival of 85% (84-85%) in all women (http://www.registri-tumori.it/cms/?q=sede_mammella).

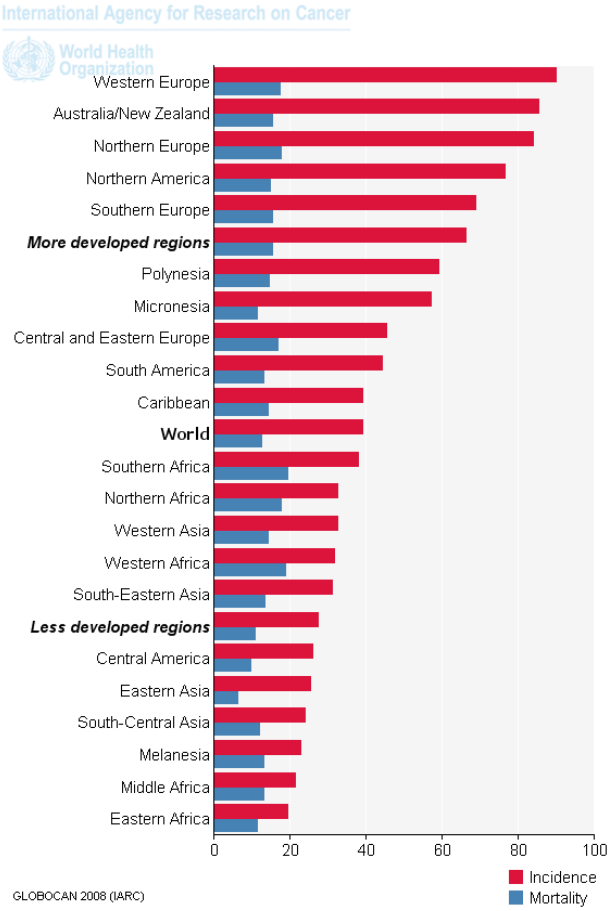


Fig. 1. Estimated age-standardized rates world (per 100,000) of BC incidence and mortality in female. Data from Globocan 2008.

1.1.2 Aetiology

Breast cancer is a multifactorial disease being related both to environmental and genetic factors. Epidemiological studies have shown that breast cancer is a disease of the affluent

societies (Ellis et al., 2003), which are characterized by the acquisition of the Western countries' lifestyle (Porter, 2008). Consequently, environmental and lifestyle factors are important risk factors. Studies on Asian migrants to United States showed an increased incidence rate across successive generations (Richie and Swanson, 2003). Women in developed country acquire risk factors, such as having few children or having them at older age, following high-caloric diet with meat and saturated animal fat and excessive consumption of alcohol (Youlden et al., 2012). Also a lower level of physical activity may contribute to increase the risk.

Gender represents the main discriminant factors, breast cancer occurs 100 times more frequently in women respect to men (Richie and Swanson, 2003), because in men the breast is a vestigial organ less sensitive to the hormones' action.

As for many other diseases, the higher risk factor is the aging process (Ferlay et al., 2010). BC is rare under 25 years, with an increasing incidence up to 50-60 years, then, there is a *plateau* and a recovery after 65 years. The likelihood of developing breast cancer is 0.4% in women younger than 39 years, 3.86% in women of age 40-59 years, 3.51% for the interval 60-69 years and 6.95% for women with an age of 70 or older (Vetto et al., 2009). This phenomenon probably reflect the menopause status and the changes in hormone cycle (Richie and Swanson, 2003).

Hormonal history is a significant risk factor. Early menarche (age less than 13 years), having no or few children, having an older age at first childbearing, lack of breast-feeding, menopause after age of 50 are all events considered to increase the risk of breast carcinoma since they lead to a greater exposure to hormones. Consequently in developed countries a decreasing rate of incidence among women aged 50-69 years in the last decade in parallel to the declining use of hormone replacement therapy has been observed (HRT) (Kumle, 2008; Richie and Swanson, 2003). Oestrogen and progestin together increase the risk more than oestrogen alone. A randomized controlled trial on postmenopausal women 'Womens' Health Initiative' (Rossouw et al., 2002) showed that the use of HRT increased the risk of develop a breast cancer of 53% for combination therapy, and 34% of increased risk for oestrogen alone, when HRT were administered for more than five years (Rossouw et al., 2002).

Approximately 10% to 15% of cases have a familial predisposition, while only a 5% of breast cancers are caused by hereditary syndrome (van der Groep et al., 2011), about 25% of

those cases can be referred to genetic mutation in two autosomal dominant genes: BRCA1 and BRCA2, which are tumour suppressor genes (Venkitaraman, 2002). Women with a first-degree relative (mother, sister, daughter) with breast cancer presented 2-fold higher risk of developing a carcinoma. There is a 5-fold increase risk for women with two first-degree relatives with breast cancer (Richie and Swanson, 2003).

Other risk factors include diet, as consumption of alcohol and obesity, since fat tissue produces hormones, which are converted to oestrogen.

Ionizing radiation exposure is another risk factor. The risk is increased in younger age and with increasing doses of radiation (Lester, 2010).

The last risk factor is represented by breast cancer itself: a percentage of 10% of patients with previous BC develop a second breast cancer, and the risk of developing a carcinoma in the contralateral breast is 3 to 7 times higher (Richie and Swanson, 2003).

1.2 Histological Classification

Most breast tumours are carcinomas, about 5% are squamous carcinomas, sarcomas, lymphomas and phyllodes tumours (Lester, 2010). Carcinomas are divided into *in situ* and invasive carcinomas. For *in situ* carcinoma the proliferation of neoplastic cells is confined to the ducts and lobules of the basal membrane. This type does not infiltrate the lymph vascular veins consequently it cannot metastasize. Conversely, invasive carcinoma has already exceeded the basal membrane invading the stroma, and the neoplastic cells can spread both to the regional or distant lymph nodes, or others sites.

In 1973 Wellings and Jensen were first to demonstrate that the vast majority of breast carcinoma originates from the epithelial cells of the terminal ductal lobular unit (TDLU), independently from histological type (Wellings and Jensen, 1973). Even if it is still common to describe the *in situ* and the infiltrating types as ductal or lobular, due to their similarity to ducts and lobules, the terms ductal and lobular do not indicate the site of origin of the carcinoma.

There are some differences in the localization sites of BC. The frequency of carcinoma in the left breast is slightly higher than the right, with a ratio of 1.07 to 1. Moreover, almost the

50% of tumours occur in the upper outer quadrant of the breast, while the lower inner quadrant is the less frequent localization site (Ellis et al., 2003).

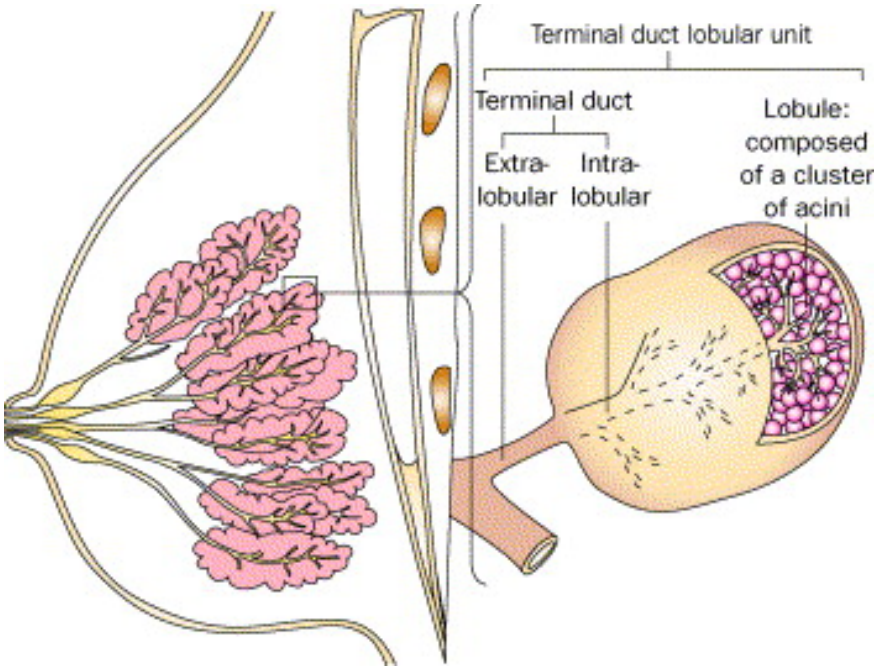


Fig. 2. Representation of the terminal ductal lobular unit (TDLU) of the breast. Taken from Davidson, T (Davidson, 2001).

The histological classification reported here is based on the 2nd edition WHO classification of malignant primary tumour of the breast (Ellis et al., 2003). The WHO classification recognizes 17 special histological types of BC, which represent the 25% of all invasive cancers (Ellis et al., 2003; Rakha and Ellis, 2011).

The *in situ* carcinoma is histologically classified as ductal (DCIS), which represent the 15-30% of all BC, and lobular (CLIS), with 1-6% of breast tumours.

There are several histological types for invasive BC. The invasive ductal carcinoma (IDC), not otherwise specified (NOS), is the commonest (40-75%) and comprises a heterogeneous group of tumours that cannot be classified as others specific histological type, so their morphological features are mixed. The invasive lobular carcinoma (ILC) represents 5-15% of invasive breast tumours. The incidence of ILC is reported to increase in postmenopausal women in the last two decades, which may be due to the use of hormone replacement therapy (HRT) (Ellis et al., 2003). At macroscopic analysis, ILC tumours show irregular and poorly delimited margins, which can be hardly to define because of the diffuse growth

pattern of cell infiltrate. Tumour cells appear individually dispersed through connective tissues, or they can appear arranged in single linear rows. This morphologic pattern was correlated with the lack of E-cadherin, responsible for cellular cohesion. The regional lymph node involvement is less frequent in ILC than IDC. Another histological type is the tubular carcinoma, which is characterized by a good prognosis. Unfortunately, this type accounts for only 2% of breast cancers. Its histopathological hallmarks are the presence of well-differentiated tubular structures with open lumina, surrounded by a single layer of epithelial cells. Conversely, medullary carcinoma, which accounts for 1-7% of all BC, is composed by poorly differentiated cells immersed in a few stroma, without glandular structures and high lymphocytic infiltrate. This type is frequently observed in hereditary tumours characterized by BRCA1 mutation. The mucin producing carcinomas comprise a variety of tumours. Mucinous carcinoma (also called colloid), for instance, is characterized by little group of small and uniform cells immersed in large amount of extracellular mucus. It accounts for 2% of all breast carcinomas and, generally, it shows a favourable prognosis. The neuroendocrine carcinoma accounts for 2-5% of breast carcinomas and they are characterised by the expression of neuroendocrine markers such as chromogranin and synaptophysin proteins in more than 50% of tumour cells. Less frequent histological types include invasive papillary carcinoma (1-2%), invasive cribriform carcinoma (~1%), apocrine carcinoma (~1%) and metaplastic carcinoma (<1%) (Ellis et al., 2003).

1.2.1 Relation of Benign Breast Lesions with Breast Carcinoma

Several epidemiological studies analysed the benign lesions and investigated on the risk of these lesion to evolve to BC (Beckmann et al., 1997; Richie and Swanson, 2003; Schnitt, 2003). Non proliferative-lesions as fibrosis or cysts do not increase the risk of breast carcinoma (Richie and Swanson, 2003). Otherwise, proliferative-lesions such as hyperplasia and papilloma are associated to a slight increase of risk. Proliferative lesions with atypia as Atypical Ductal Hyperplasia (ADH) and Atypical Lobular Hyperplasia (ALH) give a moderate increase in the risk of developing a carcinoma (Beckmann et al., 1997; Lester, 2010). The *in situ* carcinoma, if not treated, confers a high risk of developing an invasive one. (Schnitt, 2003). Women with a history of CLIS, ADH, ALH and atypical papilloma

show three to five times greater risk of breast cancer. Also the genetic mutations in BRCA1 or BRCA2 genes confer a risk of 60-80% of developing BC (Lester, 2010).

1.3 **TNM Staging**

The tumour, node, metastases system (TNM) is the most used staging classification for tumours. It was accepted from the International Union Against Cancer (UICC) and the American Joint Committee on Cancer (AJCC) and it is constantly revised and updated (Edge et al., 2010a).

The TNM system was designed to reflect the risk of distant recurrences and death after the local therapy, which was radical mastectomy and postoperative radiation. Consequently, the TNM staging for BC tries to supply a standard nomenclature to define the prognosis of BC patients, and also provide a clinical tool for the treatment decision (Edge et al., 2010a) (http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-staging).

There is clinical (c) and pathologic staging (p) for tumour, node and metastases sites. The clinical staging is based on information obtained before surgery or neo-adjuvant therapy; while the pathologic staging refers to information defined during surgery. In the following table the pTNM classification according to the 7^{TH} edition of AJCC is reported (Edge et al., 2010a).

Primary Tumour (T)		**Ragional Lymph nodes (N)**		**Distant Metastasis (M)**	
TX	Cannot be assesed	NX	Cannot be assessed	M0	No evidence of distant metastases
T0	No evidence of primary tumour	N0	No regional lymph nodes metastases	cM0(i+)	No evidence of distant metastases, but small number of cancer cells in circulating blood, bone marrow that are no larger than 0.2 mm

Tis	Carcinoma in situ (DCIS, LCIS, Paget's disease)	N1 (includes Nimi, N1a, N1b, N1c)	Cancer has spread to 1 to 3 axillary lymph node(s), and/or tiny amounts of cancer are found in internal mammary lymph nodes on sentinel lymph node biopsy	M1	Distant detectable metastases larger than 0.2 mm
T1 (includes T1a, T1b, T1c)	Tumour ≤20 mm in greatest dimension	N2 (includes N2a, N2b)	Cancer has spread to 4 to 9 lymph nodes under the arm, or cancer has enlarged the internal mammary lymph nodes		
T2	Tumour ≥20 mm but ≤50 mm in greatest dimension	N3 (includes N3a, N3b, N3c)	Metastases in 10 or more axillary lymph nodes, or in infraclavicular nodes, or in ipsilateral supraclavicular nodes, or clinically detected ipsilateral internal mammary nodes in presence of axillary positive nodes.		
T3	Tumour >50 mm in greatest dimension				
T4 (Includes T4a, T4b, T4c, T4d)	Tumour of any size with direct extension to the chest wall and/or skin				

Table 1. TNM classification according to the 7th edition of AJCC Cancer Staging Handbook (Edge et al., 2010a, b). the letter T followed by a number from 0 to 4 describes the tumour size and eventual extension to the skin or chest wall. Tis indicates the *in situ* carcinoma. The letter N followed by a number from 0 to 3 describes if the cancer has spread to lymph nodes near the breast, and how many of them are involved. The letter M followed by 0 or 1 indicates whether the cancer has spread to distant organs.

After the TNM have been estimated, this information is combined to define the anatomic stage. Non-invasive cancer is staged as 0, stage I represent the least advanced stage, while stage IV the most advanced (table 2).

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

Table 2. BC stages according to the 7th edition of AJCC Cancer Staging Handbook (Edge et al., 2010a)

1.4 **Histological Grading**

Histological grade is mainly used for invasive carcinomas and it provides important prognostic information. It refers to the cytological features of cancer cells in comparison to the normal ones. The grade is assessed through the analysis of three morphological features: the tubule formation, the nuclear pleomorphism and the mitotic count. Tubular formation has cut-off points of 75% and 10% (Ellis et al., 2003; Elston and Ellis, 1991). Nuclear pleomorphism is assessed through the nuclear size and shape of cancer cells in comparison to normal epithelia cells in adjacent breast tissue. For the identification of mitotic figures

only nuclei in clear morphologic features of metaphase, anaphase and telophase are counted. For each of these aspects, a score from 1 to 3 is given. Then, the total score is obtained adding single values (3-9) and the histological grading is defined as follow (Elston and Ellis, 1991):

- 3-5 points: well differentiated tumour, namely Grade 1 (G1);
- 6-7 points: moderately differentiated tumour, Grade 2 (G2);
- 8-9 points: poorly differentiated tumour, Grade 3 (G3);

Currently, histological grading is performed using the "Nottingham Grade" (Elston and Ellis, 1991). Although the clinical significance of grade 1 and 3 is well established, approximately 50% of tumours are classified as grade 2, which has equivocal clinical significance (Weigelt et al., 2010a). Histological grade has been included in validated prognostic algorithms to define BC therapy, as Nottingham prognostic index and Adjuvant Online! (Galea et al., 1992; Mook et al., 2009). The combination of histological grade and type provide complementary information (Rakha et al., 2008a) and a more accurate assessment of prognosis.

1.5 **Mechanism of Carcinogenesis**

Tumourigenesis in humans is a multistep process, in which the steps reflect the genetic alterations that lead the progressive transformation of normal cells to malignant ones. According to the general model of tumourigenesis of Hanahan and Weinberg, normal cells must acquire six news properties for becoming malignant (Hanahan and Weinberg, 2000). The abovementioned events are:

- self-sufficiency in growth signal, to provide mitogenic stimulatory signals of growth;
- insensitivity to anti-growth signals, to move the cell into an active proliferative status;
- evading apoptosis;
- limitless replicative potential, to circumvention the process of senescence;
- sustained angiogenesis, crucial for providing oxygen and nutrients;

- tissue invasion and metastases.

The molecular mechanisms and the main molecular actors, which lead to cell transformation, are largely unknown in BC. The molecular heterogeneity of the tumour reflects the concomitant development of different foci and witnesses that malignant transformation is the result of sequential multistep evolution (Beckmann et al., 1997). The development of BC is the result of a complex interaction between genetic predisposing factors, somatic mutations and hormones, which are playing a major role in several stages of the neoplastic transformation. BRCA1, for instance, is controlled by oestrogen and progesterone receptors. Hereditary carcinomas originate from cells with germline mutations, which alter DNA repair and normal apoptosis signals. Consequently, those cells need lower number of acquired mutations to transform into malignant ones. For example, both BRCA1 and BRCA2 genes are involved in the maintenance of DNA integrity. BRCA genes form complexes that will activate the repair machinery and also, start the homolog recombination. In hereditary carcinomas, the mutant allele of BRCA is inherited while the other allele is inactivated by a sporadic mutation. In major cases, the loss of BRCA1 is due to the loss of heterozygosity (LOH) in association to the methylation of promoter (Lester, 2010; van der Groep et al., 2011).

Concerning DNA mutations, in cancer there are two major classes of alterations: gain of functions or amplifications that activate the oncogenes; deletions or allelic loss that defect the tumour suppressor genes. Particularly in BC, the chromosomal sites which harbour oncogenes are marked by regions of DNA amplification, while point mutation of oncogenes are rare events. Beckmann *et al*. reported that although many oncogenes are reported to be involved in human cancer, only few oncogenes seem to be involved in BC development. These include three chromosome localizations: 8q24 (c-myc), 11q13 (int2, BCL2, EMS1, FGF3/4, CCND1), 17q12 (HER-2 or c-erbB2). It is important to stress that none of these oncogenes were found amplified in benign breast lesions. Actually, oncogenes amplification does not occur in the early stages of BC carcinogenesis, but it is first detected in DCIS. For example the overall rate of HER-2 amplification in DCIS is 33–61%, while it decrease towards the development of invasive ductal carcinoma (10–26%) (Beckmann et al., 1997).

However, women with atypical hyperplasia characterised by HER-2 expression have a seven-fold increased risk of develop invasive BC in comparison to cases without HER-2 amplification (Richie and Swanson, 2003).

The tumour suppressor TP53 appears constantly mutated in about 22-34% of sporadic BC, with a point mutation. Particular sites of LOH have been observed both in hereditary and sporadic breast cancer, but not in benign breast disease.

The first alteration observed in benign lesions is the increasing number of epithelial cells, which evade signals of growth inhibition and apoptosis. It seems that abnormal expression of hormonal receptors occurs even in this phase. Genetic instability is referred as a later event, since it appears in atypical dysplasia and *in situ* carcinoma. Evident aneuploidy and angiogenesis are observed subsequently in DCIS and invasive carcinoma (Beckmann et al., 1997).

Other alterations can be also frequently found in primary breast carcinomas, namely amplification of cyclin D1 and MDM2, loss of RB function, inactivation of p16 by promoter hypermethylation of 5-CpG island (Beckmann et al., 1997).

1.6 **Prognostic and Predictive Factors in Breast Cancer**

In BC, both pathological and molecular factors are commonly used in clinical routine to provide prognostic and predictive information for patients. Prognostic factors are intended to predict in objective and independent way the patients' clinical outcome, independently from systemic therapy. Predictive factors are able to predict the response to a specific treatment. Nevertheless, there is often an overlap between prognostic and predictive markers (Rakha and Ellis, 2011; Weigel and Dowsett, 2010). Moreover, the predictive markers can represent the target of therapy themselves.

Pathological factors have mainly a prognostic role, while molecular markers are generally used as predictive ones. There are many prognostic and predictive factors, which have been suggested for BC, but in this thesis only the most important and widely accepted are presented. The clinical-pathological and molecular factors have been incorporated into guidelines such as the European St. Gallen Consensus (Gnant et al., 2011; Goldhirsch et al., 2009; Goldhirsch et al., 2011), the American Society of Clinical Oncology (ASCO) guidelines (Hammond et al., 2010; Wolff et al., 2007), as well as in internet-based decision tool such as Adjuvant!Online (https://www.adjuvantonline.com) (Mook et al., 2009), to help clinicians evaluating the prognosis and the need for adjuvant therapy.

1.6.1 Traditional Morphological features

Traditionally, clinical-pathological features for patient's prognosis are: the TNM staging and the histological grade. Other pathological variables are: histological type, lymph vascular invasion, focality of the tumour, presence of *in situ* lesion, age of patients, family history and menopause status (Rakha and Ellis, 2011).

Tumour size

Tumour size is one of the strongest prognostic features, even after 20 years of follow-up (Arriagada et al., 2006). The size is closely related to the frequency of lymph nodes metastases. Increment in tumour size is accompanied by a higher probability of having lymph node involvement In fact, patients with tumours less than 1 cm in size show node positivity in 10% to 20%, at a size of 2 cm 40% of patients are nodes positive which reaches 50% at a size larger than 2 cm (Rakha and Ellis, 2011). Tumours size is also related to patients survival: tumour measuring less than 1.0 cm show a 10 years disease free survival rate of 90%, while it declines to 75% for tumours between 1-2 cm, and it falls to 60% for tumour of 2-5 cm (Kollias et al., 1997; Rakha et al., 2008a). The precise assessment of tumour size is fundamental to properly stratify patients for therapeutic treatments.

Lymph nodes

Loco-regional lymph nodes involvement is one of the most important prognostic factors in BC. Patients with positive lymph nodes have a 4-8 times higher mortality respect to lymph node negative ones (Soerjomataram et al., 2008). Prognosis depends also on the number of involved lymph nodes involved: the greater the number the poorer the survival of patients. It was also observed that 70% of patients with positive axillary nodes will develop recurrence, compared to 15-30% of lymph node negative. The involvement of nodes in the higher levels of the axilla, is related to a worse prognosis as well as the involvement of the internal mammary nodes. At diagnosis, approximately one third of operable BC patients present with positive nodes, of those 7% to 15% having more than 3 positive nodes. (Rakha and Ellis, 2011).

Histological Grade

Lower grade tumours are associated to better and longer survival, while higher grade ones showed a lower long-term survival. Cumulative 10-years survival, depending also on tumour size and nodal status, is about 90% for lower grade and 30-78% for highest score (Soerjomataram et al., 2008). Histological grade prognostic value is particularly important for the subgroups of BC in which adjuvant chemotherapy decision has to be made, such as LN-/ER+/HER2- or in patients with low volume LN metastatic disease (pN1), where the decision on the use of chemotherapy cannot be determined as for more advanced tumor stages (Rakha and Ellis, 2011).

Histological Type

Histological type provides four categories of prognosis: excellent, good, poor and very poor. Invasive cribriform, tubular, tubulo-lobular and mucinous have a excellent prognosis, showing more than 80% survival at 10 years. Tubular mixed, mixed ductal with special type, atypical medullary and alveolar lobular carcinoma present good prognosis with a 60–80% survival at 10-year. Invasive papillary tumours, classic lobular and medullary cancers have worse prognosis. Ductal, solid lobular, mixed ductal and lobular carcinoma have less than 50% survival at 10-year, while inflammatory presents 30% survival at 10-years (Soerjomataram et al., 2008). Unfortunately the majority of diagnosed BC has no special type, as invasive ductal carcinoma. As a consequence, the role of histological type in prognostication is still limited (Rakha and Ellis, 2011).

1.6.2 Traditional Molecular markers

Prognostic and predictive molecular markers routinely used in clinical practice are: oestrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER-2). The mitotic index Ki67 is not yet incorporated in clinic, despite its important value for measuring proliferation and categorizing patients into different group to help tumour classification (Rakha and Ellis, 2011).

Oestrogen receptor (ER)

ER α is the most important marker in BC, since it is an indicator of response to endocrine treatment. ER-positive tumours represent the vast majority of BC (Weigel and Dowsett,

2010). These tumours use the steroid hormone estradiol as the main growth stimulus. Nevertheless, ER is the direct target of endocrine therapy, which is performed directly using oestrogen agonists (i.e. Tamoxifen) or indirectly by blocking the conversion of androgens to oestrogen (i.e. aromatase inhibitors). ER status provides mainly predictive information, while its prognostic role is limited. In fact, ER-positive tumours are generally well differentiated and related to better outcome, but in long-term survival studies ER loses its prognostic value (2005). ER assessment defines two groups, the ER-positive and the ER-negative. Nevertheless in the first group, roughly half patients respond to hormone treatment (Rakha and Ellis, 2011; Rakha et al., 2010). ER-negative patients are addressed to other systemic treatments. It was also reported that ER-negative patients reached a more complete response with neo-adjuvant chemotherapy than ER-positive (Weigel and Dowsett, 2010).

Progesterone receptor (PR)

It is an oestrogen-regulated gene and its expression is considered an indicator of the function of ER pathway. About 40% of ER-positive tumours are PR-negative, lack of PR expression may indicate an aberrant growth factor signalling which could contribute to resistance to tamoxifen (Rakha et al., 2010). Studies in patients with metastases and in adjuvant trials with tamoxifen showed a better prognostic role of PR (Dowsett et al., 2006; Elledge et al., 2000). However tumours expressing PR but not ER are considered not true and need to retest the ER to exclude false negativity.

HER-2

The prognostic role of amplification of HER-2 in BC was first identified in 1987 (Slamon et al., 1987) . After the development of a humanised monoclonal antibody against HER-2 protein (trastuzumab, Herceptin), the assessment of HER-2 amplification became a strong predictive biomarker. Amplification of HER-2 gene and RNA/protein overexpression are correlated (Pegram et al., 2000). HER-2 amplification was found in 13-20% of BC, and approximately the 55% of these tumours are both ER and PR negative (Rakha et al., 2010). The main reason of assessing HER-2 in routine clinical practice is for treatment decision: trastuzumab could be addressed to patients with HER-2+ advanced disease and in the adjuvant setting for HER-2 early stage (Rakha et al., 2010). HER-2 status may also provide information for others systemic therapies, HER-2+ tumours are associated as well to resistance to endocrine therapies and poorer response to tamoxifen in ER+ and HER-2+

patients. An association of HER-2+ cancers and benefit from anthracycline-based chemotherapy was also demonstrated (Gianni et al., 2009; Muss et al., 1994)

Ki67

Ki67 is a nuclear non-histone protein and it is widely used as marker of proliferation, since it is universally expressed in proliferating cells and absent in quiescent cells. Ki67 is expressed during G1, S and G2 phases with a peak on mitosis, not in G0 phase of cell cycle. Many studies showed the ability of this marker as single variable to define good or worse outcome in BC patients' for low and higher Ki67 expression respectively (Cheang et al., 2009; Delpech et al., 2012; Sarode et al., 2011). However, contrasting results on the role of Ki67 in predicting patients' survival are present in literature, as reported by Stuart-Harris et al. who observed an association between Ki67 positivity and shorter survival (Stuart-Harris et al., 2008). As a consequence, Ki67 is not still recommended as prognostic marker for clinical routine (Weigel and Dowsett, 2010). Anyway, Ki67 is determinant not only for patients' outcome, but it can be used as predictor of response to chemotherapy. A recent study proposed a cut-off of 14% for Ki67 to distinguish luminal A from luminal B tumours (Cheang et al., 2009). According to this subdivision luminal A patients have been proposed to skip chemotherapy treatment because of the low level of proliferation (Goldhirsch et al., 2011).

1.6.3 Biomarkers under investigation

Tumour suppressor Retinoblastoma (RB) is a regulator of cell cycle and is inactivated in many human cancers (Manning and Dyson, 2012). RB functions as a transcriptional modulator through binding to the family of E2F transcription factors. The complex Rb-E2F acts as a growth suppressor and prevents progression through the cell cycle. Phosphorylation of the protein RB through the complexes of Cyclin-Dependent Kinases (CDK) and cyclins inhibits its activity. The majority of breast cancers are ER-positive at diagnosis and, in the process of tumourigenesis, epithelia mammary cells are ER dependent for proliferation. Retinoblastoma alterations in breast cancer includes loss of RB expression (20-35%) and loss of heterozygosity together with other alterations (7-35%) (Bosco and Knudsen, 2007). Some studies reported that loss of RB is correlated to advanced disease and often to ER-negative breast cancers (Jares et al., 1997). RB inactivity seems to be, also,

determinant for the resistance to hormone therapy (Bosco and Knudsen, 2007). Moreover, several approaches that inhibit Cyclin-Dependent Kinase 2 (CDK2) induced the arrest of phases S and G2, and determined induction of apoptosis (Shapiro, 2006). Given that inhibition of CDK2 and CDK1 may potentiate apoptosis induced by DNA-damaging agents that affect S-phase progression or by microtubule stabilizing agents, CDK inhibitors may be important in combination therapy (Shapiro, 2006).

Another widely investigated gene is p53, the most frequently mutated tumour suppressor gene in human cancer (Lai et al., 2012). The p53 regulates numerous cellular functions including cell cycle progression, apoptosis, senescence, cell motility, DNA repair, genetic instability and cell metabolism by activating transcriptionally several genes. The p53 gene is altered in 20-40% of breast tumours and its inactivation seems to be an early event in the tumourigenesis (Lester, 2010). The most frequent alterations are point mutations, which lead to the synthesis of a stable, malfunctioning and non-degradable protein, which is accumulated in the cells. Preclinical and clinical studies demonstrated that mutant p53 cancer cells are more resistant to drugs compared with those with wild-type p53. Therefore the role of p53 in mediating a chemotherapeutic response is complex and depends on both cellular context and type of therapy. (Lai et al., 2012)

The PI3K/AKT signalling cascade has emerged as a key regulator of several cancer phenotypes, including cell survival, proliferation, angiogenesis and metabolism (Chin and Toker, 2009). The PI3K pathway is the most frequently mutated pathway in BC, and it is characterized by mutation and/or amplification of PI3K catalytic subunits p110α (PIK3CA), p110β (PIK3CB) and the PI3K regulatory subunit p85α (PIK3R1) (Miller et al., 2011). The PIK3CA resulted mutated in approximately 30% of human breast cancers, particularly ER-positive tumours (Toker, 2011). The PI3K oncogenic mutation lead to the hyperactivation of its effector AKT, that in turn mediates downstream signalling by the phosphorylation of other effector proteins. Furthermore, also the deletion and loss of heterozygosity in the tumour suppressors PTEN and PHLPP (PH domain and leucine rich protein phosphatase), which inhibit the PI3K/AKT signalling, lead as well to hyperactivation of AKT (Toker, 2011).

AKT (also known as protein kinase B) is a serine/threonine protein kinase that plays a key role in multiple cellular processes, such as glucose metabolism, apoptosis, cell proliferation, transcription and cell migration. Recent studies have demonstrated different and non

redundant functions of AKT isoforms that depends on the cell and tissue type (Toker, 2011). There are three different isoforms of AKT (AKT1, AKT2, AKT3), which derived from 3 distinct genes located on three different chromosomes 14q32, 19q13 and 1q44, respectively (Chin and Toker, 2009; Toker, 2011; Wickenden and Watson, 2010). AKT1 and AKT2 are ubiquitously expressed in cells, while AKT3 is mainly distributed in neuronal tissue (Toker, 2011). AKT1 seems to play a key role in cell proliferation, survival and differentiation, by preventing the epithelial -to- mesenchymal transition (EMT). AKT2 signalling seem to be involved mainly in metabolism process. Nevertheless, AKT2 has been shown to enhance the migratory and invasive phenotypes of BC cells, by promoting cell adhesion, invasion and metastasis (Chin and Toker, 2009). AKT3 role remains unclear.

All AKT isoforms are activated by similar mechanisms through PI3K signalling (Fig.3). Upon stimulation with growth factors, PI3K synthesizes the second messenger phosphatidylinositol-3,4,5-triphosphate (PIP_3) from the precursor phosphatidylinositol-4,5-diphosphate (PIP_2). PIP_3 binds to the Pleckstrin Homology (PH) domain of AKT effectively recruiting it to the plasma membrane. The full activation of AKT is reached by phosphorylation of two residues (Thr and Ser) in the catalytic domain of the molecule. Thr308 (Thr309 for AKT2 and Thr305 for AKT3) is phosphorylated by PDK-1 (phosphoinositide-dependent-kinase-1). Whereas, the carboxyl-terminal residue Ser473 (Ser474 and Ser 472 for AKT2 and AKT3, respectively) is phosphorylated primarily by the mTORC2 (mammalian target of rapamycin complex 2) complex, even if others mechanisms have been suggested, such as PDK-1 phosphorylation, or AKT autophosphorylation. Once phosphorylated at these two residues, AKT translocated to distinct subcellular compartments and it is able to tranduces cellular signals, by phosphorylating numerous substrate proteins. Among the AKT substrates have been identified BAD, MDM2, GSK3β (Glycogen Synthase Kinase 3β) and FOXO (Forkhead transcription factors) (Wickenden and Watson, 2010).

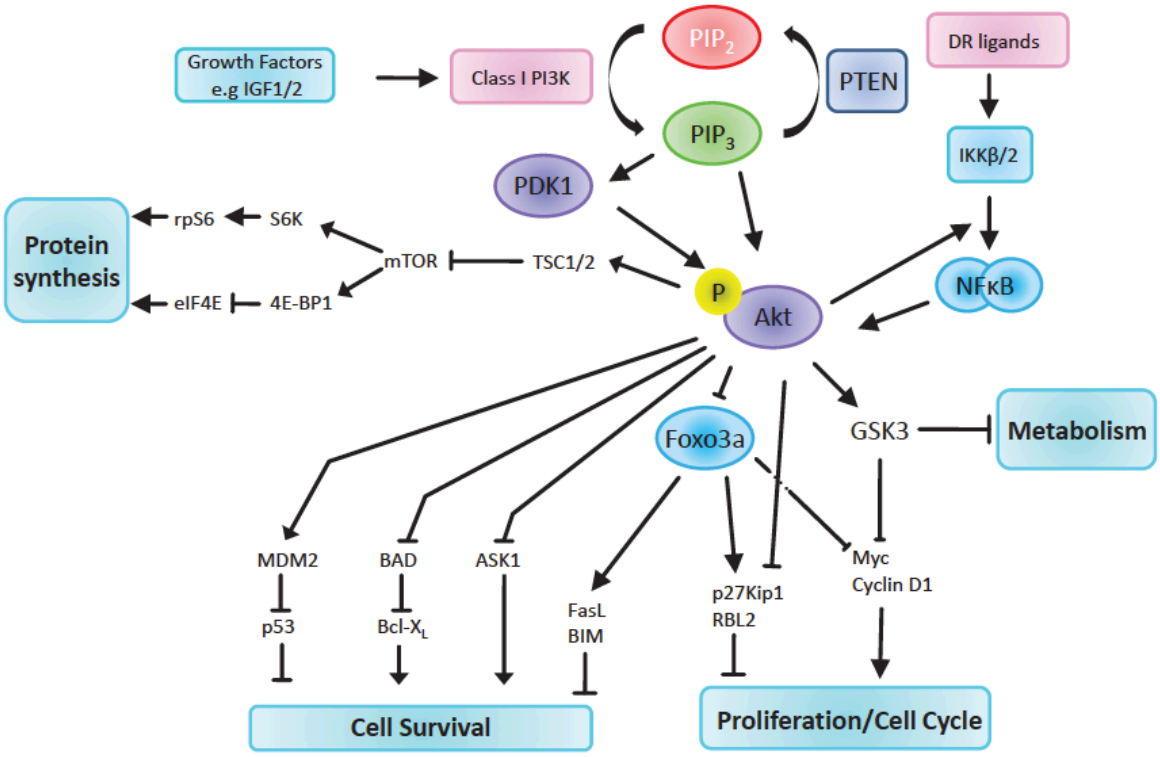


Fig. 3. PI3K/AKT signalling pathway is involved in fundamental cellular processes including protein synthesis, proliferation and survival. (DR, death receptor; FOXO, Forkhead family of transcription factors; GSK3, glycogen synthase kinase 3; IGF 1/2, insulin-like growth factor 1/2; IKKβ/2, inhibitor of kappa-B kinase β/2; mTOR, mammalian target of rapamycin; NF-κB, nuclear factor-kappa-B; P, phosphate; PDK1, 3-phosphoinositide-dependent kinase 1; PI3K, phosphatidylinositol 3-kinase; PIP2, phosphatidylinositol 4,5 diphosphate; PIP3, phosphatidylinositol-3,4,5-trisphosphate; PTEN, phosphatase and tensin homologue; S6K, S6 kinase; TSC1/2, tuberous sclerosis complex 1/2). Adapted from Wickenden and Watson (Wickenden and Watson, 2010).

Moreover, cytokeratins (CKs) constitute the largest family of intermediate filaments and are epithelial predominant in their expression (Yang et al., 1997). CKs provide a structural role in the cell, but they are also involved in signalling networks that regulate functions like cell cycle progression, cell size, cellular response to stress, protein synthesis and membrane trafficking (Iyer et al., 2013). During the process of malignancy the pattern of CKs expression is often maintained (Shao et al., 2012). For this reason the evaluation of CKs is performed in diagnostic for the detection of carcinomas (Shao et al., 2012; Yang et al., 1997). An important characteristic of these molecules is that their expression is regulated by

tissue type and differentiation. The epithelial cells of breast duct and lobules comprises two major type, namely luminal and basal/myoepithelial cells (Moll et al., 1982). Luminal cells form the inner layer that consists of differentiating cells, which are characterized by the expression of low weight CKs, as CK7, CK8, CK18 and CK19. The basal/myoepithelial cells, that represents the proliferating counterpart of breast epithelia (Iyer et al., 2013), are characterized by the expression of high weight CK, as CK5 and CK14 (Moll et al., 1982). In addition, the expression of CKs can be evaluated through immunohistochemistry, as surrogate marker of gene expression, for the characterization of BC molecular subtypes.

Vimentin is another member of the intermediate filament family of protein, whose expression in cancer has been correlated with accelerated tumour growth, invasion and poor prognosis. Vimentin has recently emerged as a marker of epithelial-mesencymal transition (EMT), a cellular reprogramming process in which epithelial cells acquire a mesenchymal phenotype that causes alteration in cell shape and motility. Increased expression of vimentin has been reported in several carcinomas, including breast cancer. Therefore, vimentin expression in cancer is likely to become a promising therapeutic target and has great potential for providing novel clinical prognostic tools. (Satelli and Li, 2011).

1.6.4 Gene Signature

In the past decade, the use of high-throughput methods has lead to BC genes signatures. Several commercial products have been developed by the use of the abovementioned signatures: MammaPrint® (Amsterdam The Netherlands), Oncotype DX® (Redwood city, CA, USA), Theros/MGI® (San Diego, CA, USA) and MapQuant DX/simplified (Stamford, CT, USA), as reported in Fig. 4. The aim of these prognostic molecular classifiers is double: to define subsets of patients with a good prognosis who could forgo chemotherapy, and identify individual patients who would benefit from chemotherapy. However, these signatures have proven to add limited information to prognostic model based on clinical-pathological parameters (Colombo et al., 2011).

MammaPrint® (van 't Veer et al., 2002) is a 70-gene signature for fresh frozen tissues. This signature is used to predict node negative BC patients' outcome within 5 years. This gene signature is currently being assessed in prospective phase III clinical trials (Cardoso et al.,

2008) and is approved by FDA as prognostic test for patients with age at diagnosis less than 61 years, with tumour less than 5 cm, lymph node negative and tumour stages I/II.

The other gene signature that is currently being tested in randomized prospective clinical trials is Oncotype DX® (Sparano and Paik, 2008). This assay is a quantitative Real Time PCR (qRT-PCR) assay of 21 genes (16 cancer-related and 5 reference genes) for formalin-fixed paraffin embedded (FFPE) tissues. Oncotype DX® can be used to predict outcome and to help therapy decision in ER- positive, node-negative patients. These signatures have limited applicability in lymph node positive and ER-negative BC patients (Weigelt et al., 2010a).

Theros® is qRT–PCR assay for FFPE tissues, which is commercialized by Biotheranostics. In this test the homeobox gene HOXB13 and the interleukin 17B receptor (IL17BR) are analysed since the expression ratios between HOXB13 and IL17BR is strongly correlated with the development of recurrences (Weigelt et al., 2010a). This test is recommended for lymph node negative and ER-positive BC patients, who are treated with surgery alone, to define risk of recurrence and benefit from endocrine therapy. However Theros® has not yet been shown to predict adjuvant hormone therapy or chemotherapy benefit and has not been included in the ASCO or NCCN guidelines for breast cancer treatment (Weigelt et al., 2010a)

MapQuant Dx™ Genomic Grade test is a 97-gene expression grade index (GGI) which characterizes high-grade vs low-grade tumours, through qRT-PCR. This test could stratify histological grade II tumours into GGI low-grade and GGI high-grade, of prognostic significance: histological grade II GGI low tumours had outcomes similar to GGI low cancers, whereas histological grade II GGI high tumours had outcomes similar to GGI high cancers. A high GGI was significantly associated with ‘excellent response to chemotherapy’ (Weigelt et al., 2010a). This test was recommended for ER-positive and LN-negative patients. This genomic grade test is available under the name MapQuant Dx® by Ipsogen, and it is now converted to eight-gene qRT–PCR test (Weigelt et al., 2010a).

Parameter	MammaPrint®	Oncotype Dx®	Theros®	MapQuant Dx®
Provider	Agendia	Genomic Health	Biotheranostics	Ipsogen
Assay	70-gene signature	21-gene recurrence score	HOXB13:IL17R ratio	8-gene genomic grade index
Platform	Microarray	Quantitative RT–PCR	Quantitative RT–PCR	Quantitative RT–PCR
Tissue type	Frozen	FFPE	FFPE	FFPE
Indication	Aged <61 years Stage I/II Node-negative Size ≤5 cm	ER-positive Node-negative	ER-positive	All (genomic grading)
Prospective trial	MINDACT	TAILORX	—	—
FDA approval	Yes	No	No	No
ASCO or NCCN guidelines	No	Yes	No	No

ASCO, American Society of Clinical Oncology; FFPE, formalin-fixed, paraffin-embedded; NCCN, National Comprehensive Cancer Network; RT–PCR, reverse-transcriptase polymerase chain reaction. Information retrieved from: www.agendia.com; http://www.eortc.be/services/unit/mindact/MINDACT_websiteii.asp; www.oncotypedx.com; www.biotheranostics.com; www.ipsogen.com; www.asco.org; www.nccn.org

Fig. 4. Summary of prognostic gene signature available in breast cancer. Taken from Weigelt B et al. (Weigelt et al., 2010a) .

1.7 Molecular Classification

In 2000 Perou et al. (Perou et al., 2000) by cDNA microarray analysis defined an "intrinsic gene list" to characterize BC tumours. The analysis was made on 38 invasive breast cancers (36 invasive ductal carcinomas and 2 lobular carcinomas), 1 *in situ* carcinoma, 1 fibroadenoma, 3 normal breast samples, and numerous replicates from the same patient. Hierarchical cluster analysis using this "intrinsic gene list" revealed the existence of four molecular BC subtypes, which were substantially different at transcriptomic level (Fig.5). The subtypes were: luminal, HER2+, basal-like and normal breast-like (Perou et al., 2000). A subsequent study performed by the same group on a larger cohort of patients, showed that the luminal group could be further divided into at least two groups: luminal A and luminal B (Sorlie et al., 2001). Moreover, Sorlie et al. demonstrated that breast tumours fall into two large molecular classes which correspond mainly to the status of oestrogen receptor expression: ER-positive and ER-negative. The difference between these two BC classes is consistent with the hypothesis that they derive from different tumour progenitor cells (Pfeffer et al., 2009).

The ER-positive (ER+) branch contains luminal subtype that expresses ER, cytokeratin 8 and 18, characterising luminal epithelial cells of the mammary gland. Luminal A cases are mostly low grade tumours and express low levels of proliferation-related genes; contrarily

luminal B tumours are usually higher grade and express more likely proliferation-related genes (Weigelt et al., 2010a).

The ER-negative (ER-) branch comprises more heterogeneous subtypes. The first is the HER-2 amplified subtype (HER-2+), which is characterised by the overexpression and amplification of HER2 gene and the absence of hormone genes expression (ER-, PR-). Basal-like tumours are ER, PR and HER-2 negative, but they express CK5/6 and CK14 typically present in the basal epithelia cells, and/or they could present the amplification of Epidermal Growth Factor Receptor (EGFR) (Rakha et al., 2008b). Normal breast-like tumours in the analysis of Perou, was characterized by the high expression of basal epithelial cells and adipose cells, with low expression of genes specific of luminal epithelial cells (Perou et al., 2000). To date the normal breast-like subtype is still under debate, because some considered it an artefact of sample representation (higher normal breast cells counterpart in frozen fresh tissues) (Colombo et al., 2011; Parker et al., 2009; Weigelt et al., 2010a).

In addition, other three molecular groups in ER-negative branch have been identified: the 'molecular apocrine', similar to the HER2+ subtype, but it showed activation of androgen receptor; the 'interferon-rich', characterized by higher expression of interferon-regulated genes; and the 'claudin-low' which comprises tumours that express low level of claudin mRNA, lack of expression of E-cadherin and higher expression for genes typically activated in the epithelial-mesenchimal transition (EMT) process (Colombo et al., 2011).

First Sorlie et al. in 2001 (Sorlie et al., 2001), and more recently a lot of other groups (Parker et al., 2009; Rakha and Ellis, 2011; Valentin et al., 2012; van de Rijn et al., 2002; Weigel and Dowsett, 2010; Weigelt et al., 2010a) demonstrated that these molecular subtypes are related to different outcomes. Among luminal tumours, the luminal A showed a better prognosis in comparison to luminal B. HER-2 and basal-like present an aggressive behaviour as well as worse prognosis. The clinical significance of normal breast-like tumours has yet to be determined.

The use of molecular classification by microarray is limited in clinical practice, because gene expression profiling is applied on fresh tissues, is expensive and is technically complex. For these reasons, investigators started to use surrogate markers to obtain molecular classification using immunohistochemistry (IHC). It has already been shown that data originating from mRNA expression profiling correlate with those obtained by IHC

(Gruver et al., 2011; Muller et al., 2011; Rakha and Ellis, 2011). However to date, a consensus on the definition of the panel of markers to be used in molecular classification has not been reached, as well as the clones of antibodies and the threshold cut-off for positivity (Geyer et al., 2012; Tang et al., 2009).

Some pathologists have used simply ER, PR, HER-2 and histological grade to classify breast tumours into luminal, HER-2 amplified and triple-negative (TN) (Muller et al., 2011; Wiechmann et al., 2009). The TN phenotype is commonly used as a surrogate for basal-like tumours. However, others have also used the proliferation markers Ki67 (Cheang et al., 2009) or the basal cytokeratins, to identify the basal-like subtype (Nielsen et al., 2004).

The PAM50 RT-qPCR assay is a single sample predictor for the molecular subtypes and may provide a way of incorporating the molecular subtypes in the diagnosis and prognostication of breast cancer patients (Parker et al., 2009). However this assay presented only a modest agreement between IHC/fluorescence in situ hybridization (FISH) defined HER2-positive cases and PAM50-defined HER2-positive cases (Weigelt et al., 2010a), that could lead to a possible misclassification of HER-2+ tumours.

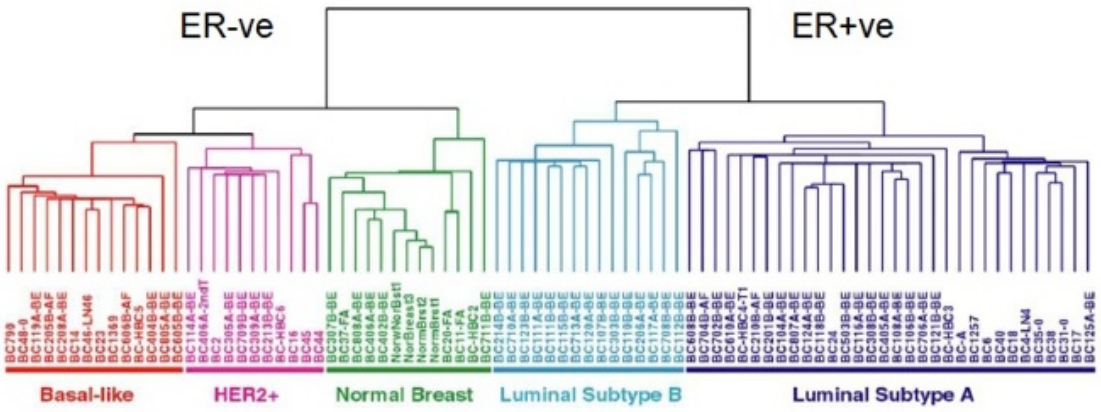


Fig. 5. Dendogram of the five major clusters belonging to BC molecular classification. The dendogram is divided in two main branches that represent the ER-negative (ER-ve) and ER-positive (ER+ve) tumours. Adapted from Sorlie et al. (Sorlie et al., 2001).

1.8 Therapeutic Treatment

The clinical management of breast cancer has been significantly changed over the last few decades. Most women with breast cancer are surgically treated, and sometimes they are

submitted also to other treatments such as radiation therapy, chemotherapy, hormone therapy, and/or targeted therapy. In the past, the first step for BC treatment was surgery, nowadays neo-adjuvant therapy has been widely accepted and utilised (Downs-Holmes and Silverman, 2011). Here an overview of the available treatments for BC patients has been reported, according to National Comprehensive Cancer Network (NCCN) (http://www.nccn.org/professionals/physician_gls/f_guidelines.asp) and American Cancer Society (ACS) (http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-treating-general-info) guidelines.

Surgery

The main goal of surgery is the removal of the cancer from the breast and the assessment of tumour stage. Surgery includes breast conserving treatment and mastectomy. Breast conserving surgery is not recommended when the tumour presents extensive involvement, or in case of inflammatory component, or multifocal disease (Downs-Holmes and Silverman, 2011). On the other side, mastectomy does not give better long-term survival if compared to breast conservation surgery (Veronesi et al., 2002). In women with no evidence of lymph node involvement before surgery sentinel lymph node biopsy has been used (Veronesi et al., 2003), but recently oncologists are considering the possibility of abolish the SLN biopsy. For many years sentinel lymph node (SLND) and axillary lymph node dissections has been the surgical standard for patients with positive sentinel lymph nodes (SLNs). However, since removal of many lymph nodes increases the risk of lymphedema after surgery, SLND without axillary dissection is suggested for patients with one or two positive sentinel nodes and tumours less than 5 cm (Downs-Holmes and Silverman, 2011). Axillary lymph node dissection (ALND) is mandatory for patients submitted to mastectomy (http://www.nccn.org/professionals/physician_gls/f_guidelines.asp), for patients with voluminous axillary disease, for tumours greater than 5 cm, after neo-adjuvant treatment, and in cases in which partial breast radiation or no radiation is planned (Downs-Holmes and Silverman, 2011).

Radiation therapy

It is used to destroy cancer cells remaining in the breast, chest wall, or underarm area after breast-conserving surgery. Radiation may also be needed after mastectomy in patients with either a cancer that is larger than 5 cm in size or when cancer is found in more than 3 lymph nodes (http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-treating-general-info; Truong et al., 2005). There are two types of radiation: external beam radiation, the usual for BC, and internal radiation therapy, also known as brachytherapy. In external radiation therapy the whole breast is usually irradiated but, depending on the size and extent of the cancer, it could also be irradiated the chest wall as well the underarm area. Internal radiation therapy consists of a radioactive substance sealed in needles, wires or catheters, which are placed directly into or near the cancer. The way the radiation therapy is given depends on the type, stage, and location of the tumour being treated.

Systemic therapy

Systemic therapy includes chemotherapy, hormone therapy, targeted therapy. Systemic therapy could be given as neo-adjuvant or adjuvant treatment. Neo-adjuvant treatment, before surgery, is administered to patients to improve operability for locally advanced disease and the chance of breast conservation for larger tumours (Downs-Holmes and Silverman, 2011). Systemic therapy given to patients after surgery is called adjuvant therapy. In this case, after surgical removal, adjuvant therapy is used to kill any undetected tumour cell that may have been left behind or migrated to other parts of the body. Moreover, systemic therapy is the main treatment for patients with metastatic BC who cannot underwent surgery.

The benefit of chemotherapy is dependent on multiple factors, such as tumour size, lymph node involvement, ER, PR, HER-2 status. Many combinations of drugs are being used. In early BC the most common regimens comprises cyclophosphamide, methotrexate, fluorouracil, doxorubicin, epirubicin, paclitaxel, and docetaxol. These and other chemotherapeutical drugs can be used in metastatic setting.

ER positive patients may be submitted to hormone therapy. This therapy is mostly used as adjuvant therapy to help reduce the risk of the recurrences after surgery, but it can also be given as neo-adjuvant treatment. There are two classes of hormonal agents for BC: selective ER modulators (SERMs) and aromatase inhibitors (AIs) (Downs-Holmes and Silverman, 2011). Tamoxifen and toromifene are SERMs, they prevent the binding of oestrogen to BC cells and are mainly given to premenopausal women. Conversely, letrozole, anastrozole and

exemestane are AIs. This class of drug can be used both in early and advanced disease, but it is effective only for postmenopausal women. AIs block the aromatase enzyme in fat tissue, which is responsible for the production of small amounts of oestrogen in postmenopausal women. It was demonstrated that postmenopausal women receive more benefit from AIs than tamoxifen in the adjuvant setting (Downs-Holmes and Silverman, 2011). Other options for patients who do not respond to tamoxifen or AIs are Fulvestrant and Androgens.

'Targeted therapy' refers to drugs that specifically target some molecules, with fewer side effects. Tratuzumab is a monoclonal antibody that targets the HER2 protein in breast tumors and offers a survival benefit for women with breast cancer that overexpresses HER-2, both in early and metastatic BC (Slamon and Pegram, 2001; Slamon et al., 2001). A recent study showed that the addition of trastuzumab to the standard chemotherapy in early BC overexpressing HER-2 reduced the risk of recurrence and death of 52% and 33%, respectively, in comparison with chemotherapy alone (Romond et al., 2005). In 2006 FDA approved trastuzumab for treatment of positive HER-2 breast tumours. Lapatinib is another drug which targets HER-2 protein, and has been found effective in patients with advanced disease resistant to trastuzumab (http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-treating-general-info).

2 AIM OF THE STUDY

Breast carcinoma encompasses a heterogeneous group of tumours with a great variability at the morphological and molecular level as well clinical outcome (Rakha and Ellis, 2011). In spite of the recent advances both in the knowledge of molecular mechanisms leading to BC initiation and in clinical management, a significant number of patients presenting similar histological and pathological features can show very distinct outcomes. Therefore the aim of this thesis is the identification and validation of new prognostic biomarkers in a retrospective case study of patients affected by primary breast carcinoma with a long follow-up.

In its first part this thesis investigates the role of BC molecular classification in patients' prognosis and prediction of clinical outcome, by comparing the traditional pathological factors with the emerging molecular classes. Molecular classification was assessed by means of immunohistochemistry, using seven surrogate markers (ER, PR, HER-2, Ki67, CK8, CK5/6 and vimentin), in a case study of 305 patients, divided in lymph node negative (LN-) and lymph node positive (LN+). Molecular characterization was also performed in some regional metastatic lymph nodes, in order to evaluate the phenotype of the cells that colonise the lymph nodes in the process of metastatization.

In the second part gene expression analysis of nine candidate prognostic biomarkers was performed. The genes of interest belong to the RB pathway (RB and CDK2), RAS pathway (HER-2, PI3K, AKT1, AKT2, AKT3, RAF1) and cellular differentiation mechanism (CK8). Gene's expression was investigated in the entire cohort of primary tumours of the case study. Genes were chosen to confirm our previous studies (Bonin et al., 2008; Bonin et al., 2006), which revealed that seven genes seemed to be involved in patient's prognosis. In the present thesis the potential prognostic role of 9 candidate genes was investigated at mRNA level through Real Time PCR. Comparison of genes expression was performed with respect to lymph node involvement and molecular classification to better characterise biological heterogeneity and clinical outcome of the molecular subtypes.

In both parts of the thesis, molecular analysis was performed on formalin fixed and paraffin embedded (FFPE) tissues specimens, in order to directly compare experimental results with clinical data.

3 MATERIALS AND METHODS

3.1 Patients and Sample Collection

The retrospective study carried out in this project was performed on a cohort of young women comprising both lymph node negative (LN-) and lymph node positive (LN+) breast cancer patients. The criteria for patients' selection were: i) women who had stage I-III invasive BC, ii) age at diagnosis 55 years or younger, iii) absence of a second primary breast cancer or other malignancies, iv) No anticancer treatment prior surgery, v) and 15-25 years of follow-up-information if alive. Women were all resident in the North-Eastern province of Trieste and had a diagnosis of primary breast cancer established between 1st January 1983 and 31st December 1993. By applying these criteria 380 women were enrolled. Patients with *in situ* BC and cases without tissue availability for microdissection of tumour (i.e. absence of tumour block or needle aspiration biopsies) were excluded from the study. Therefore, 305 patients were included in the study, 151 LN- and 154 LN+. FFPE tissues of the primary tumour referred to surgical treatment and a representative number (100) of loco-regional metastastic lymph nodes were used for the analyses. Patients were treated with radical mastectomy, modified radical mastectomy or breast-conserving surgery. All patients submitted to conserving surgery were treated with radiotherapy. In detail, in patients LN- or LN+ with 3 or less lymph nodes involved, only the breast was irradiated; on the other hand in patients with more than 3 metastatic lymph nodes and capsule invasion, breast and axilla were irradiated. All LN+ patients were treated with adjuvant chemotherapy with CMF (cyclophosphamide, methotrexate, fluorouracil) or EC/ECF (epirubicin, cyclophosphamide, fluorouracil) regimens according to standard protocols. ER-positive patients, both LN- and LN+, were submitted to hormone therapy with Tamoxifen.

Cases were enrolled before the collection of follow-up data. Patients were followed up from the date of diagnosis of BC to death or until 31 December 2008, whichever came first. Clinical end-point of the study was overall survival (OS) that was defined as the time from surgery to breast cancer specific death. The cohort of patients was followed through the general population-based Friuli Venezia Giulia Cancer Registry.

Clinical information such as type of surgery, tumour stage, TNM classification, presence of recurrence, patient's status and cause of death were obtained from medical records.

3.2 Histological Revision

Hematoxylin-eosin (H&E) staining was performed on 4 μm-thick FFPE tissue sections for all the case study according to the standard procedure (Melis, 1975). H&E slides were reviewed by a pathologist. Histological revision was performed to evaluate the histological type of tumour according to the World Health Organization (WHO 2003) (Ellis et al., 2003), and histological grade according to Elston and Ellis grading system (Elston and Ellis, 1991).

Finally, the pathologist defined for each H&E slides, both primary tumour and metastatic lymph node, the representative tumour areas that were later used for molecular analysis.

3.3 Microdissection

3.3.1 Manual Microdissection

Manual microdissection was performed to isolate tumour from perilesional tissue in cases where the amount of perilisonal tissue was higher than 15% of the total amount. The areas of interest were identified on a reference H&E-stained section and then mechanically microdissected on the paraffin block For each sample, five-twenty consecutive unstained 5 μm-thick sections were cut, depending on the size of tissue area (Faoro and Stanta, 2011). Later, dissected specimens were deparaffinised in xylene and rehydrated in ethanol for nucleic acids or protein extraction (Bonin and Stanta, 2011; Stanta et al., 1998).

3.3.2 Tissue Microarray (TMA)

Microdissection of tumours was performed with the Tissue Microarray (TMA) technique both for immunohistochemical analysis and RNA extraction. The procedure of TMA construction was similar for these purposes, the difference was limited to the needle diameter involved: 1 mm for IHC assay and 3 mm for RNA extraction. TMAs were constructed using the FFPE tissues of the entire cohort, dissecting from the selected tumour areas that were marked on the H&E slides.

For IHC purposes, tissue cylinders of 1 mm diameter were taken from the selected regions of the donor paraffin block and were punched into a recipient paraffin block using a tissue-arraying instrument (Galileo TMA CK3500, Integrated Systems Engineering, Milano,

Italy). Regarding primary tumours, 8 tissue arrays containing 384 spots were built to assemble the 305 BC patients. In detail one representative spot for each patient was taken. Multiple tissue cores were sampled if: the tumour presented histological heterogeneous regions, or the tissue spot was not evaluable in subsequent IHC. For the analysis of metastatic lymph nodes, one representative spot was sampled. Accordingly, three TMAs with 102 tissue core were assembled. Once prepared, TMA was placed upside-down onto a glass slide and incubated at 40°C for about 45 minutes to allow binding of the donor cores to the paraffin wax of the block (Faoro and Sapino, 2011). The glass slide was used to level the block surface by gently pushing the cores into the block. After cooling, 4 µm thick sections were cut, mounted on Superfrost® Plus (Thermo Scientific) microscope slides and heated at 37°C overnight for subsequent IHC analyses.

For RNA extraction, 3 mm tissue core was taken from areas adjacent to the previously extracted punches and inserted into a new recipient block, for each tumour. After melting paraffin wax and core tissues and subsequent cooling as above mentioned, 20 sections of 5 µm thickness were cut for each punch and placed into a new 1.5 ml tube.

3.4 Molecular Analysis at Protein Level

3.4.1 Immunohistochemical Staining

The IHC staining was performed on 4 µm-thick tissue sections of the TMA blocks. The antibodies used and the experimental conditions are listed in the table below (Table 4). The immunostaining procedures were performed manually using the Vectastain Universal Elite ABC kit (Vector Laboratories) for the antibodies: ER, PR, HER2, cytokeratin 8 (CK8), cytokeratin 5/6 (CK5/6) and vimentin. Incubations were performed in a humidified chamber. Briefly, tissue sections were deparaffinised in xylene for 30 minutes and then hydrated in a decreasing alcohol series (100%, 96% and 50%). Endogenous peroxidase activity was blocked by incubating the tissue sections in 0.3% hydrogen peroxide solution (Sigma-Aldrich) for 10 minutes. Heat mediated antigen retrieval was performed in water bath at 120°C, with low and high pH buffer for 20 minutes. To avoid unspecific binding, the sections were incubated for 20 minutes with blocking serum (Vectastain Universal Elite ABC kit, Vector Laboratories). Incubation with the primary antibody was performed for 1 hour as described in Table 3. The slides were washed three times for 3 minutes each in PBS and 0.1% Triton X-100 (PBST), incubated for 60 minutes with the biotinylated secondary

antibody and for 30 minutes with Vectastain ABC system. Slides were washed three times in PBST. For the visualization Dab Substrate kit for Peroxidase (Vector Laboratories) was used. Finally, sections were counterstained with Mayer's Hematoxilin (Bio Optica). Positive and negative controls were used in each IHC assay. Positive control slides from BC tissues were employed for ER, PR, Her2 and Ck8 staining; while slides obtained from tonsil was used for Ck5/6, Vimentin and Ki67 antibodies. Negative controls were prepared using the same TMAs slides, without the primary antibody incubation.

Immunostaining for Ki67 and HER-2 were performed in Lab Vision Autostainer 480S (Thermo Scientific) with the UltraVision LP Large Volume Detection System HRP Polymer (Lab Vision Corporation, Thermo Scientific) according to manufacturer's protocol. In brief, after deparaffinisation and rehydration steps, tissue sections were washed two times in 0.05M TBS + Tween 20 (Bio-Optica) buffer. Tissue slides were treated with high pH antigen retrieval for 20 minutes in heated water bath and washed 4 times in buffer. To reduce non-specific background staining, endogenous peroxidase was blocked with hydrogen peroxide for 10 minutes. After 4 washes, the primary antibody was applied; subsequently tissue sections were incubated with primary antibody enhancer for 20 minutes at room temperature. After washing, the HRP Polymer was applied for 30 minutes at room temperature. For visualization the chromogen 3,3' Diaminobenzidine (DAB) (LabVision, Thermo scientific) was used. The sections were counterstained with Mayer hematoxylin as usual.

Primary antibodies	Clone source	Localization	Antigen retrieval	Diluition	Incubation
ER	SP1 Ventana	nucleus	0.1 M Tris-Borate 1mM EDTA pH8 (120° C water bath, 20 minutes)	PD	1 hour RT
PR	1E2 Ventana	nucleus	0.1 M Tris-Borate 1mM EDTA pH8 (120° C water bath, 20 minutes)	PD	1 hour RT
HER-2	CB11 Thermo Scientific	membrane	0.1 M Tris-Borate 1mM EDTA pH8 (120° C water bath, 20 minutes)	1:300	30 minutes RT
Ki67	MIB1 Dako	nucleus	0.1 M Tris-Borate 1mM EDTA pH8 (15 minutes MW)	1:200	30 minutes RT
CK 8	M20 Abcam	Cytoplasm/ membrane	10 mM buffer Citrate pH6 (120° C water bath, 20 minutes)	1:250	1 hour RT
CK 5/6	D5&16B4 Aczon biotech	Cytoplasm/ membrane	0.1 M Tris-Borate 1mM EDTA pH8 (120° C water bath, 20 minutes)	1:100	1 hour RT
Vimentin	V9 Ventana	Cytoplasm/ membrane	10 mM Citrate buffer pH6 (120° C water bath, 20 minutes)	PD	1 hour RT

Table 3. Primary antibody used for IHC analysis. PD=prediluted, RT=room temperature

3.4.2 Evaluation of Immunohistochemistry and Definition of Molecular Classes

Immunostaining was evaluated using the light microscopy, counting the positive cells across three high power fields (HPF) at 40X magnification by two different observers in a blinded fashion. The mean values were used for analyses.

Semi-quantitative evaluation of each antibody staining was scored according to the intensity of the signal and by assessing the percentage of immunoreactive tumour cells in a high magnification field over the total number of tumour cells in that magnification field. The intensity of the signal for each antibody was categorized into 4 groups (0-3), namely no expression, weak expression (1+), moderate expression (2+), or strong expression (3+).

Tumours were considered positive for ER and PR if more than 10% of tumour nuclei were stained (Li et al., 2011; Muller et al., 2011). HER-2 expression was scored according to ASCO guidelines using the following criteria: 1+, a faint and incomplete membranous

staining for HER-2 in more than 10% of cells was considered as negative; 2+, a complete membranous staining, either non-uniform or weak in at least 10% of cells; score of 3+ when complete membrane staining intensity in more than 30% of cells was present (Wolff et al., 2007). HER-2 results were considered positive in cases of 3+ immunohistochemical staining. A score of 2+ was interpreted as equivocal and automatically analysed for gene amplification through silver *in situ* hybridization (SISH), as described in the next paragraph. Tumours positive in more than 10% of cells were defined positive for CK8, CK5/6. Vimentin was considered positive if more than 1% of cancer cells were stained. For Ki67 samples were scored by counting the percentage of positive cells and stratified into two groups according to their proliferation activity: 0-14% were classified as low proliferation index (<14%), and ≥14% as high proliferation index, according to St. Gallen consensus (Cheang et al., 2009; Goldhirsch et al., 2011).

The tumours were then classified into four main molecular classes according to the staining profile of the antigen markers: ER, PR, HER-2, CK8, CK5/6, Ki67 (Cheang et al., 2009; Dawood et al., 2011; Goldhirsch et al., 2011; Gruver et al., 2011; Park et al., 2012). In detail, cases:

- ER+ and/or PR+, HER-2-, low Ki67, CK8+ were classified as Luminal A;
- ER+ and/or PR+, HER-2-, high Ki67 or ER+ and/or PR+, HER-2+, any Ki67 were classified as Luminal B;
- ER-, PR-; HER-2+, any Ki67 were defined HER-2 type;
- ER-, PR-, HER-2-, CK5/6+, any Ki67 were classified as basal-like.

The expression of vimentin, was studied in the entire cohort.

Tumours which did not meet the above mentioned panel criteria, were defined as ‘unclassified’.

In Fig. 6 are shown the pictures of positive staining for each primary antibody.

3.4.3 Silver In Situ Hybridization (SISH)

It is used to quantitatively detect HER-2 gene amplification for cases classified as 2+ in IHC. SISH analysis was performed using the *ultra View* SISH DNP Detection Kit (Ventana Medical System) in Benchmark XT automated slide stainer instrument, according to manufacturer’s instruction. The Pathway HER-2/neu DNA Probe (Ventana) and HER-2/neu

Region of Chromosome 17 (Ventana), both labelled with dinitrophenol (DNP) were used. In brief, slides of microarray block were deparaffinised in xylene, and subsequently rehydrated in graded alcohol. Afterwards, tissue sections were pre-treated with heating and/or enzymatic digestion. For denaturation, slides were heated to 95° C for 12 minutes and hybridization was performed at 52° C for 2 hours. After hybridization, appropriate stringency washes with 2X SSC (0.3M NaCl, 0.03M Na citrate pH 7.0), 50% formamide were performed 3 times at 59°C. HER-2 DNA probe and Chromosome 17 (Ventana), both labelled with dinitrophenol (DNP), were detected using the Rabbit anti-DNP primary antibody (*ultra View* SISH DNP Detection Kit) incubating at 37°C for 20 minutes. Evaluation of signal was carried out using a goat anti-rabbit antibody conjugated to horseradish peroxidase (HRP) for 16 minutes at 37° C. The chemistry of the SISH reaction is driven by the sequential addition of silver acetate, hydroquinone and H_2O_2. The silver ions (Ag+) are reduced by hydroquinone to metallic silver atoms (Ag). The silver precipitation is deposited in the nuclei and a single copy of the HER2 gene is visualized as a black dot. Finally, slides were counterstained with hematoxylin.

Counting was performed as recommended by the ASCO/CAP guidelines (Wolff et al., 2007), a copy number greater than 6 or HER-2 gene/chromosome 17 copy number ratio higher than 2.2 was considered positive. A representative picture of SISH for the two probes is reported in Fig. 7.

3.5 Molecular Analysis at RNA level

3.5.1 RNA Extraction from FFPE

Total RNA was extracted from FFPE specimens of primary breast carcinomas using a proteinase K-based protocol (Bonin and Stanta, 2011; Stanta et al., 1998). Sections were then deparaffinised in xylene for 30 minutes, and rehydrated in graded ethanol (100%, 96% and 50%) for 2 minutes twice. Samples were then digested in 150-400 µl of RNA digestion buffer containing 6 mg/ml proteinase K, 1.12 M Guanidine thiocyanate, 20 mM Tris HCl pH 7.5, 0.5% N-Lauroyl Sarcosine, 40 mM β-mercaptoethanol at 55°C overnight in a thermomixer (Eppendorf). Total RNA was purified by acid phenol/chloroform (Sigma-Aldrich) extraction followed by ethanol precipitation. Total RNA was resuspended in an appropriate volume of DEPC treated water (15 - 30 µl, depending on the amount of starting

tissue). Total RNA concentration was assessed through spectrophotometer reading (Amersham Biosciences). RNA extracts were stored at -80°C in aliquots.

3.5.2 DNase Treatment and Reverse Transcription (RT)

To avoid false positive caused by genomic DNA (gDNA) amplification in RT-PCR-based analysis, gDNA should be removed by DNase digestion. DNase treatment was performed in RNA extracts as already reported (Dotti and Bonin, 2011). Briefly, 4 µg of total RNA were digested with RNase free DNase for 15' at 25°C in 20 µl final volume containing 5U of DNAse I (GE Healthcare) and 1X DNase buffer (40 mM Tris-HCl, pH 7.5, 6 mM $MgCl_2$). The enzyme was blocked with 2 µl of 25 mM EDTA and heat inactivated at 65°C for 10'.

DNase treated RNAs were reverse-transcribed into cDNA using Moloney Murine leukemia virus (MMLV) reverse transcriptase and random hexamers as already reported (Nardon et al., 2009). In detail, 1.2 µg of total digested RNA was added to 3.35 nmoles of random examers in a final volume of 9 µl. The mixture was incubated at 65°C for 10 minutes and then immediately chilled on ice. At this point 11 µl of the RT mixture were added, yielding a final concentration of 1X First Strand Buffer (50 mM Tris-HCl pH 8.3; 75 mM KCl; 3 mM $MgCl_2$ - Invitrogen), 10 mM DTT (Invitrogen), 4 units of Rnase Inhibitor (Promega) 4.5 mM $MgCl_2$, 1mM dNTPs (Amersham) and 250 units of MMLV enzyme (Invitrogen). The mixture was left at room temperature (25°C) for 10 minutes to allow the annealing of the random hexamers, then reverse transcription was carried out at 37°C for 50 minutes. Subsequently, enzyme was blocked by heating at 70°C for 10 minutes. Complementary DNA was stored at -20°C in aliquots.

3.5.3 Gene Expression

Genes of interest are: RB and CDK2 (RB pathway), HER-2, PI3K, AKT1, AKT2, AKT3, RAF1 (RAS pathway) and CK8. Gene expression analysis was performed by real-time PCR using the TaqMan chemistry in the instrument Mastercycler® ep Realplex (Eppendorf, Hamburg, Germany).

For every target gene intron-spanning primers were designed. Amplicon lengths were between 60 and 100 base pairs. Primer and probe sequences were built by the use of Primer Express software (Applied Biosystem, Darmstadt, Germany) and PrimerQuest software

(Integrated DNA Technologies). The details regarding primer/probe sequences and reaction conditions are reported in Table 4. Amplification conditions (temperature of annealing and $MgCl_2$ concentration) were set up using a cDNA pool of the case study. Subsequently, a calibration curve was made for each gene using the cDNA pool with six 1:4 serial dilutions, starting from 250 ng of cDNA. Efficiencies of real time amplification, for each gene, were checked plotting Ct values of PCR amplified serial dilutions of cDNAs, against the log10 of the theoretical initial RNA input. Efficiency was defined as $[10^{\wedge(-1/slope)}-1] \times 100$, where the slope is obtained from the linear regression line fitted thought the determined points. To correct for quantification errors depending on differences in sample-to-sample RNA quality, β-Actin was chosen as reference gene for normalization.

Each PCR assay was performed in duplicate using the JumpStart^{TM}Taq ReadyMix^{TM} For Quantitative PCR (Sigma-Aldrich, St. Louise, USA) according to the manufacturer's instructions. In each run respectively 30 ng (β-ACTIN, CDK2, HER-2, AKT3), 40 ng (RAF1), 50 ng (AKT2), 60 ng (RB, PI3K, AKT1) of cDNA were used in a final reaction volume of 20 μl, containing the JumpStart Ready Mix, the ROX reference dye, the adjusted $MgCl_2$, the forward and reverse primer (0.3 μM), the MGB probe labelled with FAM dye (0.2 μM). The cycle condition were: one step of 10 minutes at 95°C for denaturation and polymerase activation, followed by 45 cycles of two-step program of 15 seconds at 95°C for denaturation and 1 minute at 60°C for annealing/extension. To exclude contamination, negative controls without cDNA were included in each assay. RNA extracted from the breast cancer cell line MCF-7 was used as positive control.

Genes expression levels were normalized against the chosen housekeeping gene and expressed as ratio according to the ΔΔCt model as previously reported (Pfaffl, 2001).

Gene	Primer Sequences	Product size (bp)	$MgCl_2$ (mM)	PCR efficiency
β-ACTIN	F: CGGCCCCTCCATCGT R: AAAGGGTGTAACGCAACTAAGTCAT P: CACCGCAAATGCTTC	66	1.5	106%
RB	F: TCCCATGTTGCTCAAAGAACCA R: CCGTGCACTCCTGTTCTGA P: TCACCTCGAACACCC	94	1.5	100%
CDK2	F: CCTCCCCTGGATGAAGATGGA R: CCGCTTGTTAGGGTCGTAGTG P: CAGCATTTGCGATAACAA	66	1.5	100%
HER-2	F: AGCGCTTTGTGGTCATCCA R: GCAGTGAGCGGTAGAAGGT P: CCCAGCCAGTCCCTTG	75	1.5	110%
PI3K	F: TGCCGAGAGATTTTCCCACAAT R: CCTGAAGCTGAGCAACATCCT P: ATTGACAGCAGTAATTTT	85	1.5	101%
AKT1	F: CCACTGTCATCGAACGCACCT R: CACAGTCTGGATGGCGGTTGT P: ATGTGGAGACTCCTGAAG	77	2	88%
AKT2	F: AAGCAGAGGCTTGGTGGG R: TTGATGCTGAGGAAGAACCTG P: CCAGCGATGCCAAGGAGG	71	2	96%
AKT3	F: TCGAGAGAGCGGGTGTTCT R: TGTAGATAGTCCAAGGCAGAGACAA P: ACGTGTGCGGTCCTC	77	2.5	102%
RAF1	F: TCAGGAATGAGGTGGCTGTTC R: TGTACCCCATGAAAAGCAGAA P: CGCAAAACACGGCATG	66	1.5	104%
CK8	F: GGCTCCAGGCTGAGATTGAG R: GCTCGGCATCTGCAATGG P: CCAGAGGGCTTCCC	72	1.5	86%

Table 4. Primers and MGB Probe sequences, condition of reactions and efficiency of Real Time analysis. (F: Forward, R: Reverse, P: Probe).

3.6 Cell lines

Cell lines cultures were used for to settting up the PCR conditions for candidate genes.

According to the molecular classification, three human BC cell lines were selected: MCF-7, SKBR-3 and MDA-MB-231. Briefly, MCF-7 is an oestrogen dependent epithelial breast cancer cell line, which expresses also progesterone and androgen receptors and it is classified as luminal A subtype (Soule et al., 1973). SK-BR-3 is a breast cancer epithelial cell line that is characterized by the overexpression of the HER2 gene product. MDA MB-231 is an oestrogen receptor independent breast cancer epithelial cell line and it is classified as basal-like (or triple negative) subtype. (http://www.lgcstandards-

atcc.org/ATCCCulturesandProducts/CellBiology/CellLinesandHybridomas/tabid/981/Default.aspx).

Cells were cultured in adhesion and maintained in low glucose Dulbecco's modified Eagle's Medium (DMEM) with L-glutamine (Euroclone S.p.A, Pero Milano, Italy), supplemented with 10% (v/v) fetal bovine serum (FBS; Euroclone), 100U/ml penicillin and 100 μg/ml streptomycin (Euroclone), 3 μg/ml Gentamicin. Cells were grown at 37°C in an incubator supplied with 5% CO_2. Confluent monolayer cells were washed with phosphate-buffered saline (PBS) and treated with trypsin (0.25% (w/v) trypsin, 0.5 mM EDTA in PBS) at 37°C for 3-4 minutes or until cells were dispersed. An appropriate volume of complete growth medium was added to block trypsin. The solution of cells were removed from the flask, collected in a falcon tube and centrifuged at 800 rpm for 5 minutes. Cell pellet was then resuspended in 1 ml new warm medium and an aliquot was collected to count cells in a Burker cell counting chamber. Cells were finally plated in a new flask in a proper concentration. For sub-cultivation, cells were resuspended in complete growth medium, according to the following ratio:

- MCF7 at a ratio of 1:4;
- Sk-BR-3 at ratio of 1:2;
- MDA-MB 231 at ratio of 1:3;

Cells were frozen at a density of approximately 2 x 10^6 cells/ml in FBS with 10% DMSO (Dimethyl sulfoxide) and then stored at -80°C.

3.6.1 RNA Extraction from Cell Lines

MCF-7, MDA-MB-231 and SK-BR-3 RNAs were extracted from 2x10^6 cell pellets with 1 ml of TRIzol ® Reagent (Invitrogen) following the manufacturer's instruction. In brief, for the separation phase, 200 μl of chloroform (Sigma-Aldrich) was added and the tube was vortexed for 15 seconds. After an incubation of 3 minutes on ice, samples were centrifuged at 12000 x g for 15 minutes at 4°C and the three phases separation was obtained. The upper acqueous phase containing the RNA was collected and placed in a new tube. RNA was then precipitated with 500 μl of isopropanol. After that, glycogen was added and samples were centrifuged at 1200 x g for 10 minutes at 4 °C. The surnatant was removed and the pellet

was washed with 1 ml of cold 70% ethanol. Finally, the RNA pellets were resuspended in DEPC treated water, quantified by spectrophotometric reading and stored at -80°C.

3.6.2 DNase Treatment, Reverse Transcription and PCR

DNase treatment was performed as previously described in paragraph 3.5.2.

Reverse transcription was performed with SuperScript II ™ Reverse Transcriptase (Invitrogen) according to manufacture's instructions. Oligo dT priming (Integrated DNA Technologies) was used instead of random hexamers. Briefly, 4.5 μg of total RNA were mixed with 1 μl dNTPs (10mM each), l μl oligo dT (0.5 μg/μl) in a final volume of 12 μl. The mixture was heated at 65 °C for 5 minutes and rapidly chilled on ice. Later, 8 μl of RT mixture were added, yielding a final concentration of 1X First Strand Buffer (50 mM Tris-HCl pH 8.3; 75 mM KCl; 3 mM $MgCl_2$ - Invitrogen), 10 mM DTT (Invitrogen), 4 units of Rnase Inhibitor (Promega) and 200 units of SuperScript ™ II RT (Invitrogen) in a final volume of 20 μl. The mixture was incubated at 42°C for 1 hour, then the reaction was blocked by heat inactivation at 70 °C for 15 minutes.

Thirty ng of cDNA were used in end-point PCR (Thermal Cycler, Biorad),to set up annealing temperature through a thermal gradient program. Each amplification was performed in a mixture containing 1X PCR buffer (10 mM Tris-HCl, pH 8.3; 50 mM KCl), 1.5 mM $MgCl_2$, dNTPs (each 0.2 mM), primers (15 pmol each) and 0.9 units of AmpliTaq® Gold DNA Polymerase (Applied Biosystems), in a final volume of 25 μl. PCR program was: hot-started at 95°C for 7 minutes, subsequently denaturation for 30 seconds at 95°C, a specific gradient temperature for annealing for 1 minutes, extension for 30 seconds at 72°C. A program of 35 cycles was used to amplify the targets. Ten μl of each PCR reaction were loaded onto 2% agarose gel, stained with ethidium bromide, and visualized under UV transilluminator.

3.7 Statistical Analysis

Associations between clinical-pathological factors and categories of markers were tested for significance using the chi-square test (or Fisher's exact test depending on the sample size) for categorical variables. The distribution of data for continuous variables was tested by

kurtosis test to establish the statistical tests (parametric or non-parametric) to use. For continuous variables normally distributed the parametric Student's t-test and one-way ANOVA test were used. For those variables asymmetrically distributed non parametric Kruskal-Wallis test were used. The Spearman's rank correlation coefficient was used to test the strength of correlation for non-parametric variables. For the comparison of immunohistochemical markers between primary tumour and paired metastatic lymph node the sign test for matched pairs was used. Real time qRT-PCR normalized values for the genes were dichotomized for subsequent analysis with respect to their median value of expression. Tumours with gene expression levels lower or higher than the median value were classified as low or high status of expression, respectively. BC specific survival was defined as the time from surgery to breast cancer specific death or end of follow-up, whichever came first. The log-rank test and Kaplan-Meier curves were used to estimate and compare the dependence of BC patients' specific survival on single variables or on combinations of variables, namely the molecular classification, immunohistochemical markers, the expression of genes, the clinical and pathological variables taken one at time. To estimate the joint effects of the analysed covariates on patients' survival and confirm the results of the log-rank test, the data were analysed by fitting the Cox proportional hazard regression model. The power of the study was calculated by exponential test, which compares two exponential survivor functions calculated on the hazard ratio in the groups of interest: Ki67 status in luminal patients LN- and LN+, and CK8 expression in loco-regional metastasis.

All p-values are two-sided with values <0.05 regarded as statistically significant, but for Kruskal-Wallis test the adjusted significance value was considered for statistics ($p<0.05$). P-values between 0.05 and 0.07 were considered "borderline".

Statistical analyses were performed with the Stata/SE 12 package (Stata, College Station, TX).

4 RESULTS

The first part of the study was related to the molecular characterization of primary tumours and a set of metastatic lymph nodes of BC patients with respect to clinical and pathological features. The second part dealt with the analysis of 9 genes of interest as candidate prognostic biomarker, in association with the molecular classification.

4.1 Description of Clinical and Pathological Features

The cohort of BC patients included 305 young women with an age at diagnosis less than or equal to 55 years 151 patients were negative for axillary loco-regional lymph node involvement (LN-), and 154 were positive for axillary loco-regional lymph nodes metastases (LN+) at the time of diagnosis (Table 6). The cohort of patients presented an average age at diagnosis of 46.9 years (range 26-55). No significant differences in age at diagnosis were observed between the LN- (46.8 y, range 32-55 y) and LN+ (47.0 y, range 26-55 y), but the frequency of patients under 35 was higher in LN+. Fifteen patients (5%) were lost at follow-up during the period of observation because of emigration. For those the follow-up end at the time of emigration.

Patients were treated with radical mastectomy, modified radical mastectomy or breast-conserving surgery, with or without axillary lymph node dissection. All patients submitted to conserving surgery were treated with radiotherapy. In detail, in patients LN- or LN+ with 3 or less lymph nodes involved, only the breast was irradiated; in patients with more than 3 metastatic lymph nodes and capsule invasion, breast and axilla were irradiated. All LN+ patients were treated with adjuvant chemotherapy with CMF (cyclophosphamide, methotrexate, fluorouracil) or EC/ECF (epirubicin, cyclophosphamide, fluorouracil) regimens according to standard protocols. ER-positive patients, both LN- and LN+, were submitted to hormone therapy with Tamoxifen. No specific treatment with trastuzumab was performed in patients HER2+.

The mean follow-up time was 14 years (range 0-25), in detail it was 16 years (range 0-25) for LN- group and 11 years (range 0-24) for the LN+ group ($p<0.001$). In the LN- group 49 women (33%) recurred, while in the LN+ group 98 patients (69%) developed

recurrences. For 4 LN- patients (3%) and 11 LN+ ones (7%) information about recurrences was missing.

Significant differences between LN- and LN+ subgroups were detected for frequency of women aged less than 35 years at diagnosis ($p=0.004$), histological subtypes ($p=0.001$), histological grade ($p<0.001$), tumour size ($p<0.001$), type of surgery ($p<0.001$), recurrences ($p<0.001$) and number of patients who died from BC ($p<0.001$). All details are reported in Table 5.

Features	Entire cohort N=305 n (%)	LN- N=151 n (%)	LN+ N=154 n (%)	p
Age, years				
≤35	21 (7)	4 (3)	17 (11)	0.004
>35	284 (93)	147 (97)	137 (89)	
Mean Age, years (range)	46.9 (26-55)	46.8 (32-55)	47.0 (26-55)	0.2
Histology				
Ductal	250 (82)	115 (76)	135 (87)	0.001
Lobular	28 (9)	13 (9)	15 (10)	
Medullary	8 (3)	8 (5)	0	
Mucinous	7 (2)	4 (3)	3 (2)	
Tubular	12 (4)	11 (7)	1(1)	
Grade				
1	39 (13)	34 (22)	5 (3)	<0.001
2	145 (47)	86 (57)	59 (38)	
3	121 (40)	31 (21)	90 (59)	
Tumour size, cm				
≤2	181 (60)	109 (72)	72 (48)	<0.001
2-5	106 (35)	39 (26)	67 (44)	
≥5	15 (5)	3 (2)	12 (8)	
Missing	3	0	3	
Lymph nodes				
1-3 lymph nodes	97 (63)	0	97 (63)	
≥4 lymph nodes	56 (37)	0	56 (37)	
Stage				
I	107 (35)	107 (71)	0	
II	125 (41)	42 (28)	83 (54)	
III	72 (24)	2 (1)	70 (46)	
Missing	1	0	1	
Type of surgery				
Mastectomy	221	88	133	<0.001
Breast-conservation	84	63	21	
Recurrence *				
No	143 (49)	98 (67)	45 (31)	<0.001
Yes	147 (51)	49 (33)	98 (69)	
Missing	15	4	11	
BC specific death	128 (42)	39 (26)	89 (58)	<0.001

Table 5. Clinical and pathological characteristics of BC patients. Entire cohort=305, LN-=151, LN+=154 (* Recurrences diagnosed during follow-up period; p= level of significance for association).

4.2 Molecular Classification

The immunohistochemistry results for ER, PR, HER-2, Ki67, CK8, CK5/6 and vimentin, were dichotomized in negative and positive expression for each marker, according to the previously described criteria (material and methods section).

Subsequently, 305 BC patients were molecularly classified into 4 subtypes in agreement with their protein expression pattern (paragraph 3.4.3).

A total of three cases (1%) failed to be classified in one of the molecular subtypes because not in line with the previously described criteria. In detail, one case exhibited positive staining for PR, CK8, high Ki67, but not ER, HER2 and Ck5/6. The second case showed only high positivity for Ki67, but negativity for the other markers (probably it should be included in triple negative). The third case exhibited positivity only for CK8. All details regarding the staining for the 7 markers are reported in Table 6.

The remaining 303 cases were classified as follows: luminal A was the most frequent molecular subtype with 140 patients (46%), followed by luminal B with 102 cases (34%), 36 basal-like (12%) and 24 HER-2+ (8%). Results were reported in Table 7 and 8.

Most cases in luminal A and B showed positive PR expression (96% and 87% respectively). Among the luminal subtypes a small percentage of cases (11% and 13% respectively for luminal A and B) showed positivity for basal cytokeratins. In the luminal B subtype 37 cases (36%) showed HER-2 amplification, of those 4 samples (11%) had a Ki67 rate less than 14%, while 33 (89%) cases had a higher Ki67 rate.

Most HER-2+ (83%) and basal-like (86%) tumours were characterized by high Ki67 expression (>14%). Moreover, in the HER-2+ subtype a large portion of cases (79%) was positive to CK8 and fewer (21%) were positive to CK5/6. Also in basal-like subtype, more than half cases showed CK8 positivity (64%).

Vimentin was prevalently expressed in basal like cases, 75% of positive tumours, in comparison to other classes ($p<0.001$). Luminal A subtype showed the lower rate of positivity for vimentin (24%), while luminal B and Her-2 had similar proportion of vimentin expression.

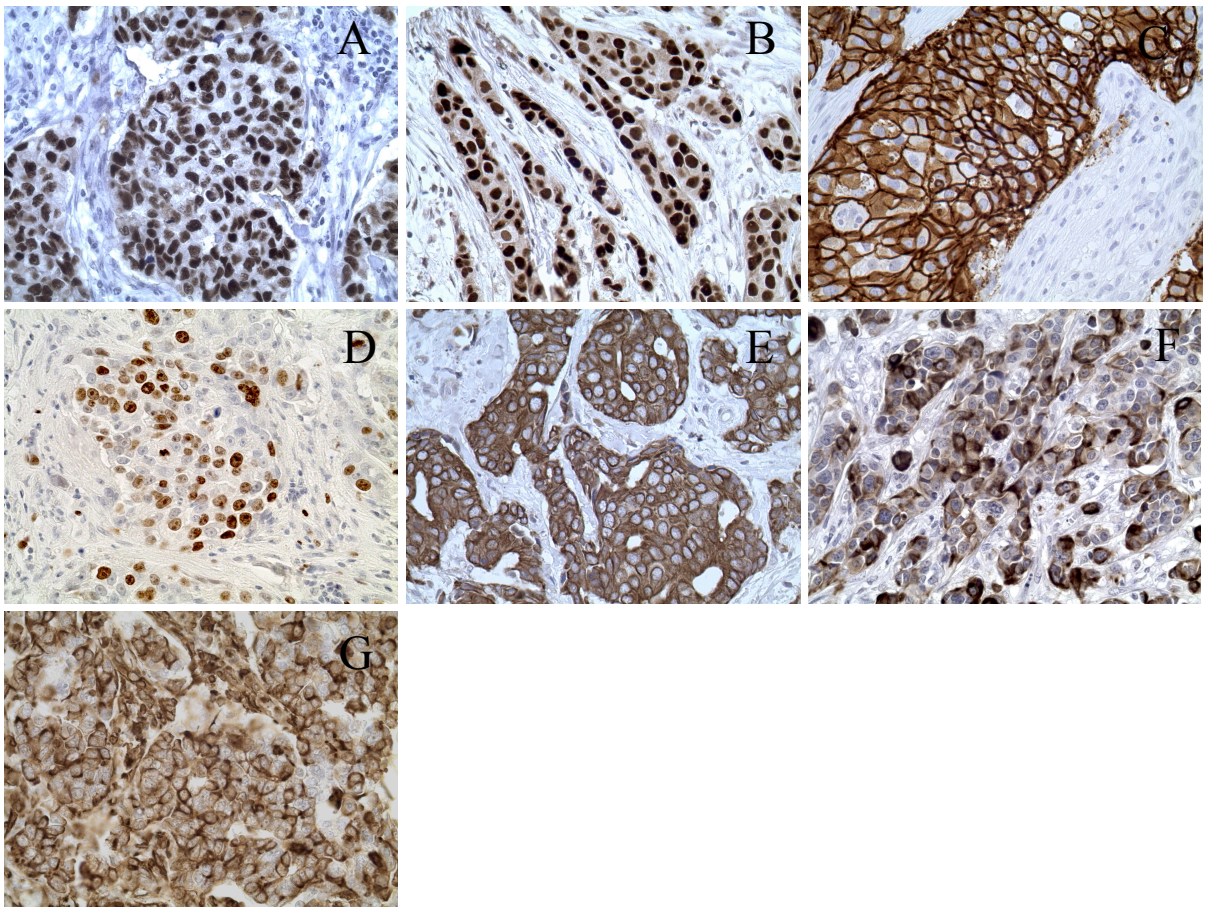


Fig. 6. Representative picture of immunohistochemistry expression for A: ER, B: PR, C: HER2, D: Ki67, E: CK8, F: CK5/6 and G: Vimentin. Pictures are at 20X magnification.

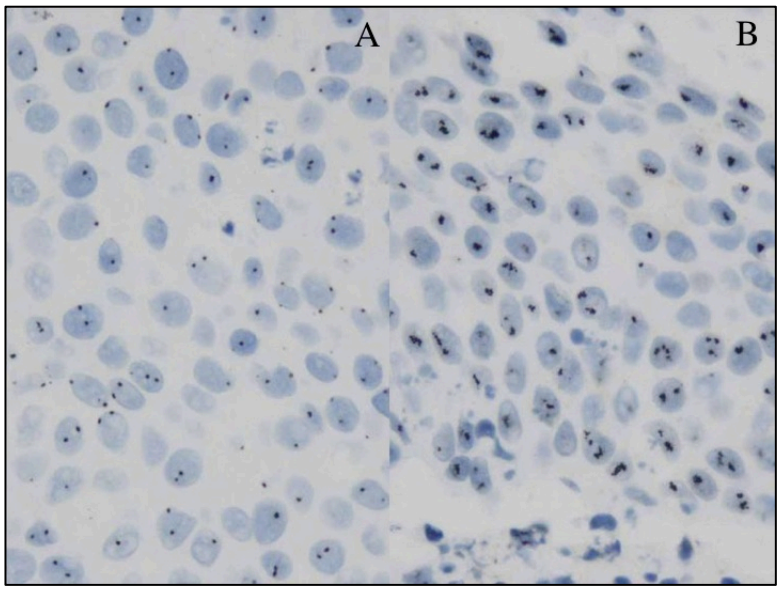


Fig. 7. Representative picture of SISH for chromosome 17 (A) and HER-2 (B).

Marker	Entire cohort N=305 n (%)	LN- N=151 n (%)	LN+ N=154 n (%)	p
ER				
Negative	63 (21)	25 (17)	38 (25)	0.08
Positive	242 (79)	126 (83)	116 (75)	
PR				
Negative	81 (27)	27 (18)	54 (35)	0.001
Positive	224 (73)	124 (82)	100 (65)	
HER-2				
Negative	244 (80)	140 (93)	104 (67)	<0.001
Positive	61 (20)	11 (7)	50 (33)	
Ki67				
<14%	156 (51)	104 (69)	52 (34)	<0.001
≥14%	149 (49)	47 (31)	102 (66)	
CK8				
Negative	29 (10)	11 (7)	18 (12)	0.3
Positive	276 (90)	140 (93)	136 (88)	
CK5/6				
Negative	235 (77)	109 (72)	126 (82)	0.04
Positive	70 (23)	42 (28)	28 (18)	
Vimentin				
Negative	196 (64)	106 (70)	90 (58)	0.03
Positive	109 (36)	45 (30)	64 (42)	

Table 6. Immunohistochemistry results for the markers used in primary tumours characterization. Entire cohort=305, LN-=151, LN+=154 (* Recurrences diagnosed during follow-up period; p= level of significance for association).

4.3 Clinical and Pathological Features of Molecular Classes

The clinical and pathological features stratified per molecular subtypes are shown in Table 7. Significant associations between molecular classification and clinical-pathological features were observed for age at diagnosis with 35 years cut-off (p=0.031), histological type (p<0.001), tumour grade (p<0.001), lymph node involvement (p<0.001), tumour stage (p<0.001) and development of later recurrences (p=0.02).

Patients younger than 35 were related to more aggressive tumours (Soerjomataram et al., 2008) and were more frequently in luminal B and HER2 subtypes.

The ductal (NOS) histological type was the most represented in all the classes. Most lobular carcinomas were detected in luminal, as well as the less aggressive histological types, mucinous and tubular (Ellis et al., 2003). Most medullary carcinomas were grouped in basal-like type, while the remaining belong to luminal A type.

Tumour differentiation was significantly different among the molecular subtypes, grade 1 was present only in luminal, especially in luminal A, while grade 3 mostly in HER2 and basal like ($p<0001$).

Regarding lymph node involvement, luminal A subtype was prevalently associated to lymph node negative patients (70%), while the opposite situation occurred for luminal B, where 73% of cases presented positive lymph nodes. HER-2+ subtype was more frequent in LN+ (79%). The basal-like tumours were equally distributed between LN- and LN+ groups. Patients with luminal B and HER-2+ tumours presented more frequently loco-regional metastases at the diagnosis than patients with luminal A ($p<0.001$ for both). Moreover, HER2+ and basal-like tumours presented a higher number of metastatic lymph nodes (≥4), even if the association between molecular subtypes and number of positive lymph nodes was not statistically significant.

Most luminal A tumours were stage I at the diagnosis, while most luminal B were stage II and almost half of HER2+ stage III. Basal-like tumours were equally distributed in the three stages. Tumour size was not statistically different among molecular subtypes, but it was significantly higher in lymph node positive group ($p<0.001$). Comparing the LN- from LN+ groups of patients, we observed for all molecular subtypes a general, non-significant trend of higher tumour size (≥5 cm) for patients with positive lymph nodes.

In the entire case study, luminal A subtype showed the lowest percentage of recurrences, while more than 75% of HER-2 tumours relapsed. Luminal B and basal-like showed similar intermediate percentage of recurrences. When we analysed LN- and LN+ groups separately we observed similar frequency of recurrences for luminal subtypes: 33% and 26% of recurrences in luminal A and luminal B in LN-; 64% and 67% for luminal A and B in LN+ respectively.

Vimentin expression was prevalently associated to LN+ tumours ($p=0.03$), higher histological grade ($p<0.001$), tumour stage III ($p=0.07$) and presence of later recurrences ($p=0.003$) in all the case study. Moreover, the medullary histological type showed the highest percentage of vimentin expression (88%), with respect to ductal and lobular (both 36%), tubular (25%) and mucinous, which showed no expression ($p<0.001$).

In addition, almost half of patients (44%) who died from BC were vimentin positive, against 29% of positive expression in women who are still alive at the end of follow-up ($p=0.014$). Local or distant recurrences were recorded in 42% of luminal A, 55% of

luminal B, 77% of HER-2 and 57% of basal-like tumours. The pattern of metastatic sites was reported in Table 8. Organ distribution of metastasis among molecular subtypes was almost similar, but for bone and brain. HER-2+ type exhibited a low rate (12%) of bone metastasis in comparison to others subtypes (p=0.03), in detail for bone involvement significant differences were found between HER2+ and luminal B (p=0.01), and basal-like (p=0.04). Higher rate of brain metastases were found in both Her-2+ (17%) and basal-like (11%) (p=0.03). In detail, a significant difference for brain metastases was observed between HER2+ and luminal A (p=0.02), basal-like and luminal A (p=0.04).

Features	Luminal A N=140 n (%)	Luminal B N=102 n (%)	HER-2 N=24 n (%)	Basal-like N=36 n (%)	p
Age (years)					
≤35	4 (3)	10 (10)	4 (17)	2 (6)	0.03
>35	136 (97)	92 (90)	20 (83)	34 (94)	
Age (years) range	47.4 30-55	46.6 30-55	45.9 28-55	47.0 34-55	0.7
Histology					
Ductal	108 (77)	88 (86)	23 (96)	28 (78)	<0.001
Lobular	16 (12)	10 (10)	1 (4)	1 (3)	
Medullary	2 (1)	0	0	6 (16)	
Mucinous	3 (2)	3 (3)	0	1 (3)	
Tubular	11 (8)	1 (1)	0	0	
Grade					
1	34 (25)	5 (4)	0	0	<0.001
2	77 (55)	49 (48)	9 (37)	9 (25)	
3	29 (20)	48 (48)	15 (63)	27 (75)	
Tumour size(cm)					
≤2	95 (68)	50 (49)	12 (50)	22 (63)	0.1
2-5	39 (28)	44 (43)	11 (46)	11 (31)	
≥5	5 (4)	7 (7)	1 (4)	2 (6)	
Missing	1	1	0	1	
Lymp node involvement					
no	98 (70)	28(27)	5 (21)	18 (50)	<0.001
yes	42 (30)	74 (73)	19 (79)	18 (50)	
Lymph nodes					
1-3	28 (67)	51 (69)	10 (53)	7 (39)	0.1
≥4	14 (33)	23 (31)	9 (47)	10 (56)	
Missing	0	0	0	1	
Stage					
I	71 (51)	17 (16)	4 (16)	13 (37)	<0.001
II	48 (34)	55 (54)	9 (38)	12 (34)	
III	21 (15)	30 (30)	11 (46)	10 (29)	
Missing	0	0	0	1	

Recurrence No Yes Missing	 78 (58) 56 (42) 6	 43 (45) 53 (55) 6	 5 (23) 17 (77) 2	 15 (43) 20 (57) 1	0.02
Status BC specific death	 45 (32)	 48 (47)	 14 (58)	 19 (53)	0.001
ER Negative Positive	 0 140 (100)	 0 102 (100)	 24 (100) 0	 36 (100) 0	<0.001
PR Negative Positive	 6 (4) 134 (96)	 13 (13) 89 (87)	 24 (100) 0	 36 (100) 0	<0.001
HER-2 Negative Positive	 140 (100) 0	 65 (64) 37 (36)	 0 24 (100)	 36 (100) 0	<0.001
Ki67 <14% ≥14%	 140 (100) 0	 6 (6) 96 (94)	 4 (17) 20 (83)	 5 (14) 31 (86)	<0.001
CK8 Negative Positive	 0 140 (100)	 0 102 (100)	 5 (21) 19 (79)	 23 (64) 13 (36)	<0.001
CK5/6 Negative Positive	 125 (89) 15 (11)	 89 (87) 13 (13)	 19 (79) 5 (21)	 0 36 (100)	<0.001
Vimentin Negative Positive	 107 (76) 33 (24)	 65 (64) 37 (36)	 14 (58) 10 (42)	 9 (25) 27 (75)	<0.001

Table 7. General clinical-pathological and IHC characteristics of breast tumours according to molecular classification. Luminal A= 140, Luminal B=102, HER-2+= 24, Basal-like= 36. (p= level of significance for association).

Metastatic site	Luminal A n (%)	Luminal B n (%)	HER-2 n (%)	Basal-like n (%)	p
Bone	22 (16)	29 (28)	3 (12)	11 (31)	0.03
Liver	27 (19)	27 (26)	7 (29)	12 (33)	0.7
Lung and Pleura	24 (17)	20 (20)	6 (25)	13 (36)	0.2
Brain	2 (1)	8 (8)	4 (17)	4 (11)	0.03
Endocrine *	4 (3)	4 (4)	1 (4)	1 (3)	0.9
Distant LN **	4 (3)	2 (2)	1 (4)	1 (3)	0.9
Other ***	5 (4)	6 (6)	1 (4)	1 (3)	0.9
Local	22 (16)	13 (13)	7 (29)	5 (14)	0.3

Table 8. Metastasis sites and local recurrence in accordance to BC molecular subtypes over the patients who recurred (N=146). * Endocrine sites include thyroid, ovary and kidney. ** Distant lymph nodes are considered the nodes that do not belong to the axilla, including the supraclavicular lymph nodes. * Other sites comprise peritoneum and pericardium. (p= level of significance for association).**

4.4 Survival Analysis

4.4.1 Role of Clinical and Pathological Features

At the end of the follow-up 128 women died of BC, 16 died of any cause unrelated to BC, 18 were lost at follow-up, because of emigration, and 143 patients were alive. In the LN- group 7 women (5%) died from causes different from BC and 39 patients (26%) died of breast cancer specific-death. Within the LN+ group 9 women (6%) had a non-BC-specific death, while 89 patients (58%) died from BC.

Log-rank test for clinical and pathological variables among the entire cohort of patients revealed that histological type of tumour ($p=0.04$) (Fig. 8A), histological grade ($p<0.001$) (Fig. 8B), tumour size ($p=0.02$) (Fig. 8C), lymph node involvement ($p<0.001$) (Fig. 8D), number of positive lymph nodes ($p=0.01$) (Fig. 8E) and stage of tumour ($p<0.001$) (Fig. 8F) were significant for BC specific survival. Age at diagnosis was not significant ($p=0.1$). Also the presence of later recurrences negatively affected BC specific survival ($p<0.001$). In detail patients with negative nodes, mucinous and tubular histological type, tumour size

less than 2 cm, a number of positive lymph nodes equal or less than 3, with tumour stage I and well differentiated tumour (G1-G2) showed a significant higher BC specific survival.

The Cox proportional hazard model was performed for histological type of tumour, tumour grade, tumour size, lymph node status and age at diagnosis. It emerged that histological type of tumour played a protective role on patient's survival (hazard ratio 0.8, $p=0.04$), tumour grade affected negatively BC specific survival (hazard ratio 1.5, $p=0.006$) and also lymph node involvement constituted a higher risk for BC death (hazard ratio 2.2, $p<0.001$) (probability of the model $p<0.001$).

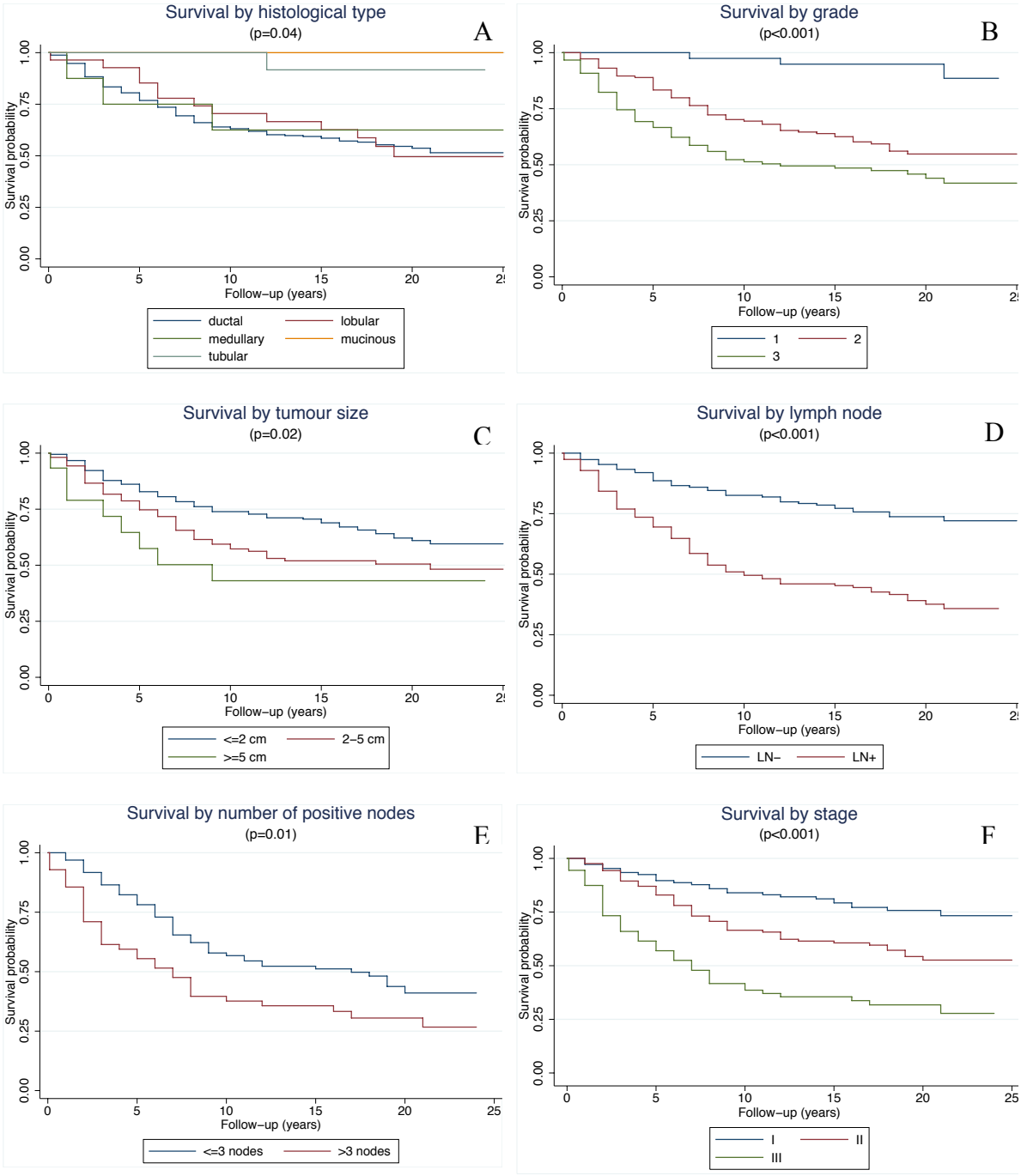


Fig. 8. Kaplan-Meyer survival curves for the entire cohort (N=305, 128 BC specific deaths) over the following clinical and pathological features: histological type of tumour (A), histological tumour grade (B), tumour size (C), lymph node involvement (D), number of positive nodes among LN+ patients (E), tumour stage (F).

4.4.2 Role of Molecular Classification

BC specific survival curves for molecular classes in the entire cohort of patients are shown in Fig. 9A and the corresponding survival estimates are listed in Table 9. A significant better BC specific survival was recorded for luminal A tumours and a worse outcome for HER-2 subtype (p=0.0003). In detail, significant differences were found between luminal A and luminal B (p=0.009), luminal A and HER-2+ (p<0.001), luminal A and basal-like (p=0.01), luminal B and HER-2+ (p=0.04). No significant differences between HER-2+ and basal like (p=0.3), luminal B and basal-like (p=0.5) were detected.

Furthermore, all patients with HER-2+ tumours showed a early BC specific death within 5 years, except for one patient. In addition, women with basal-like subtype died from BC within 10 years, only one woman died at 13 years of follow-up. Women who had luminal tumours continued to die from BC even after 20 years of follow-up, with a constant reduction of survival estimates.

When we separately analysed BC specific survival in the groups of LN- and LN+ patients, a statistically significant difference among molecular classes was no longer observed (p=0.09 and p=0.2 for LN- and LN+), as shown in Fig 9B-C. Luminal A and luminal B subtypes displayed similar survival in LN- and LN+ subgroups respectively, showing that differences between A and B subgroups are likely related not to molecular subtypes, but to differentiation and stage at the diagnosis, which were higher in most luminal B cases.

Luminal A and luminal B classes showed a significant better survival in LN- in comparison to LN+ (p<0.001 and p=0.001, respectively). On the other hand, both HER-2 and basal-like subtypes did not show statistically significant differences for survival over lymph node stratification (p=0.2 for both).

Moreover, luminal tumours expressing PR showed a better outcome in comparison to those lacking its expression (p=0.02) (Fig. 10A), in particular in luminal A tumours (p=0.03) (Fig. 10B). In luminal B neither the expression of PR (p=0.3) or HER-2 (p=0.5) resulted significantly associated to BC specific survival. The expression of Ck5/6 in luminal tumours was not related to the patients' outcome (p=1.0 and p=0.8 for luminal A and luminal B, respectively).

In HER-2+ and basal-like subtypes Ki67 did not affect significantly BC specific patients' survival (p=0.5 and p=0.4, respectively).

A separate mention was dedicated to vimentin, since it was not involved in molecular classes definition. The log-rank test revealed that positive vimentin expression was associated to shorter BC specific survival in the entire cohort of patients (p=0.009) and in LN- group (p=0.02), but not in LN+ patients (p=0.6). Stratifying patients into the four molecular classes, the logrank test showed shorter survival in luminal (p=0.03) and luminal B patients (p=0.04) for positive vimentin expression, as showed in Fig. 10 C-D. This marker was not significant for BC specific survival in the others molecular classes (p=0.5, p=0.9, p=0.6 respectively for luminal A, HER-2+ and basal-like).

Multivariate Cox regression model was run for the pathological covariates lymph node involvement, tumour size, histological type, histological grade, age at diagnosis and the molecular classes. The analysis revealed that lymph node involvement ($p<0.001$) and histological grade (p=0.05) were significant prognostic factor for worse outcome. Instead, histological type (p=0.04), luminal A (p=0.04) and luminal B (p=0.03) remained a significant prognostic factor for good patients' outcome, as reported in table 10.

Time	Luminal A (%) (95% CI)	Luminal B (%) (95% CI)	HER2 (%) (95% CI)	Basal-like (%) (95% CI)
5-year	87 (80-91%)	85 (76-90%)	44 (24-63%)	58 (40-72%)
10-year	75 (67-81%)	66 (55-74%)	44 (24-63%)	49 (32-64%)
15-year	72 (64-79%)	56 (46-66%)	44 (24-63%)	46 (29-62%)
20-year	67 (58-74%)	48 (36-59%)	35 (15-57%)	46 (29-62%)
25-year	65 (56-73%)	46 (35-57%)	-	46 (29-62%)

Part I

Time	Luminal A (%) (95% CI)	Luminal B (%) (95% CI)	HER2 (%) (95% CI)	Basal-like (%) (95% CI)
5-year	92 (84-96%)	96 (75-99%)	60 (13-88%)	72 (46-87%)
10-year	85 (76-90%)	96 (75-99%)	60 (13-88%)	61 (35-79%)
15-year	81 (71-87%)	85 (64-94%)	60 (13-88%)	56 (31-75%)
20-year	78 (68-85%)	75 (52-88%)	60 (13-88%)	56 (31-75%)
25-year	75 (64-83%)	75 (52-88%)	-	56 (31-75%)

Part II

Time	Luminal A (%) (95% CI)	Luminal B (%) (95% CI)	HER2 (%) (95% CI)	Basal-like (%) (95% CI)
5-year	76 (60-86%)	80 (69-88%)	40 (18-61%)	43 (20-64%)
10-year	52 (36-66%)	54 (42-65%)	40 (18-61%)	37 (15-58%)
15-year	52 (36-66%)	45 (33-57%)	40 (18-61%)	37 (15-58%)
20-year	42 (26-58%)	38 (26-50%)	27 (6-54%)	37 (15-58%)
25-year	-	-	-	-

Part III

Table 9. BC specific survival estimates by molecular classes in the entire cohort (*part I*) (N=303), in LN- group (N=149) (*part II*) and LN+ group (N=153) (*part III*). (CI= Confidence Interval).

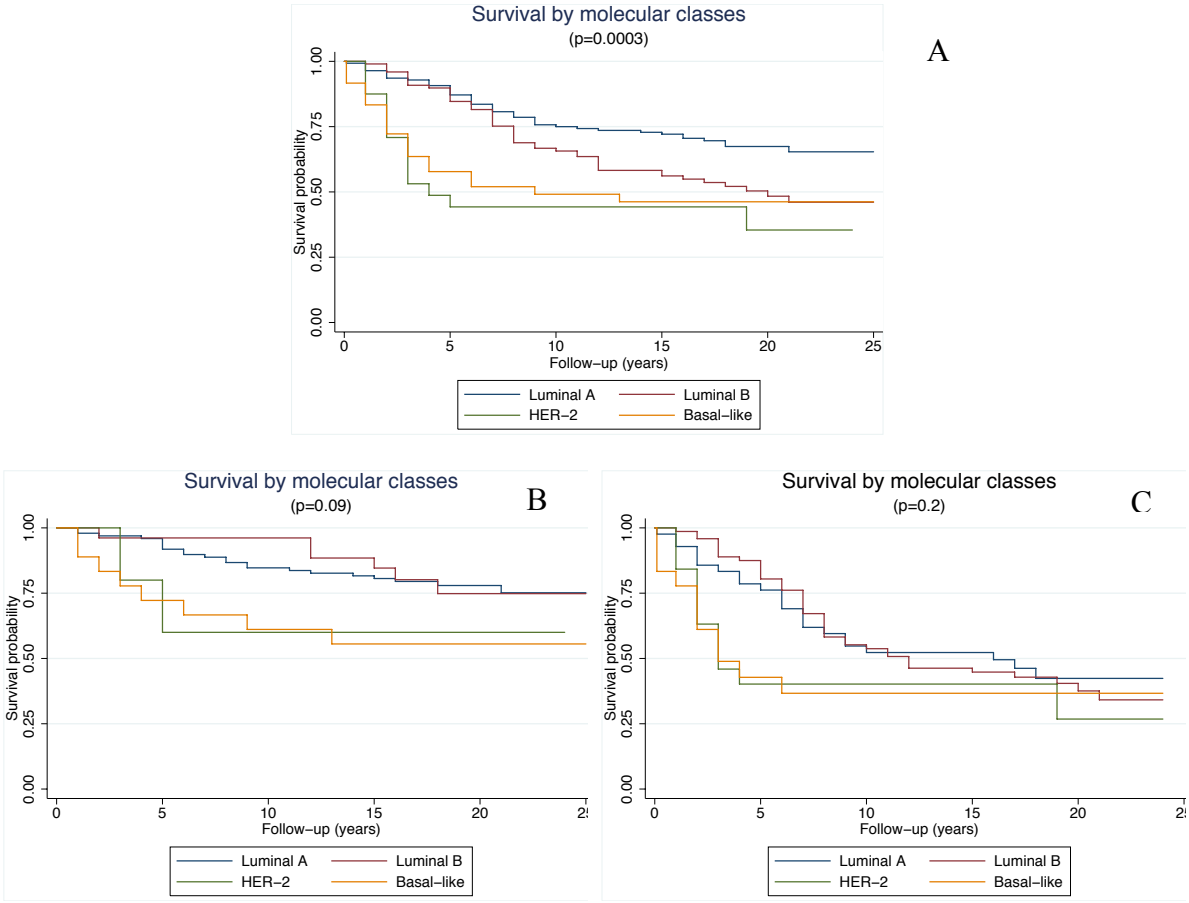


Fig. 9. Kaplan-Meier survival curves with respect to molecular classes in the cohort of BC patients (A) (N=303, 126 BC specific deaths), lymph node negative tumours (B) (N=149, 38 BC specific deaths), classes in lymph node positive tumours (C) (N=153, 88 BC specific deaths).

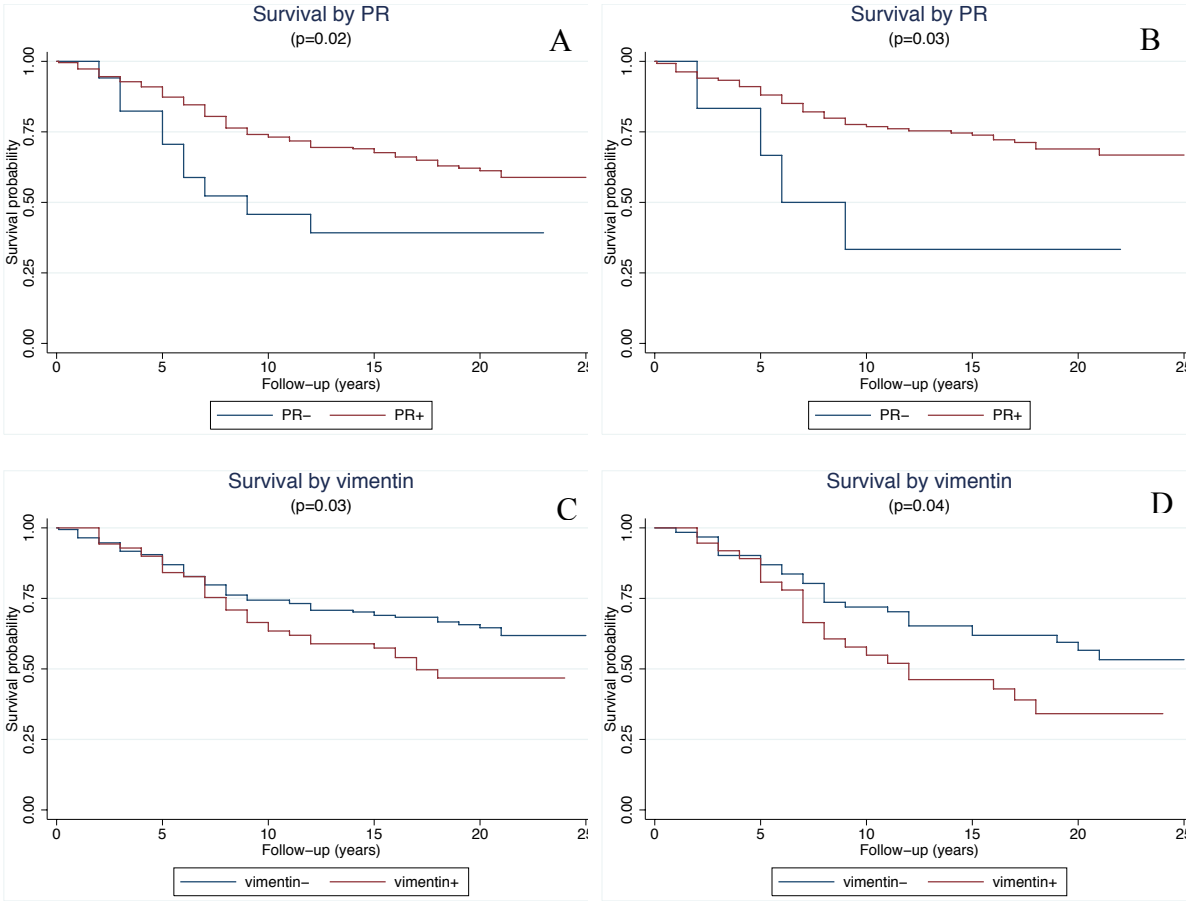


Fig. 10. Kaplan-Meyer survival curves according to PR and vimentin expression in patients classified as 'luminal'. Survival by PR in all luminal tumours (A) (N=242, 93 BC specific deaths) and in luminal A tumours (B) (N=140, 45 BC specific deaths). Survival by vimentin in all luminal tumours (C) and in luminal B tumours (B).

Risk Factors	Hazard Ratio	95% CI*	p
Age at diagnosis	1.0	0.9-1.0	0.4
Histologic type	0.8	0.6-1.0	0.04
Grade	1.4	1.0-1.9	0.05
Tumour size	1.2	0.9-1.7	0.2
Lymph node	2.4	1.5-3.7	0.000
Luminal A	0.2	0.1-0.9	0.04
Luminal B	0.2	0.1-0.8	0.03
HER-2+	0.3	0.1-1.4	0.1
Basal-like	0.4	1.0-1.6	0.2

Table 10. Results of Cox proportional hazard model of BC specific survival (N=303, 128 BC specific deaths) (*Confidence Interval, p= level of significance for association).

4.4.3 Luminal Tumours

The long follow-up period allowed us to observe that the ER-positive patients showed the presence of very long-term recurrences, contrary to ER-negative tumours. For this reason, we focused on luminal A and luminal B tumours and we compared patients who had 'early' and 'late' BC specific death. We considered as 'early' BC specific deaths those occurred within 5 years, 'intermediate' the one between 6 and 10 years of follow-up, 'late' BC specific deaths those that happened after 10 years of follow-up. Among the 93 patients with luminal tumours, who died from BC, 33 (35%) women had 'early' BC specific death, 35 (38%) patients died from BC between 6-10 years of follow-up, and 25 (27%) patients had a "late" one.

The Log-rank test was used to test the influence of the pathological features and markers among the 'early', the 'intermediate' and the 'late' groups of BC specific death, separately. All details are reported in Table 11.

- In patients who had 'early' BC death, histological type of tumour (p=0.3), and age at diagnosis (35 years cut-off) (p=0.9) did not significantly affect BC patients' survival. Otherwise, tumour grade (p<0.001), tumour size (p=0.006), lymph node involvement (p<0.001) number of positive nodes (p=0.02) and tumour stage (p<0.001) resulted significant for BC specific survival. In detail, tumour with G3 histologic grade, size larger than 2 cm, more than 3 positive lymph nodes involved at diagnosis and stage III are more likely to cause a BC specific death within 5 years. BC specific deaths were not significantly different between luminal A and B (p=0.6). PR negative patients showed a worse survival within 5 year (p=0.06) with respect to positive ones. No significant influence on BC survival was detected for HER-2 (p=0.6), Ck5/6 (p=0.4) and vimentin (p=0.3).
- In the 'intermediate group', clinical parameters with significant influence on survival were: histologic grade (p=0.01) lymph node involvement (p<0.0001), tumour size (p=0.03) and tumour stage (p<0.001). In detail, tumour of grade 3, larger than 2 cm, LN+, showed a worse BC survival in this period of time. Conversely, histological type of tumour (p=0.9), number of positive lymph nodes (p=0.8), age at diagnosis (p=0.9), type of luminal (p=0.1) did not affect BC specific survival. Moreover, PR negative patients had shorter BC survival (p=0.06). HER2, CK5/6 and vimentin expression did not affect patient's survival (p=0.4, p=0.8, p=0.08 respectively).
- Histological type (p=0.4), histologic grade (p=0.1), number of positive lymph nodes (p=0.8) did not significantly affect 'late' BC specific death. Tumour size did not result significant, although small tumours (size ≤2 cm) tended to have late BC specific death with 19 dead patients out of 25 (76%). Tumour stage also showed a similar trend, with 9 out of 25 patients (36%) of stage I who died from BC after 10 years (p=0.3). Tumour size and histologic grade showed an inversion in comparison to the results obtained within 5 years of follow-up. Only lymph node involvement was statistically significant (p=0.04). Patients with positive lymph nodes (73%) are more likely to have early BC specific death with respect to 'late' one (52%). Finally, patients with luminal B tumours tend to have more frequently 'late' BC specific death than luminal A (p=0.01). The role of HER-2 and vimentin expression showed a border influence on BC survival, tumours overexpressing HER-2 and positive for vimentin had a shorter survival (p=0.08 both). CK5/6 expression did not affect 'late' BC death (p=0.8).

Features	Early BC death N=33 n (%)	Intermediate BC death N=35 n (%)	Late BC death N=25 n (%)
Age (years)			
≤35	2 (6)	2 (6)	1 (4)
>35	31 (94)	33 (94)	24 (96)
Histology			
Ductal	31 (94)	31 (89)	19 (76)
Lobular	2 (6)	4 (11)	5 (20)
Medullary	0	0	1 (4)
Grade			
1	0	1 (3)	2 (8)
2	14 (42)	20 (57)	17 (68)
3	19 (58)	14 (40)	6 (24)
Tumour size (cm)			
≤2	13 (39)	15 (45)	19 (76)
2-5	16 (48)	16 (48)	6 (24)
≥5	4 (12)	2 (6)	0
Missing	0	2	0
Lymph node involvement			
no	9 (27)	7 (20)	12 (48)
yes	24 (73)	28 (80)	13 (52)
Lymph nodes			
1-3	12 (50)	20 (71)	9 (69)
≥4	12 (50)	8 (29)	4 (31)
Stage			
I	4 (12)	5 (14)	9 (36)
II	13 (39)	18 (51)	11 (44)
III	16 (48)	12 (34)	5 (20)
PR			
Negative	5 (15)	4 (11)	1 (4)
Positive	28 (85)	31 (89)	24 (96)
HER-2			
Negative	29(88)	29 (83)	19 (76)
Positive	4 (12)	6 (17)	6 (24)
CK5/6			
Negative	31 (94)	30 (86)	22 (88)
Positive	2 (6)	5 (14)	3 (12)
Vimentin			
Negative	22 (67)	21 (60)	16 (64)
Positive	11 (33)	14 (40)	9 (36)
Luminal A	18 (55)	17 (49)	10 (40)
Luminal B	15 (45)	18 (51)	15 (60)

Table 11. General clinical-pathological and IHC characteristics of luminal A and luminal B patients who died from BC subdivided in 'early' (BC death within 5 years) 'intermediate' (BC death between 6-10 years) and 'late' (BC after 10 years).

4.4.4 Role of Ki67 among Luminal Tumours

As the subdivision between luminal A and B is based on the count of Ki67 positive nuclei (14% cut-off), in case of no HER-2 overexpression, we deepened the distribution of Ki67 with respect to BC specific survival, especially in luminal tumours. Considering the rate of positive Ki67 nuclei as a continuous variable, we investigated the relationship between Ki67 percentage and BC specific survival. Ki67 resulted significantly higher in all LN+ patients ($p<0.001$) and in LN+ luminal patients ($p<0.001$). Nevertheless Ki67 was not able to predict the outcome in luminal tumours. Analysing Ki67 percentage separately in LN- and LN+ patients, there were not significant differences in Ki67 positivity between living and BC dead patients in luminal cancers ($p=0.08$ in LN- and $p= 0.4$ in LN+). It seems that survival in luminal patients was not affected by Ki67 percentage of positive nuclei, neither in LN- negative ($p= 0.5$), nor in LN+ patients ($p= 0.9$). This data were also confirmed in the solely luminal B group of patients. The distribution of cases by Ki67 in comparison to BC specific death showed a continuous increment without apparent cut-off levels (Tab. 12).

However, using the Ki67 14% cut-off the Log-rank test revealed as expected a significant longer survival for luminal patients with Ki67<14% ($p=0.008$) (Fig. 11).

Ki67 percentage	Cases n (%)	Specific BC deaths n (%)
<5%	58 (24)	12 (21)
6-10%	43 (19)	16 (37)
11-13%	43 (19)	18 (42)
14-20%	33 (14)	14 (42)
>20%	65 (27)	33 (51)
Tot	242	93 (38)

Table 12. Ki67 percentage of positive nuclei in comparison with BC specific death in luminal patients classified as luminal

Stratifying Ki67 percentage into three categories: ≤10%, 11-20% and > 20%, we found a significant p value (p=0.003) for BC specific survival respect to the 14% cut-off (p=0.008) as well. The corresponding Kaplan-Meyer curves are reported in Fig. 11A and B. Fig. 11A shows that at 4 years of follow-up the curve for 0-10% Ki67 is well separated from the one for 11-20%, with a corresponding better survival. Conversely, in Fig. 11B the separation for good and worse survival according to the 14% Ki67 cut-off is shifted over 7 years of follow-up.

To further investigate the role of Ki67 and other markers in luminal group tumours, we performed the Cox multivariate analysis in luminal tumours, which considered the clinical-pathological variables and the molecular markers PR, Ki67, HER-2 and vimentin. Neither the percentage rate of Ki67, nor the positivity to PR and HER-2 affected the survival of luminal patients, but the histological type (p=0.004), the tumour grade (p=0.01), the tumour size (p=0.07) and the involvement of the lymph node at the diagnosis (p=0.001) were significant. In Table 13 are reported the results of Cox analysis in luminal tumours.

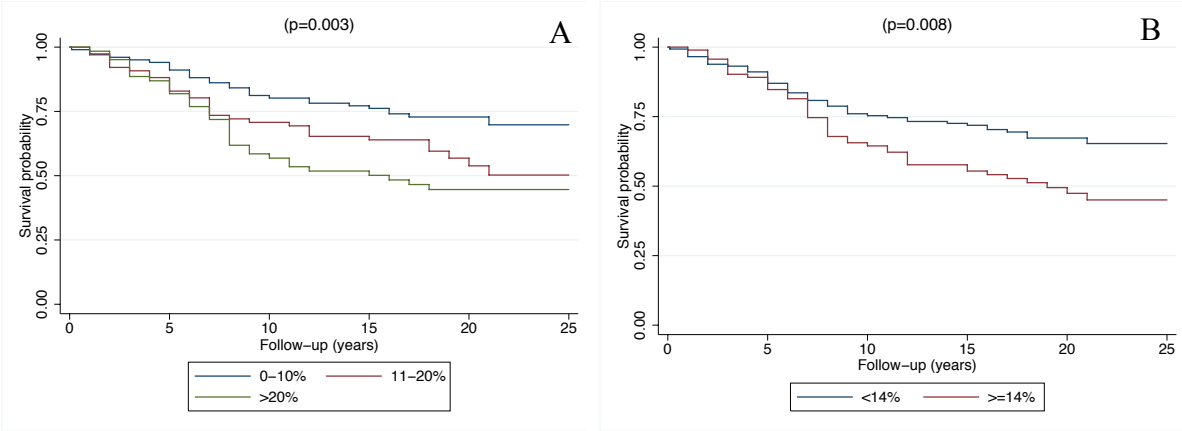


Fig. 11. Kaplan-Meyer survival curves for Ki67 percentage of positive cells in luminal tumours (N=242, 93 BC specific deaths). Ki67 rate is categorized into three ranges (ki67 ≤10%; 11-20%; >20%) (A), Ki67 is dichotomized according to the 14% cut-off (B).

Risk Factors	Hazard Ratio	95% CI	Probability
Age at diagnosis	1.0	1.0-1.1	0.2
Histological Type	0.7	1.5-4.2	0.04
Grade	1.5	1.1-2.2	0.01
Tumour size	1.4	1.0-2.0	0.07
Lymph node	2.5	1.5-4.2	0.001
PR	0.9	0.4-1.9	0.7
Ki67	1.0	1.0-1.1	0.9
HER-2	0.7	0.4-1.1	0.1
Vimentin	1.2	0.9-1.0	0.9

Table 13. Results of Cox proportional hazard model of BC specific survival on patients with luminal tumours (N=242, 93 BC specific deaths) (*Confidence Interval, p= level of significance for association).

4.5 Molecular Classification in Loco-Regional Metastasis

In order to investigate the behaviour of molecular classes with respect to the cells which colonise the metastatic lymph nodes, we compared the staining of the markers assessed in primary tumours to the corresponding metastatic lymph nodes in a set of 100 LN+ patients.

Immunohistochemical assays were performed for ER, PR, HER-2, Ki67, CK8, CK5/6 and vimentin using the same conditions and procedures of the primary tumours. The evaluation of the staining of each marker was already described in paragraph 3.4.3.

Five out of 7 markers resulted significantly different between primary tumour and the corresponding loco regional metastasis (Table 14). Since some molecular classifiers were lost and others acquired at the level of metastatic lymph nodes, the molecular classification resulted quite difficult in the latter.

The expression of each marker was compared in primary tumour and matched metastatic lymph node. In detail, ER and PR showed similar rate of positive cases in primary tumour and in the corresponding lymph node (p=0.1 and p=0.3 respectively). Nevertheless, the

ttest for matched pairs revealed significant differences in the mean percentage of positive nuclei for ER (54% and 65% in primary tumour and metastasis, p=0.0003) and PR (40% and 53% in primary tumour and metastasis, respectively, p=0.0001). Conversely, significant differences were observed for HER-2 (p=0.03), Ki67 (p<0.001), CK8 (p<0.001), CK5/6 (p<0.001) and vimentin (p=0.04). HER-2 resulted more expressed in primary tumour (34%) in comparison to metastatic nodes (23%). Similarly, higher rate of Ki67 (Ki67>14%) was observed in primary tumours (66%) compared to metastatic nodes (40%). Moreover, the mean percentage of Ki67 positive nuclei in primary tumour was 25% and in metastatic lymph node it was 17% (p=0.003). CK8 also decreased in the metastatic lymph node (58%) with respect to primary tumour (89%). Vimentin expression was higher in primary tumour (55%) versus corresponding metastatic nodes (45%), CK5/6 showed a significant increment of positive cells in metastasis (47%) with respect to primary tumour (19%).

The following results were recorded:

Luminal A

Among 28 cases of primary tumour classified as luminal A, 18 (64%) resulted luminal A also in the lymph node (ER+ and /or PR+, Ki67<14%, HER-2-); 8 (29%) cases passed to luminal B (ER+ and or PR+, Ki67≥14% or ER+ and/or PR+, HER-2+), 1 (4%) acquired basal-like phenotype at the lymph node level (ER-, PR-, HER-2-, Ck5/6+) and 1 case resulted unclassifiable because it was negative for all the marker except PR.

Among 28 cases of primary luminal A tumours, in the corresponding metastatic nodes 15 (54%) cases lost CK8 expression and 11 (39%) cases gained Ck5/6 expression. Nine cases (32%) of lymph nodes showed Ki67>14% with respect to the corresponding primary tumour. Eleven (39%) primary tumours were vimentin positive, but only 7 metastatic lymph nodes (25%).

Luminal B

45 primary tumours were luminal B. The immunostaining results in the corresponding metastatic nodes revealed that 24 (53%) lymph nodes remained luminal B (ER+ and or PR+, Ki67≥14% or ER+ and/or PR+, HER-2+), while 21 (47%) were classified as luminal A (ER+ and /or PR+, Ki67<14%, HER-2-).

Among the 45 luminal B primary tumours, 30 (67%) cases lost Ck8 expression in the corresponding regional metastatic node, while 15 (33%) cases acquired the CK5/6

positivity. In addition, 11 (44%) metastatic lymph nodes lost HER-2 overexpression. 43 cases (96%) showed Ki67>14% in primary tumours, while only 17 cases (38%) in metastatic nodes. Vimentin was expressed in 19 (45%) primary tumour and 10 (22%) corresponding regional metastasis.

HER-2+

Among 14 cases of HER-2+ primary tumour, 8 (57%) cases remained HER-2, 2 (14%) cases switched to basal-like subtype in the lymph node (ER-, PR-, HER-2-, Ck5/6+) and 4 (29%) cases became luminal B (ER+ and/or PR+, HER-2+).
Two cases (14%) lost the positivity for HER-2 expression. Overall 4 (29%) cases acquired the ER expression and 3 (21%) the PR positivity. HER-2+ primary tumours were all positive for CK8, except 1 case, while in the relative metastatic lymph nodes 6 (43%) lost the CK8 positivity. A similar result was observed for CK5/6 staining, with only one positive case in primary tumours and 6 (29%) cases in the corresponding metastasis. 3 cases (21%) showed Ki67<14%, while 9 cases (64%) in metastatic lymph nodes. Vimentin expression was comparable with 6 cases in primary tumour and 5 in corresponding metastasis.

Basal-like

Of 13 cases of primary basal-like tumours 6 (46%) cases remained basal at the lymph node level, while 2 (15%) cases lost Ck5/6 expression and for this reason they can be classified as triple negative (ER-, PR-, HER-2-). Considering together basal-like and TN, 61% of regional metastatic nodes remained of the same molecular class as primary tumour. One (8%) case changed its phenotype from basal-like to HER-2+ subtype. Other 2 (15%) cases resulted not classifiable (PR+, ER-, HER-2+, CK5/6+, CK8- and PR+, ER-, HER-2-, CK5/6+ and CK8-). Finally 2 (15%) cases switched to luminal B subtype (ER+ and or PR+, Ki67≥14%).

Overall 3 (23%) cases have lost the positivity for the basal cytokeratin 5/6, while no differences were observed for Ck8 expression. Two (13%) cases acquired the positivity for HER-2+ expression and 4 (31%) the positivity for PR. Ki67<14% was observed in, 1 case (8%) of primary tumour versus 4 (31%) cases of metastatic nodes. Vimentin was expressed in 8 cases both in primary tumour and corresponding metastasis and in 1 patient the expression was lost in the lymph node.

Most specimens preserved their molecular phenotype in the corresponding metastatic regional node: 64% for luminal A, 53% for luminal B, 57% for HER-2+ and 61% for basal-like tumours. Therefore, relevant changes were detected for single markers, generally the acquisition of basal characteristics (CK5/6+) and the loss of the luminal (CK8-) features in the metastatic lymph nodes were observed.

Marker	Primary Tumour N=100	Metastatic lymph node N=100	p
ER			
negative	27%	23%	0.1
positive	73%	77%	
PR			
negative	39%	31%	0.3
positive	61%	69%	
HER-2			
negative	66%	77%	0.03
positive	34%	23%	
Ki67			
<14%	34%	60%	<0.001
≥14%	66%	40%	
CK8			
negative	11%	42%	<0.001
positive	89%	58%	
CK5/6			
negative	81%	53%	<0.001
positive	19%	47%	
Vimentin			
negative	55%	70%	0.04
positive	45%	30%	

Table 14. Comparison of immunohistochemistry results for the seven markers between primary tumour and corresponding metastatic lymph node (N=100). (p= level of significance for association).

4.5.1 Relationship of Markers with Clinical and Pathological features

The relationship between markers expression and clinical pathological features was evaluated for both primary tumours and corresponding metastatic nodes. The clinical-pathological parameters analysed were: histologic grade, size of tumour, number of lymph nodes involved, tumour stage, histologic type of tumour, presence of later recurrences. Only the significant associations are here reported.

Primary tumours lacking of ER and PR expression were significantly associated to higher histological tumour grade, with respectively 81% and 69% of G3 for ER- and PR- (p=0.002 for ER, p=0.06 for PR). The same result was also observed for ER and PR assessed in the metastatic lymph node (p=0.002). HER-2 (p=0.04) and Ki67 (p=0.02) positivity in primary tumours were mainly associated to less differentiated lesions (71% and 65% for HER-2+ and Ki67+). The same was observed for HER-2 in the metastatic lesions (p=0.001), with 87% of G3 tumours expressing HER-2 versus 47% of the same histological grade but HER-2 negative. For tumour size we did not find any significant relationship.

Moreover, ER expression in primary tumours was significantly associated to the number of positive lymph nodes involved (p=0.06). Indeed, 66% of tumours lacking of ER expression showed more than 3 positive axillary nodes at diagnosis, with respect to 45% of positive ER tumour with more than 3 positive nodes. A similar result was also observed for PR: the lack of PR expression was associated to the higher number of metastatic lymph nodes at diagnosis, even if the data was not statistically significant (p=0.09). Interestingly, HER-2 positive expression assessed in the loco-regional metastasis was found marginally associated to the higher number of positive nodes (p=0.06), but nor the expression evaluated in the primary tumour.

For what concerns tumour stage, we observed that PR negative expression was associated to higher tumour stage: 69% of stage III had PR-, while 49% of same stage had PR+ (p=0.05). The relationship between histologic type of tumours and markers expression revealed that vimentin was significantly higher expressed in ductal type with 82% of cases, the remnant was found in lobular type, while the tumours classified as mucinous were vimentin negative (p=0.02). Furthermore, all loco-regional metastasis coming from lobular tumours were negative for HER-2 (p=0.009). Finally we detected a significant association between CK8 expression assessed in the metastatic nodes and presence of later recurrences (p=0.03).

4.6 Survival Analysis for Markers in Metastatic Lymph Node

Log-rank test revealed that CK8 was the only marker significantly associated to BC specific survival: a high expression of CK8 in metastatic lymph nodes was associated to a longer survival ($p<0.001$) in the group of examined LN+ patients, as reported in Fig. 12A.

Among 100 patients with positive regional lymph nodes, 60 women died from BC, of those 27 (45%) showed a positive CK8 expression, while 33 (55%) were negative. Contrarily, CK8 positivity in the primary tumours did not affect significantly BC specific survival ($p=0.9$) (Fig. 12B).

The Cox regression analysis confirmed the CK8 result obtained in univariate survival analysis. Namely, the multivariate survival analysis was run for the following variables: histological type, tumour grade, tumour size, number of positive lymph nodes and CK8 expression in metastatic nodes. The number of positive lymph nodes showed a significant worse effect on survival ($p=0.04$), while CK8 expression had protective effect ($p<0.001$), as shown in table 15. We obtained a power of the study equal to 0.9985, using a sample number of 100 patients, a significance of 0.05 and two-sided test.

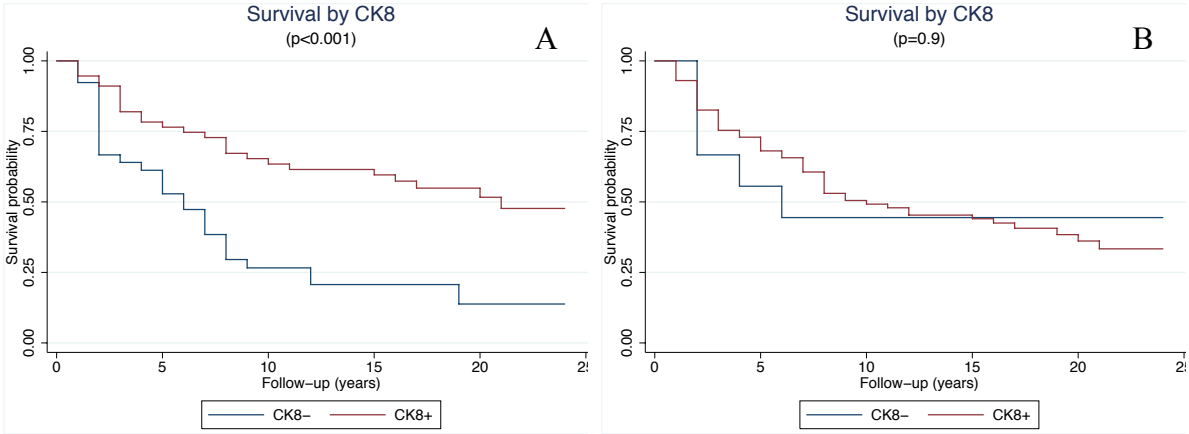


Fig. 12. Kaplan-Meyer survival curves with respect to CK8 expression evaluated by means of IHC in metastatic lymph nodes (A) and in corresponding primary tumour (B) (N=100, 60 BC specific deaths).

Risk Factors	Hazard Ratio	95% CI*	p
Age at diagnosis	1.0	1.0-1.1	0.3
Histologic type	0.7	0.5-1.1	0.2
Grade	1.0	9.6-1.7	0.9
Tumour size	0.9	0.5-1.4	0.5
Number of positive nodes	1.9	1.0-3.3	0.04
CK8 in metastatic nodes	0.4	0.2-0.6	0.0000

Table 15. Results of Cox proportional hazard model of BC specific survival in the group of LN+ patients (N=100, 60 BC specific deaths) (*Confidence Interval, p= level of significance for association).

4.7 Gene Expression and Relationship with Clinical and Pathological Features

The expression of the nine genes, was investigated in relation to the following BC clinical-pathological factors: histological type (ductal, lobular, medullary, mucinous and tubular), tumour grade (1, 2, 3), tumour size (smaller than 2 cm, between 2 and 5 cm, larger than 5 cm), lymph node involvement (no or yes), number of positive nodes (less or more than 3 lymph nodes) tumour stage (I, II, III), presence of later recurrence (no or yes), age at diagnosis (younger or older than 35 years) and patient status at the end of follow-up (alive or dead from BC).

4.7.1 RB and CDK2

RB was slightly higher in lobular and mucinous histological types, lower in medullary type ($p=0.08$) (Fig. 13A). No statistically significant differences were found for RB expression among histological grades, tumour size, stage ($p=0.1$ for all of them) and age at diagnosis ($p=0.7$). RB was significantly higher in LN- tumours ($p=0.01$) and in patients who did not develop later recurrences ($p=0.01$) (Fig. 13B-C).

CDK2 expression was higher in ductal and lower in tubular carcinomas ($p=0.02$) (Fig. 14A). CDK2 expression level was higher in: less differentiated tumour grade ($p=0.04$) (Fig. 14B), tumour size larger than 5 cm ($p=0.05$) (Fig 14C), LN+ patients ($p<0.001$) (Fig. 14D), stage III tumours ($p<0.001$) (Fig. 14E), and patients who underwent later

recurrences (p=0.02) (Fig. 14F). No significant differences were found for CDK2 expression and the number of involved nodes (cut-off 3 lymph node) (p=0.7), as well age at diagnosis, dichotomized at 35 years (p=0.6).

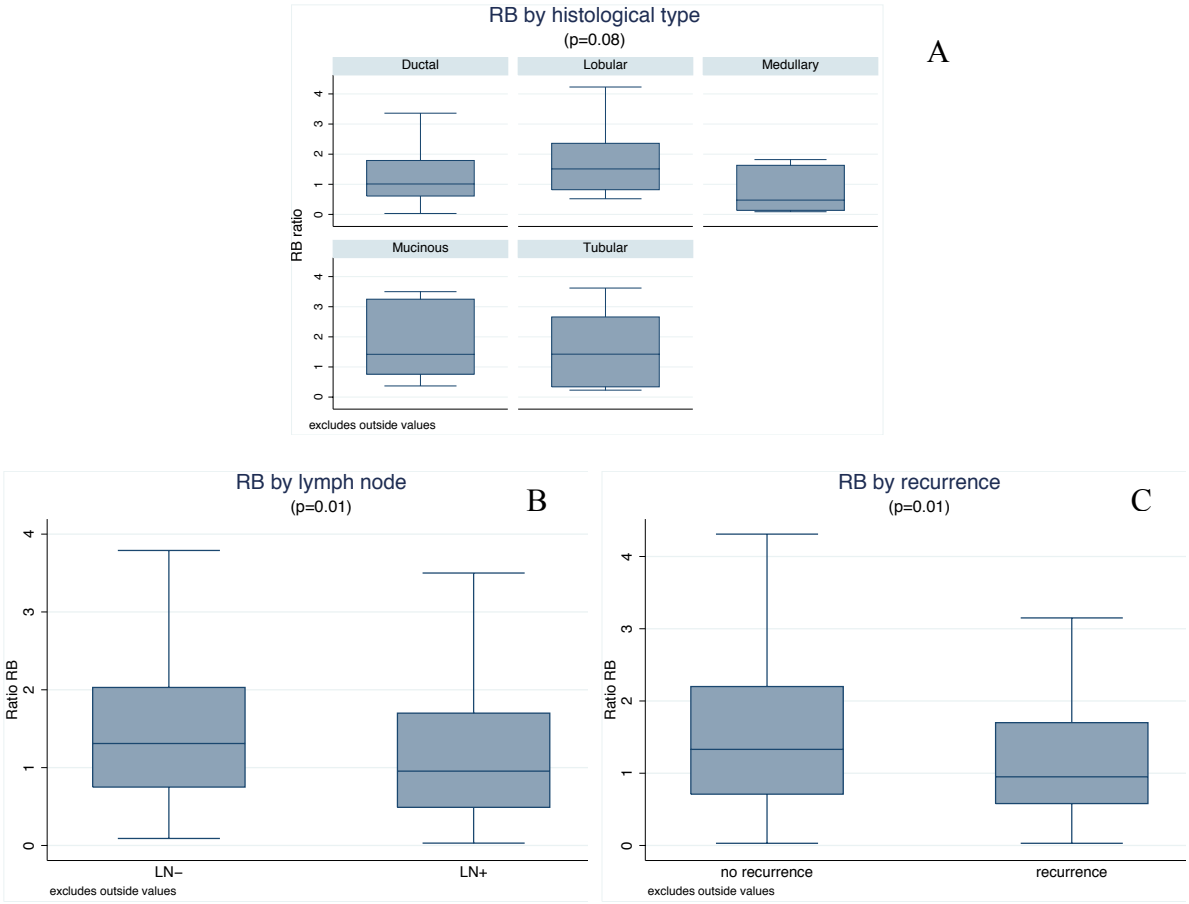


Fig. 13. Box plot representing significant RB expression ratio stratifying patients according to histological type of tumour (A), lymph node involvement (B), presence of later recurrence (C). The median value is indicated as a black line in each group.

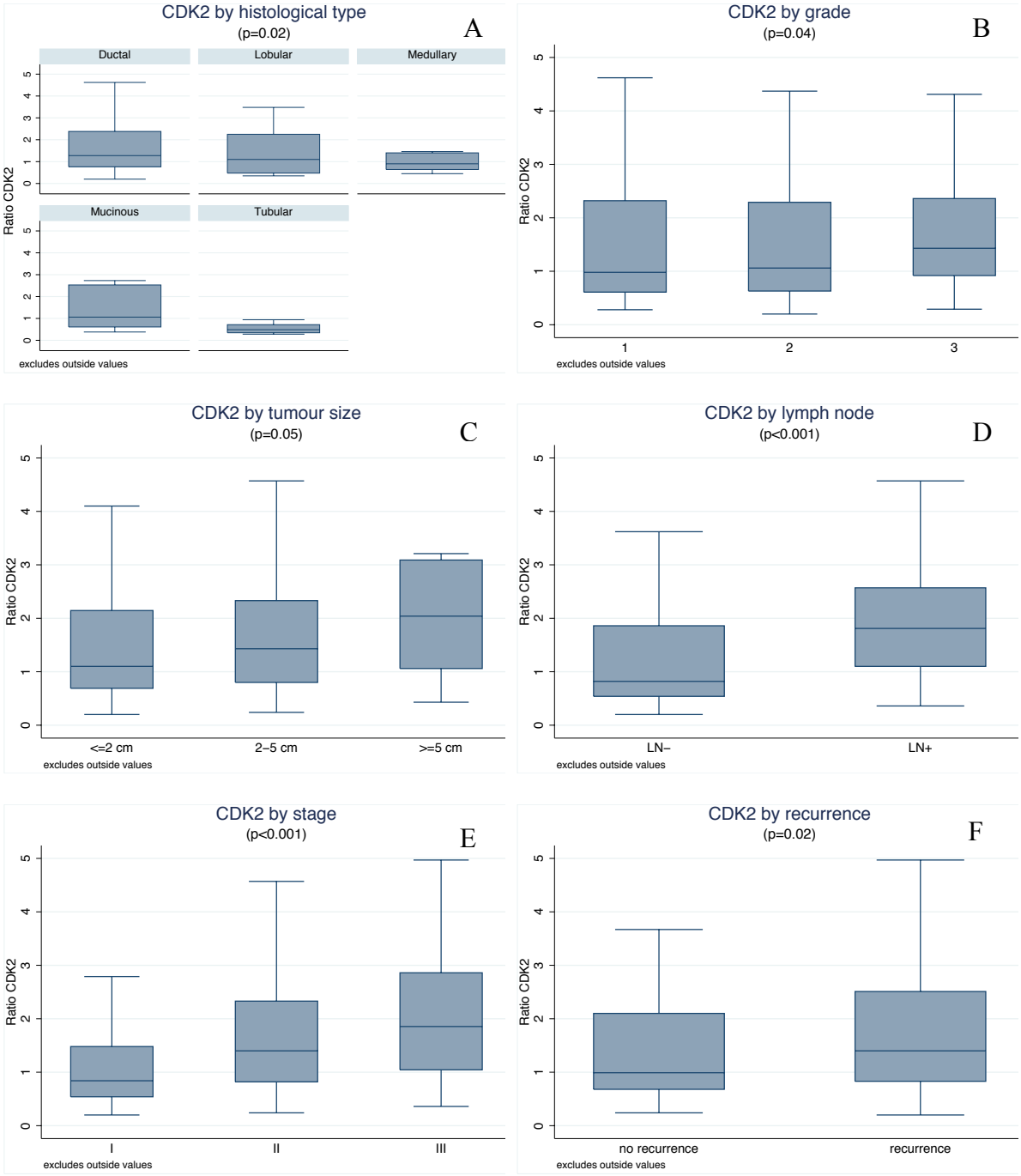


Fig. 14. Box plot representing significant CDK2 expression ratio stratifying patients according to histological type of tumour (A), histological tumour grade (B), tumour size (C), lymph node involvement (D), tumour stage (E), presence of later recurrence (F). The median value is indicated as a black line in each group.

4.7.2 HER-2, PI3K, AKTs and RAF1

HER-2 was more expressed in: ductal carcinomas (p=0.003) (Fig. 15A), tumour with lymph node involvement (p<0.001) (Fig. 15B), stage III BC (p<0.001) (Fig. 15C). Conversely, no significant differences were observed for histological grade (p=0.4), size of tumour (p=0.9), number of positive nodes (p=0.1), presence of later recurrences (p=0.3) and age at diagnosis (35 years cut-off) (p=0.6).

PI3K was similarly expressed among the histological classes of tumour (p=0.2). Conversely, PI3K higher expression was detected in tumour characterized by: lower histological grade (p<0.001) (Fig. 16A), smaller tumour size (p=0.004) (Fig.16B), no lymph node involvement (p<0.001) (Fig. 16C), stage I (p<0.001) (Fig. 16D), absence of later recurrences (p<0.001) (Fig. 16 E), patients older than 35 years at diagnosis (p<0.001) (Fig. 16F).

AKT1 was highly expressed in less differentiated tumour (p=0.05) (Fig. 17A), lymph node positive (p<0.001) (Fig. 17B), stage II and III (p<0.001) (Fig 17C), in patients who developed later recurrences (p=0.02) (Fig. 17D). No significant differences for AKT1 expression among histological types (p=0.2), tumour size (p=0.7), number of positive nodes (p=0.7) and age at diagnosis (p=0.3) were detected.

AKT2 resulted higher in tumours with grade I (p=0.002) (Fig. 18A), LN- (p<0.001) (Fig. 18B), stage I (p=0.003) (Fig. 18C), without recurrences (p=0.01) (Fig. 18D), in patients with more than 35 years old at diagnosis (p=0.02) (Fig. 18E). Otherwise, no different expression was recorded for histological type (p=0.9) and tumour size (p=0.6).

AKT3 expression showed significant differences among histological types, with medullary and mucinous tumours presented the highest expression level (p=0.002) (Fig. 19A). Moreover, higher AKT3 expression was detected in case of grade 1 (p<0.001) (Fig. 19B), in smaller tumours (p=0.003) (Fig. 19C), in LN- cases (p<0.001) (Fig 19D), in cancer of stage I (p<0.001) (Fig 19E), in patients who did not develop later recurrences (p<0.001) (Fig 19F). AKT3 expression did not vary with age at diagnosis (p=0.4).

RAF1 expression was slightly increased only in LN+ tumours (p=0.06) (Fig. 20). No significant differences in its expression were detected for the other variables: histological type (p=0.3), grade, tumour size, number of positive nodes, stage (p=0.7 for all of them), presence of recurrence (p=0.4), age at diagnosis (p=0.8).

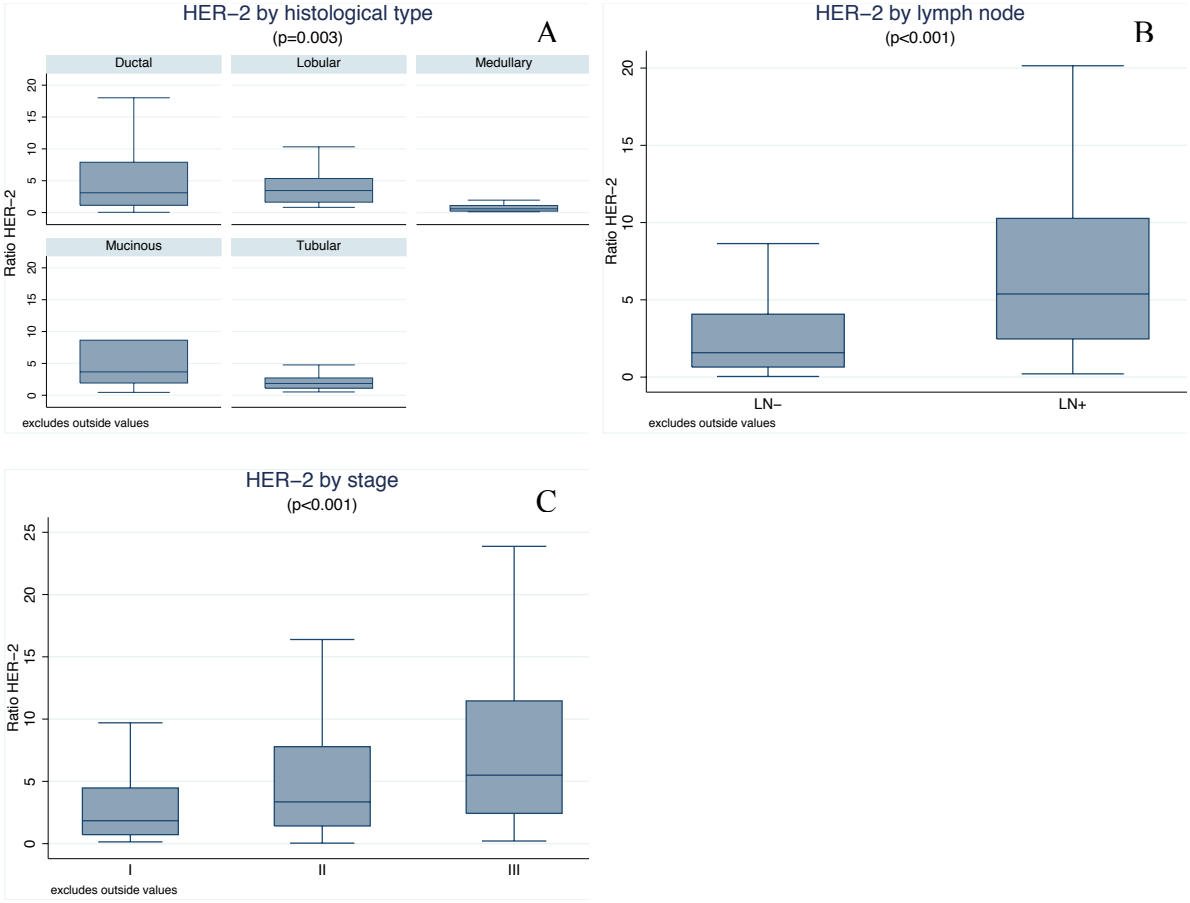
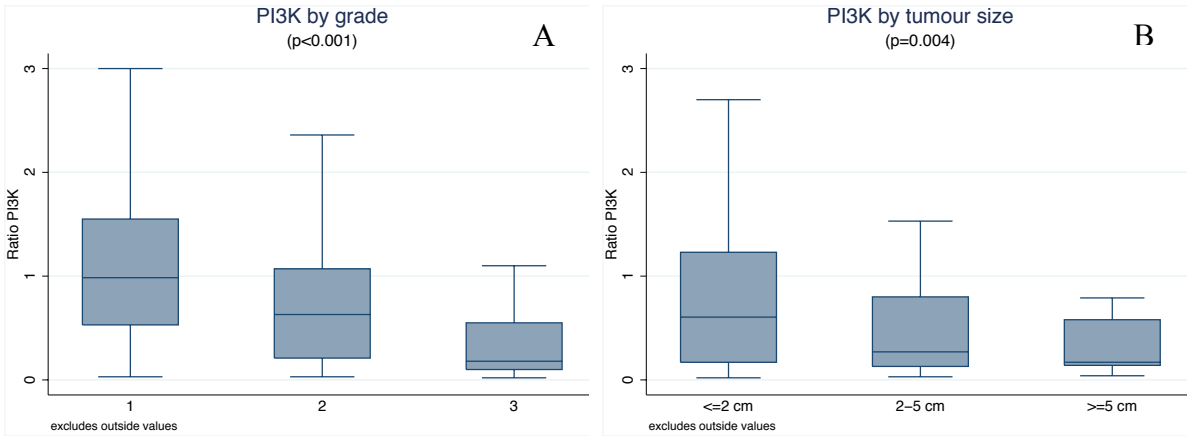


Fig. 15. Box plot representing significant HER-2 expression ratio stratifying patients according to histological type of tumour (A), lymph node involvement (B), tumour stage (C). The median value is indicated as a black line in each group.



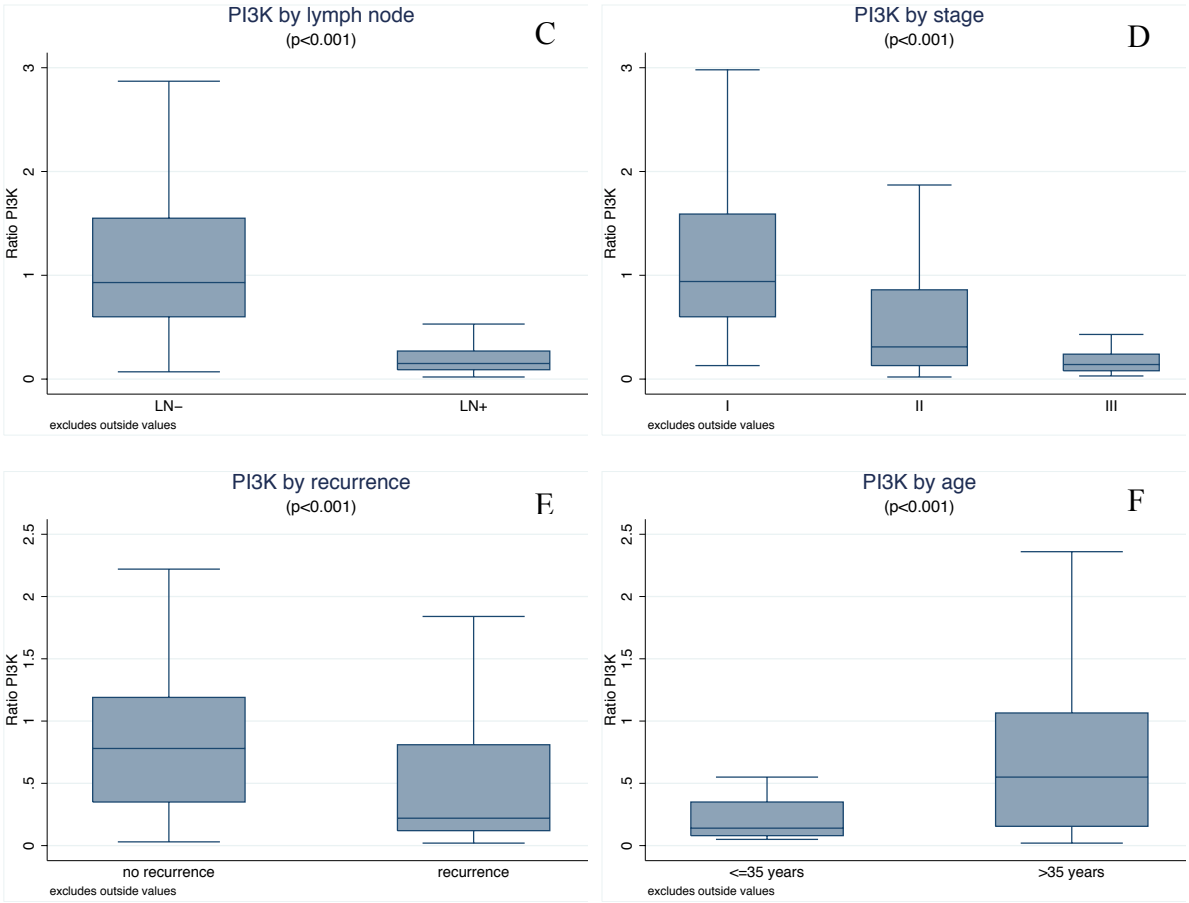
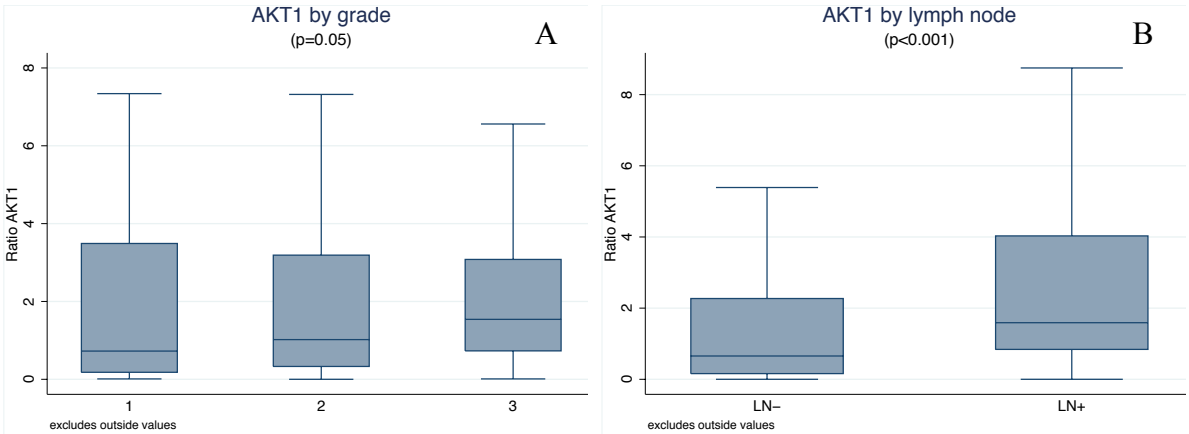


Fig. 16. Box plot representing significant PI3K expression ratio stratifying patients according to histological tumour grade (A), tumour size (B), lymph node involvement (C), tumour stage (D), presence of later recurrence (E), age at diagnosis, (F). The median value is indicated as a black line in each group.



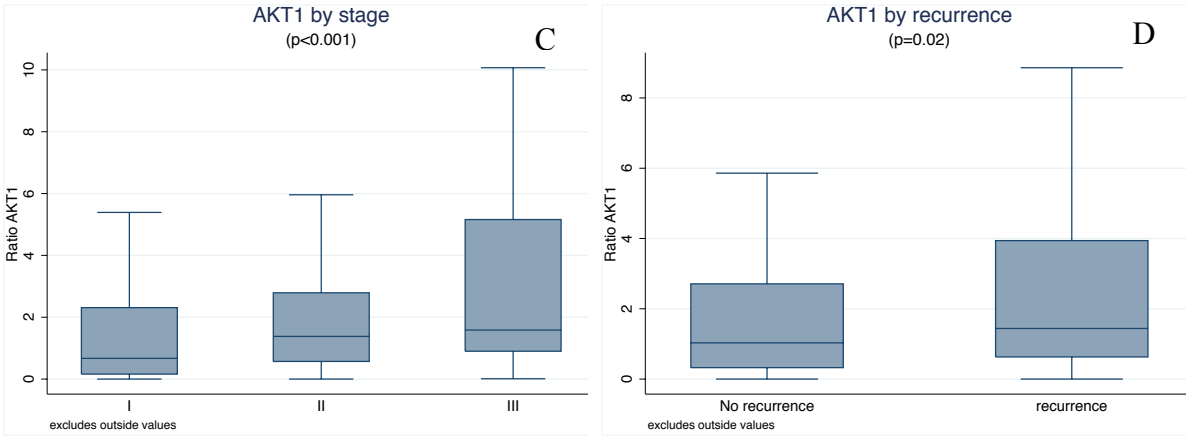
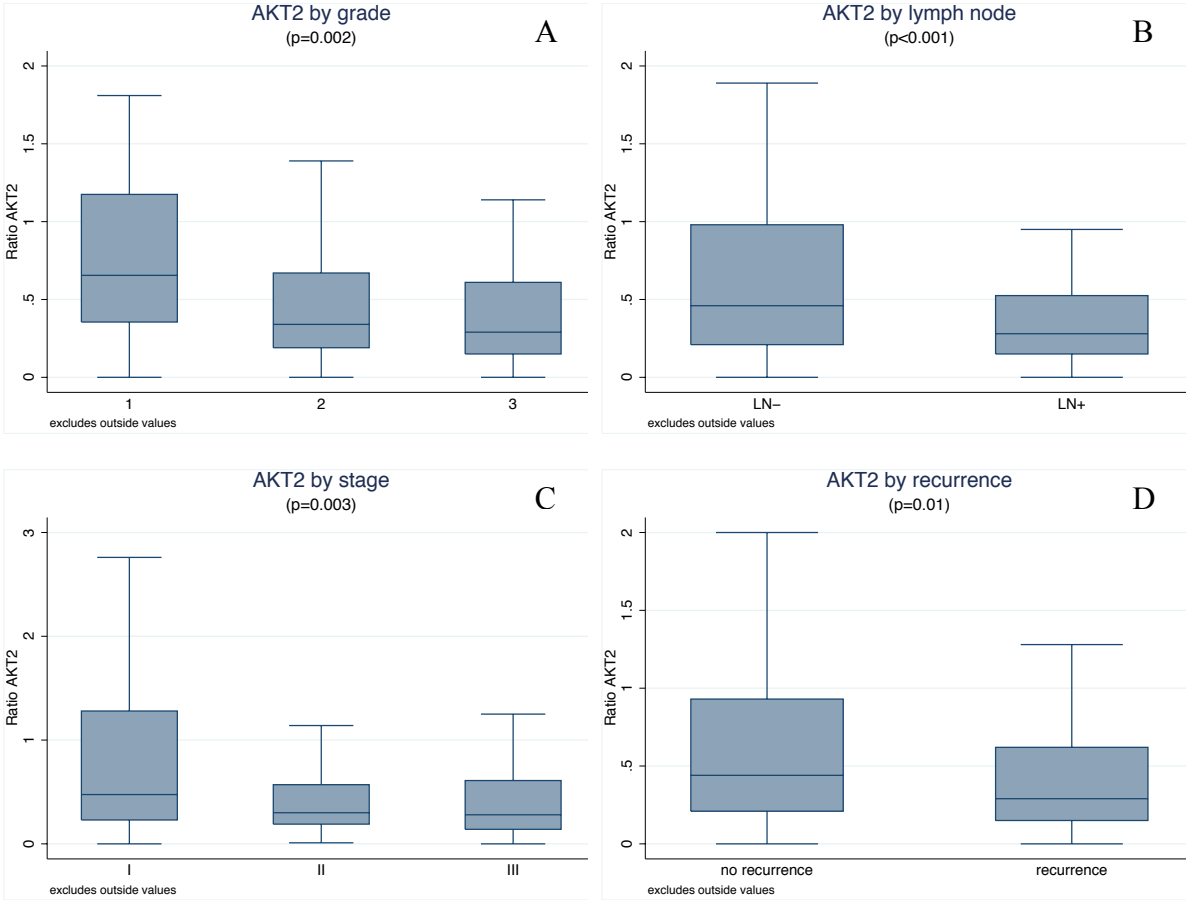


Fig. 17. Box plot representing significant AKT1 expression ratio stratifying patients according to histological tumour grade (A), lymph node involvement (B), tumour stage (C), presence of later recurrence (D). The median value is indicated as a black line in each group.



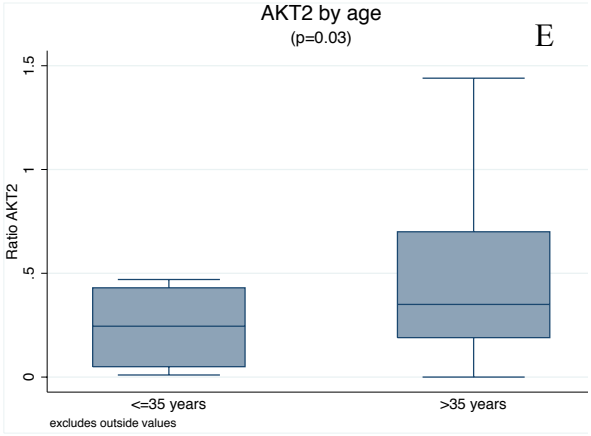
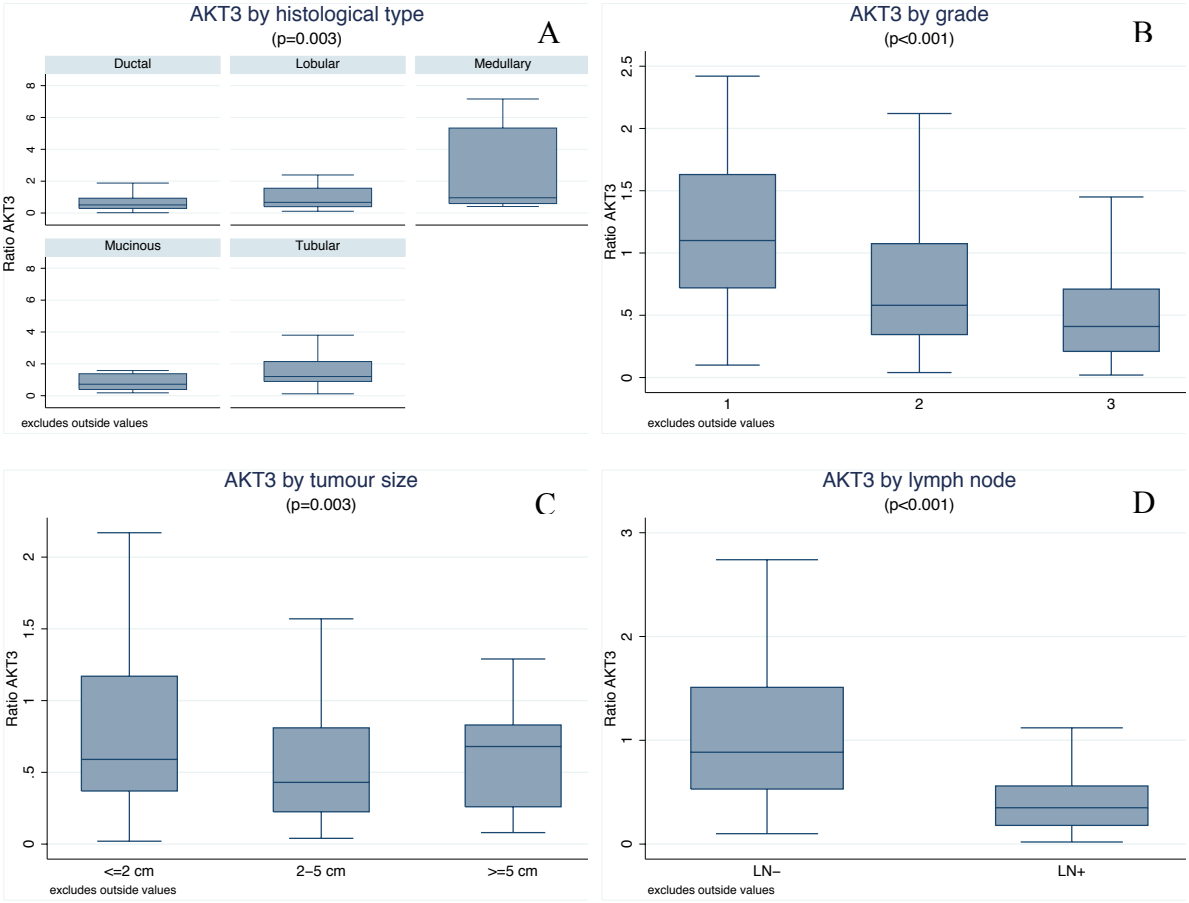


Fig. 18. Box plot representing significant AKT2 expression ratio stratifying patients according to histological tumour grade (A), lymph node involvement (B), tumour stage (C), presence of later recurrence (D), age at diagnosis (E). The median value is indicated as a black line in each group.



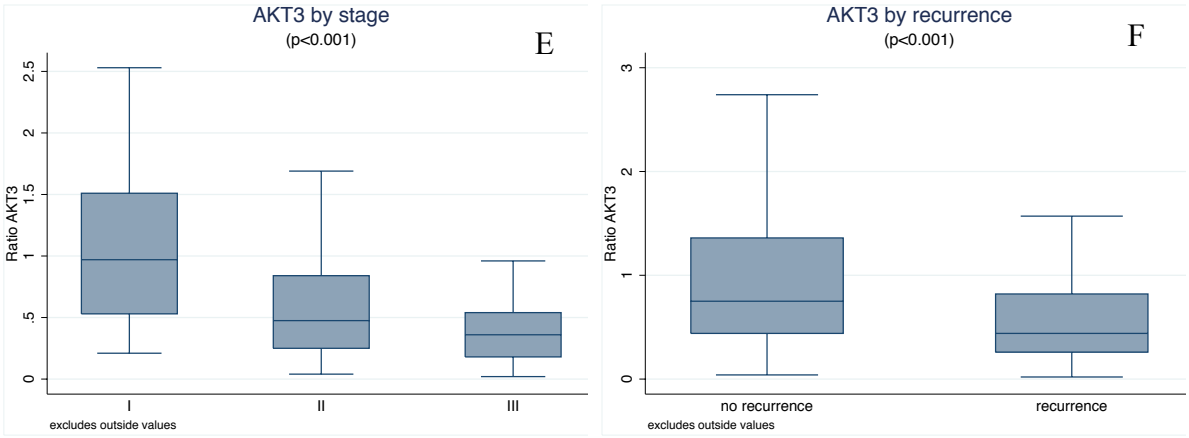


Fig. 19. Box plot representing significant AKT3 expression ratio stratifying patients according to histological type of tumour (A), histological tumour grade (B), tumour size (C), lymph node involvement (D), tumour stage (E), presence of later recurrence (F). The median value is indicated as a black line in each group.

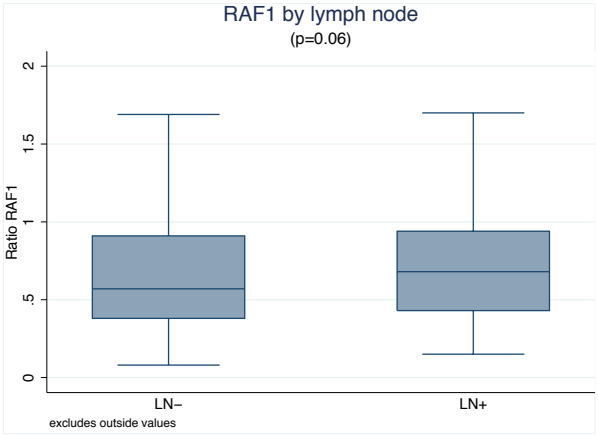


Fig. 20. Box plot representing significant RAF1 expression ratio stratifying patients according to lymph node involvement. The median value is indicated as a black line in each group.

4.7.3 CK8

Higher CK8 expression was observed in: mucinous tumours (p=0.04) (Fig. 21A), intermediate size tumours (p=0.01) (Fig 21B), LN+ patients (p<0.001) (Fig 21C), stage III tumours (p<0.001) (Fig. 21D). Conversely, CK8 expression did not discriminate over tumour grade (p=0.1), number of positive nodes (p=0.3), presence or absence of recurrence (p=0.3) and age at diagnosis (p=0.5).

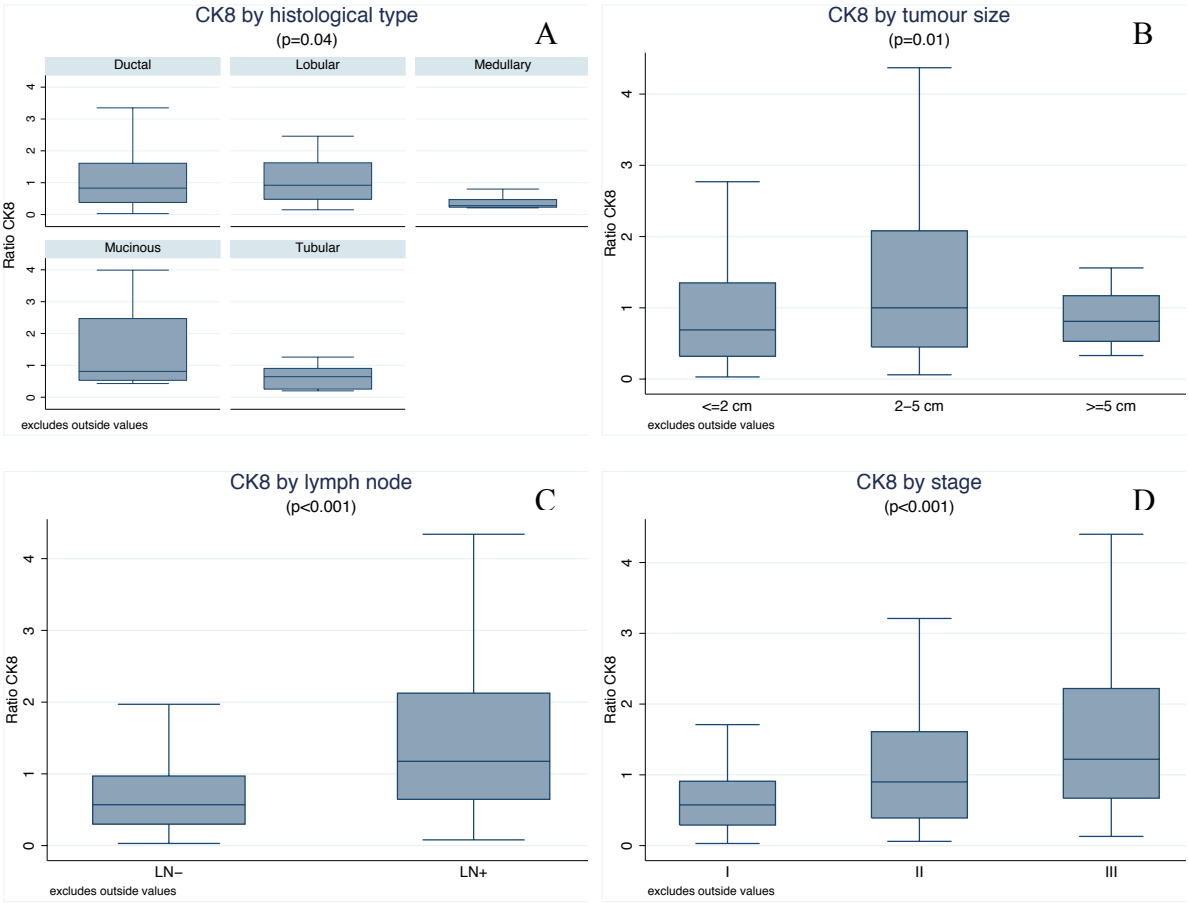


Fig. 21. Box plot representing significant CK8 expression ratio stratifying patients according to histological type of tumour (A), tumour size (B), lymph node involvement (C), tumour stage (D). The median value is indicated as a black line in each group.

4.8 Gene Expression in Molecular Classes

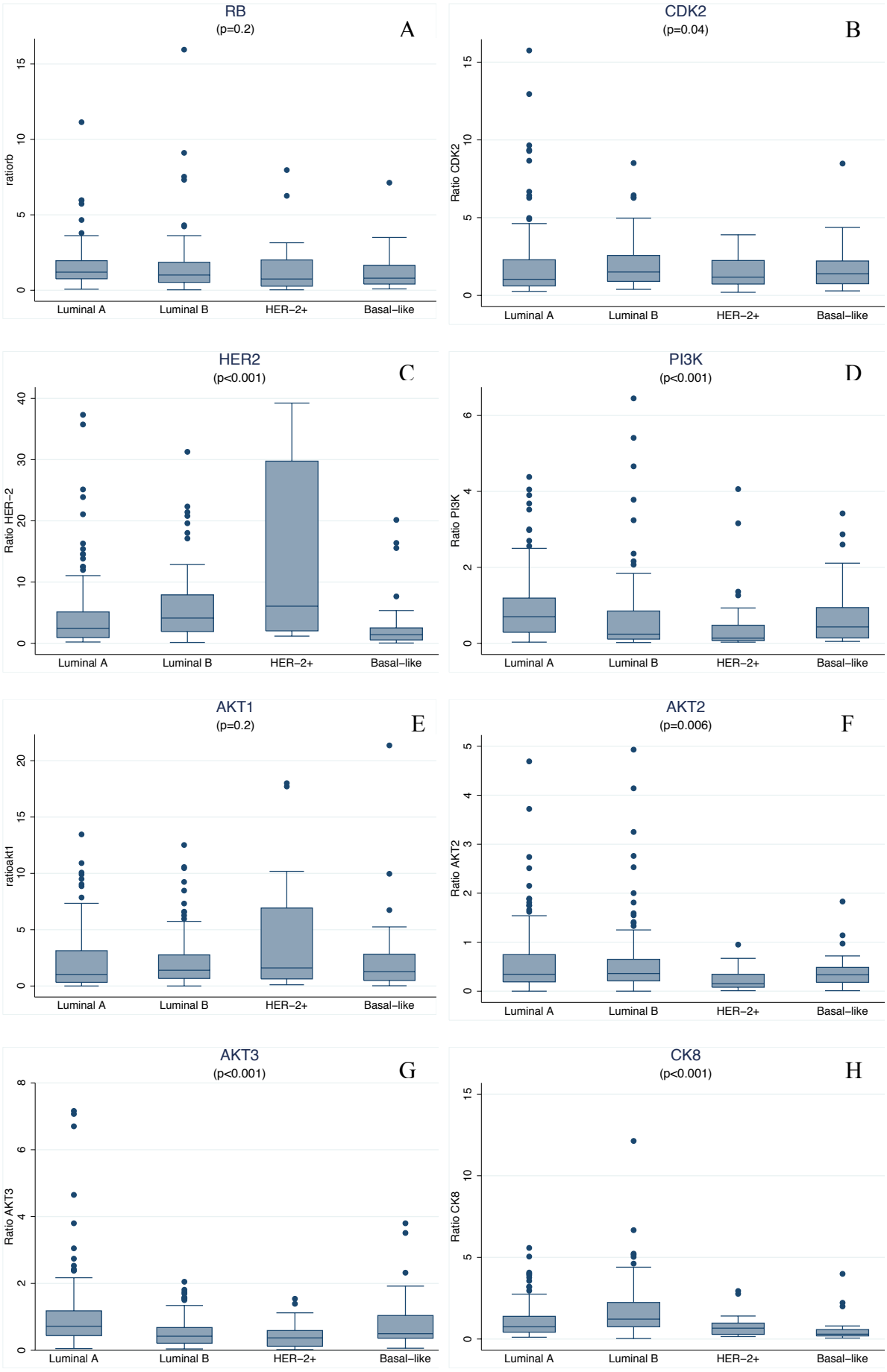
Subdividing patients into the 4 molecular classes we observed a significant difference in the expression levels of: CDK2 ($p=0.04$) HER-2 ($p<0.001$), PI3K ($p<0.001$), AKT2 ($p=0.006$), AKT3 ($p<0.001$) and CK8 ($p<0.001$) among molecular subtypes. No different distribution for RB ($p=0.2$), AKT1 ($p=0.2$) and RAF-1 ($p=0.3$) was observed. The corresponding box plot are shown in Fig. 22A-I. A positive correlation was found between HER-2 protein by IHC and HER-2 mRNA level (Spearman's $\rho=0.51$ and $p<0.01$). No correlation for CK8 expression at the protein and mRNA level was found ($\rho=0.3$).

CDK2 was significantly higher expressed in luminal B in comparison to luminal A ($p=0.006$). No significant differences between the other classes were observed. Higher mRNA levels of HER-2 were predominantly found in HER-2+ subtype with respect to

luminal A, luminal B and basal-like (p=0.0001 for all). In addition HER-2 was more expressed in luminal B in comparison to luminal A and basal-like (p=0.0001 for both), but also in luminal A in comparison to basal-like (p=0.04). The highest expression level of PI3K was recorded in luminal A tumours in comparison with luminal B (p=0.0001), HER-2 (p=0.0002) and basal-like (p=0.03). Also PI3K expression was observed slightly higher in basal-like with respect HER-2 (p=0.06). Carcinomas classified as HER-2+ showed the lowest AKT2 expression in comparison to luminal A (p=0.0009), luminal B (p=0.001) and basal-like (p=0.02). No differences between luminal A and B (p=0.7), between luminal A and basal-like (p=0.3), neither between luminal B and basal-like (p=0.5) was found. AKT3 was higher in luminal A with respect to luminal B (p=0.001) and HER-2 (p=0.0002), nor with basal-like (p=0.3). Moreover, AKT3 was significant higher in basal-like in comparison to luminal B (p=0.04) and HER-2+ (p=0.02). No differences were detected for AKT3 expression between luminal B and HER-2+ (p=0.1). The highest CK8 expression was detected in luminal B class versus luminal A (p=0.0002), HER-2+ (p=0.0003) and basal-like (p=0.0001). Also luminal A expressed higher CK8 in comparison with basal-like (p=0.0001), but not in comparison to HER-2+ (p=0.1). In fact HER-2+ subtypes showed higher CK8 with respect basal-like (p=0.02).

Dividing patients for lymph node involvement, in LN- group only HER-2 and CK8 were still differently distributed among molecular classes (Fig. 23A-B). Namely, HER-2+ and luminal B tumours displayed the highest HER-2 and CK8 expression respectively ($p<0.001$ for both).

In LN+ group three genes were differentially distributed among molecular classes (Fig 24 A-C). HER-2 was prevalently expressed in HER-2+ subtype (p=0.002). AKT2 was mainly expressed in luminal and basal-like tumours in comparison to HER-2+ subtype (p=0.03). CK8 was mainly expressed in luminal tumours ($p<0.001$), but without differences between luminal A and B (p=0.3)


RB
(p=0.2)
A
ratiorb
0
5
10
15
Luminal A
Luminal B
HER−2+
Basal−like
CDK2
(p=0.04)
B
Ratio CDK2
0
5
10
15
Luminal A
Luminal B
HER−2+
Basal−like
HER2
(p<0.001)
C
Ratio HER−2
0
10
20
30
40
Luminal A
Luminal B
HER−2+
Basal−like
PI3K
(p<0.001)
D
Ratio PI3K
0
2
4
6
Luminal A
Luminal B
HER−2+
Basal−like
AKT1
(p=0.2)
E
ratioakt1
0
5
10
15
20
Luminal A
Luminal B
HER−2+
Basal−like
AKT2
(p=0.006)
F
Ratio AKT2
0
1
2
3
4
5
Luminal A
Luminal B
HER−2+
Basal−like
AKT3
(p<0.001)
G
Ratio AKT3
0
2
4
6
8
Luminal A
Luminal B
HER−2+
Basal−like
CK8
(p<0.001)
H
Ratio CK8
0
5
10
15
Luminal A
Luminal B
HER−2+
Basal−like

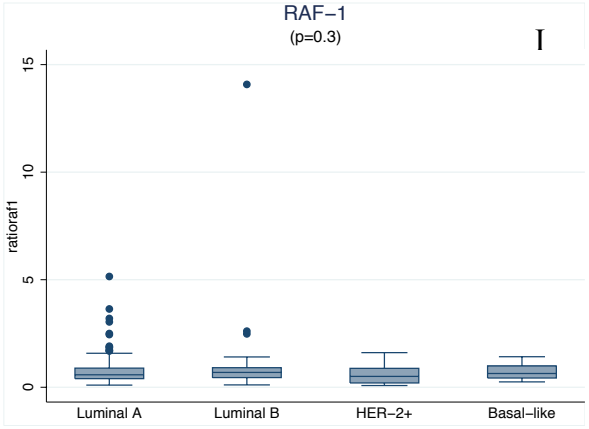


Fig. 22. Box plot representing significant genes expression among molecular classes in the BC patients' cohort (N=303). RB (A), CDK2 (B), HER-2 (C), PI3K (D), AKT1 (E), AKT2 (F), AKT3 (G), CK8 (H), RAF1 (I). The median value is indicated as a black line in each group.

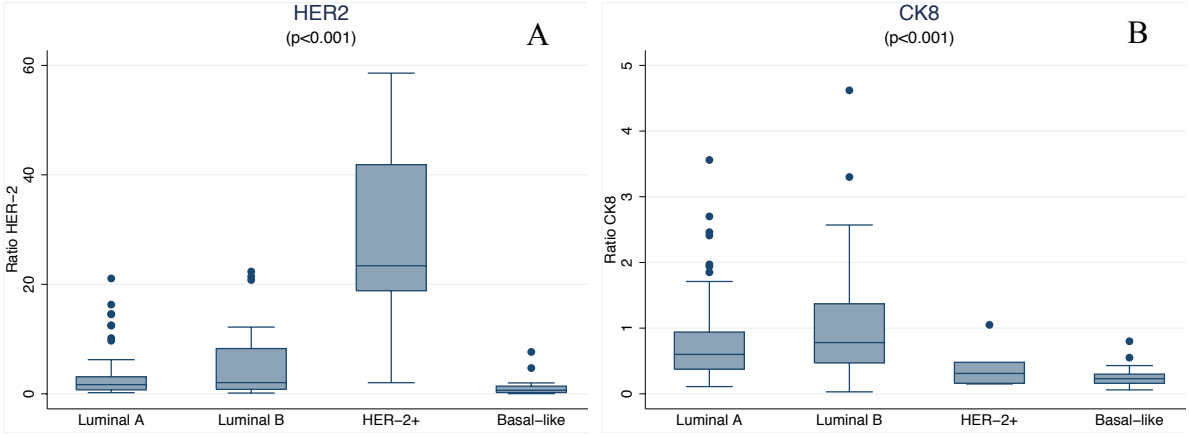
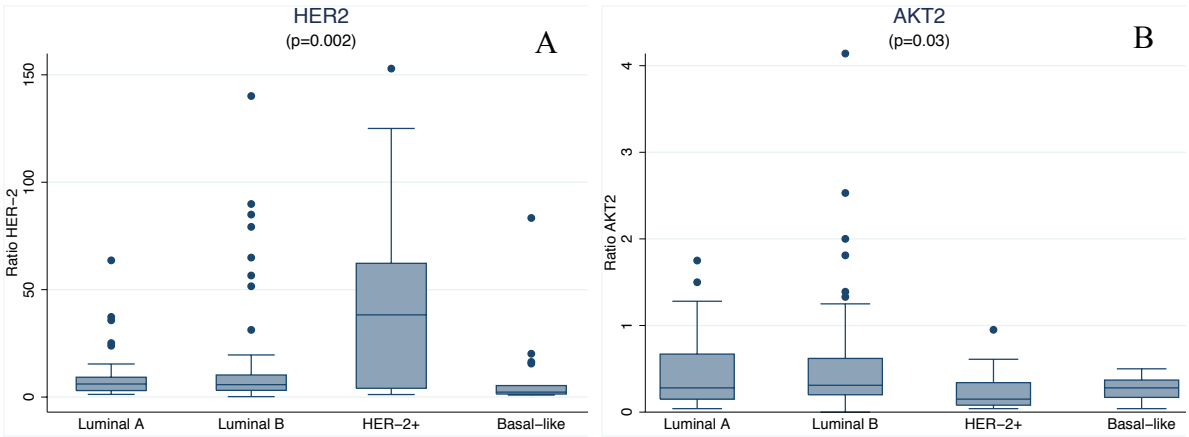


Fig. 23. Box plot representing significant genes expression among molecular classes in the LN- group of BC patients' cohort (N=149). HER-2 (A), CK8 (B). The median value is indicated as a black line in each group.



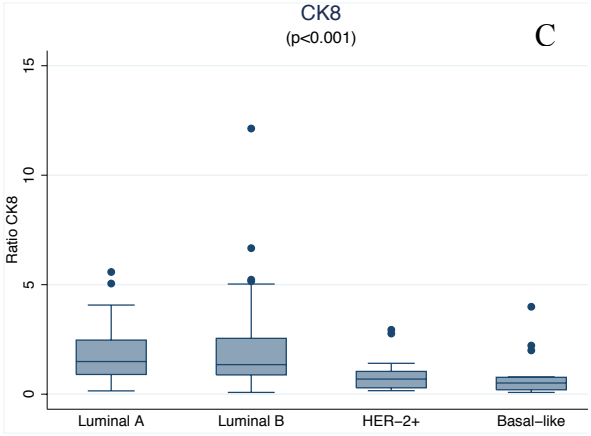


Fig. 24. Box plot representing significant genes expression among molecular classes in the LN+ group of BC patients' cohort (N=153). HER-2 (A), AKT2 (B), CK8 (C). The median value is indicated as a black line in each group.

4.9 Survival Analysis

4.9.1 Candidate Genes

For survival analysis genes' expression level was dichotomized in low expression and high expression with respect to the median value of ratio of each gene, in the entire case study.

The log-rank test revealed that 2 out nine genes affected significantly BC specific survival in the cohort of our patients: AKT1 (p=0.04) and AKT2 (p=0.04) (Fig 25 A-B). In detail 78% of patients with high AKT1 expression level died specifically of BC. A less marked difference was detected for AKT2 expression, 56% of patients with high AKT2 status died from BC versus 44% having a low one.

The expression of the other genes was not significant: RB (p=0. 2), CDK2 (p=0.5), HER-2 (p=0.3), PI3K (p=0.3), AKT3 (p=0.1), RAF-1 (p=0.7), CK8 (p=0.5).

When BC patients were stratified over lymph node involvement, we obtained different results. In LN- group, RB and AKT3 seemed to have a protective effect on patients' survival (p=0.03 and p=0.01, respectively), since higher expression levels were associated to longer survival, as shown in Fig. 26 A-B. Particularly, 64% of patients with a low RB expression had a BC specific death, against 36% with a high RB status. Similarly, 67% of LN- patients with low AKT3 expression died from BC with respect to 33% of women with high AKT3. The other genes were not significantly related to BC patients' survival (CDK2

p=0.4, HER2 p=0.3, PI3K p=0.3, AKT1 p=0.4, AKT2 p=0.4, RAF1 p=0.4 and CK8 p=0.09).

In LN+ group none of the analysed gene showed influence on BC specific patients' survival (RB p=0.8, CDK2 p=0.6, HER2 p=0.6, PI3K p=0.4, AKT1 AKT2, AKT3 and RAF1 p=0.9).

To confirm the results obtained in the Log-rank survival analysis, Cox proportional hazard model was performed using as covariates the pathological variables (histological type, histological grade, tumour size, lymph node involvement, age at diagnosis) and the genes that resulted statistically significant in the Log-rank test. When the entire cohort of BC was analysed, the histological type of tumour (p=0.06), the lymph node involvement (p<0.001) and histological grade (p=0.008) affected significantly BC specific survival, but not tumour size (p=0.7), age at diagnosis (p=0.5), AKT1 (p=0.2) and AKT2 (p=0.8) (probability of the model p=0.000) (Table 16 *Part I)*. Considering solely LN- group, the Cox regression model for the abovementioned pathological variables (with the exception of lymph node) and for RB and AKT3 genes revealed that histological grade (p=0.03) and marginally RB (p=0.08) affected BC outcome (probability of the model p=0.04) (Table 16 *Part II*).

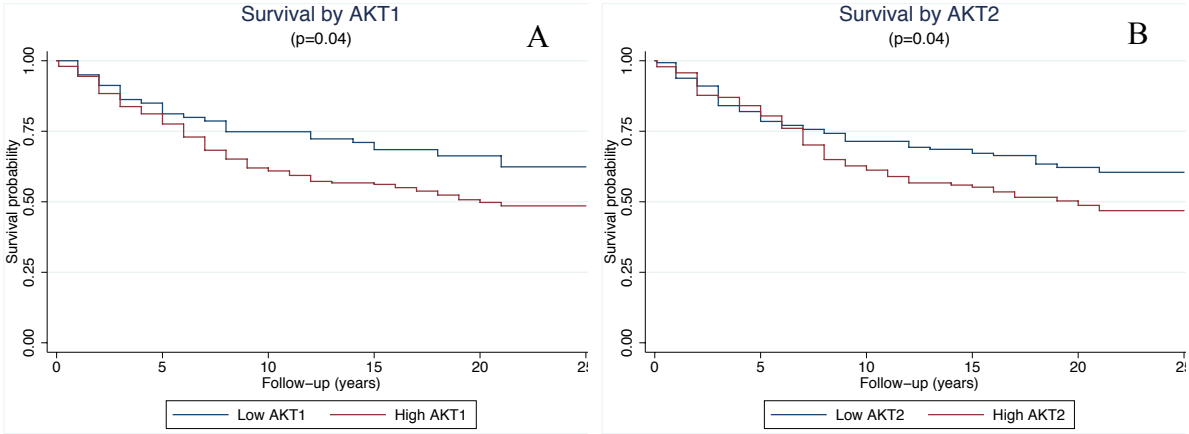


Fig. 25. Kaplan-Meyer survival curves for AKT1 expression (A) and AKT2 expression (B) in the entire cohort (N=305, 128 BC specific deaths). Gene expression was dichotomized in low expression and high expression with respect to the median value of each gene.

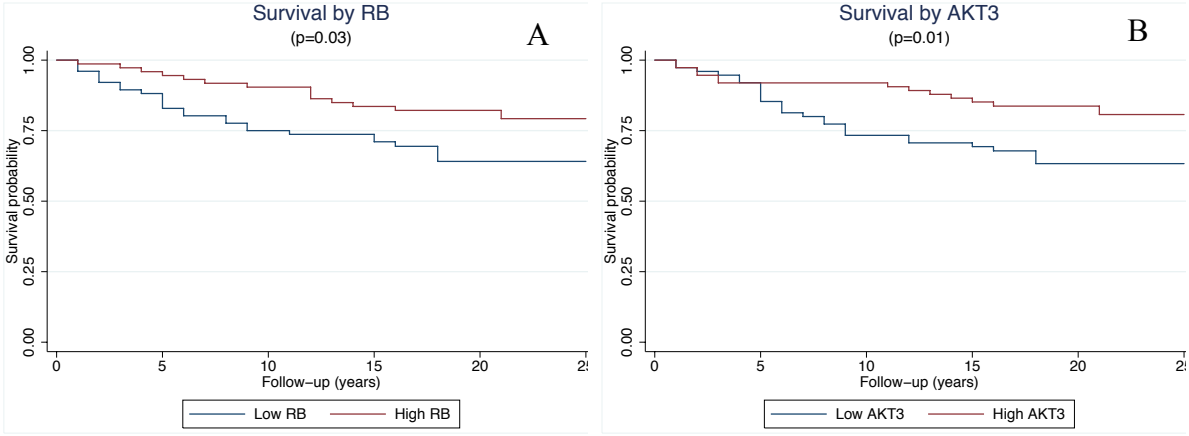


Fig. 26. Kaplan-Meyer survival curves for RB expression (A) and AKT3 expression (B) in the LN- group of patients (N=151, 39 BC specific deaths). Gene expression was dichotomized in low expression and high expression with respect to the median value gene.

Risk Factors	Hazard Ratio	95% CI*	p
Age at diagnosis	1.0	0.9-1.0	0.4
Histological type	0.8	0.6-1.0	0.06
Grade	1.5	1.1-2.0	0.008
Tumour size	1.1	0.8-1.5	0.7
Lymph node	2.6	1.5-4.4	0.000
AKT1	0.7	0.4	0.2
AKT2	1.1	0.7-1.6	0.8

Part I

Risk Factors	Hazard Ratio	95% CI*	p
Age at diagnosis	1.0	0.9-1.1	0.8
Histological type	0.9	0.6-1.2	0.4
Grade	1.7	1.0-2.7	0.03
Tumour size	1.3	0.6-2.4	0.5
RB	0.6	0.3-1.1	0.08
AKT3	0.6	0.3-1.2	0.1

Part II

Table 16. Results of Cox proportional hazard model of BC specific survival for *Part I*: entire cohort (N=305, 128 BC specific deaths); *Part II:* lymph node negative group (N=151, 39 BC specific deaths). (*Confidence Interval, p= level of significance for association).

4.9.2 Candidate Genes and Molecular Classification

The role of the genes as candidate prognostic markers for BC specific survival was investigated among the 4 molecular classes taken one a time. A better patients' survival was found for higher level of CDK2 in basal-like tumours (p=0.06) (Fig. 27), lower level of AKT2 in luminal B (p=0.03) (Fig 28), and higher level of AKT3 in luminal A tumours (p=0.001) (Fig. 29A). In detail, 63% of women classified as basal-like with low CDK2 expression level died from BC; 78% of luminal B patients with higher AKT2 had a BC specific death; 66% of women with low AKT3 died from BC among luminal A patients.

Dividing patients for lymph node involvement, CDK2 was no more significantly related to patient's survival (p=0.2 for both LN- and LN+). AKT2 also lost its prognostic role when it was analysed only in LN- and LN+ group (p=0.2 and p=0.6 respectively). AKT3 influence on BC survival were confirmed both in LN- and LN+ subgroups of luminal A patients (Fig 29B-C). A better overall survival was displayed in patients with high AKT3 status (LN- p=0.005 and LN+ p=0.04). 73% and 60% of luminal A patients with low AKT3 in LN- and LN+, respectively, died of BC.

Cox proportional hazard regression was performed for CDK2 status and the clinical-pathological variables in basal-like tumours, but the model was not significant (p=0.09). For luminal B tumours the regression was run using as covariates the clinical-pathological

data and AKT2 status, in Cox model (p=0.02) only lymph node involvement affected marginally (p=0.06) survival, but not AKT2.

The protective role of high AKT3 status in luminal A tumours was confirmed by Cox proportional hazard regression (p=0.000), as shown in Table 17. Lymph node involvement (p=0.008), tumour size (p=0.004) and histological grade (p=0.02) conferred an increased risk of BC specific death among luminal A patients.

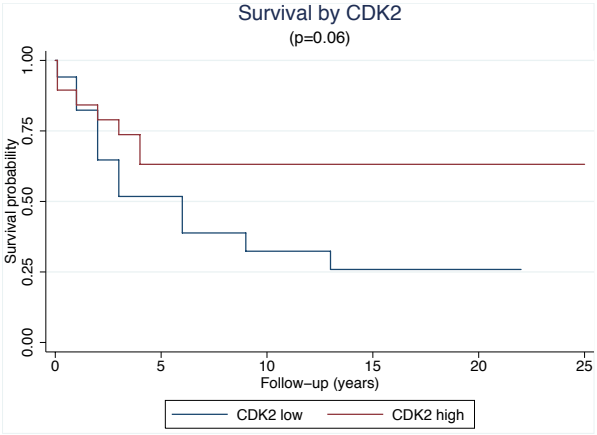


Fig. 27. Kaplan-Meyer survival curves for CDK2 expression over basal-like tumours (N=36, 19 BC specific deaths). Gene expression was dichotomized in low expression and high expression with respect to the median value of CDK2 gene expression.

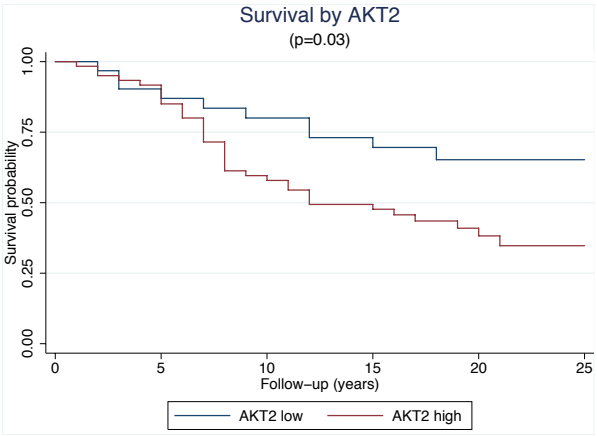


Fig. 28. Kaplan-Meyer survival curves for AKT2 expression over luminal B tumours (N=142, 48 BC specific deaths). Gene expression was dichotomized in low expression and high expression with respect to the median value of AKT2 gene expression.

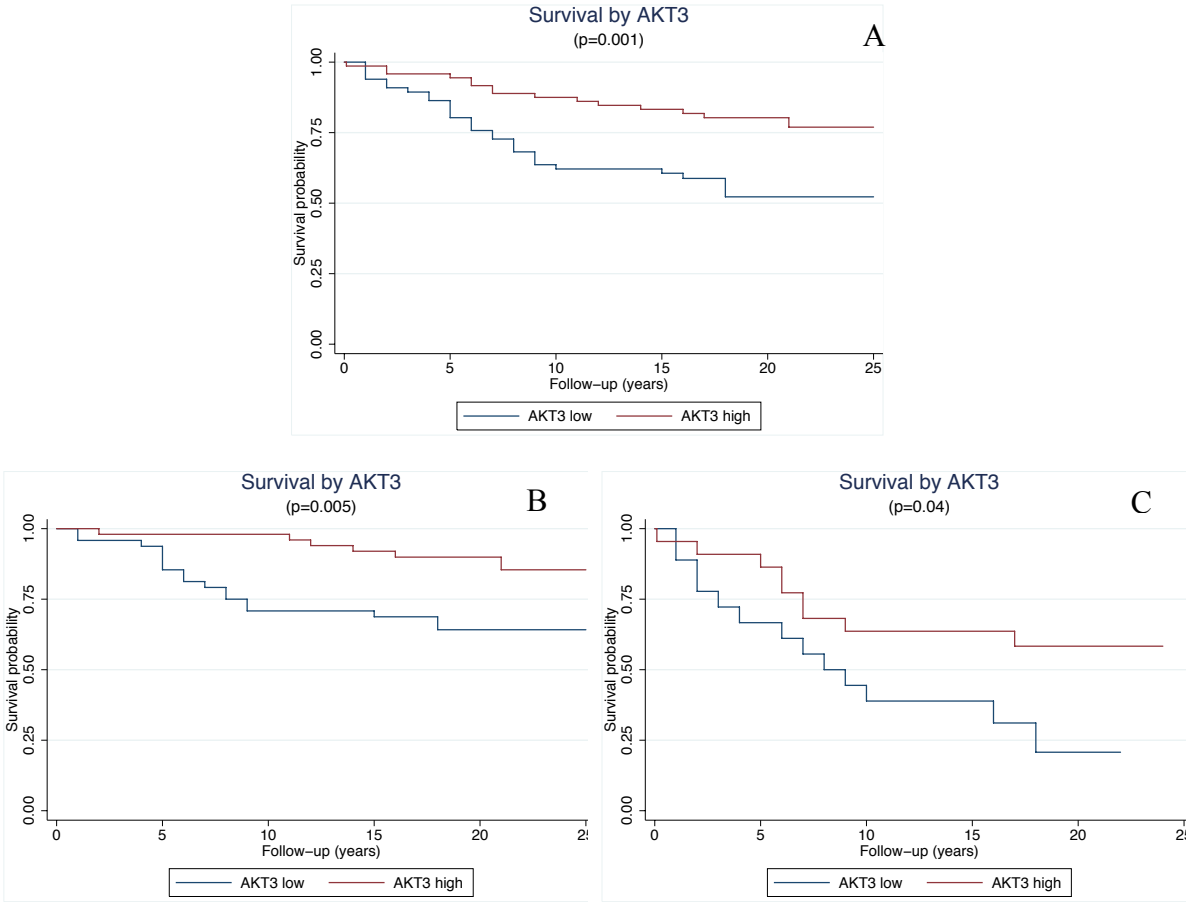


Fig. 29. Kaplan-Meyer survival curves for AKT3 expression: (A) over entire luminal A tumours (A) (N=140, 45 BC specific deaths); (B) over LN- luminal A tumours (N=98, 22 BC specific death); (C) over LN+ luminal A tumours (N=42, 23 BC specific death). Gene expression was dichotomized in low expression and high expression with respect to the median value of AKT2 gene expression.

Risk Factors	Hazard Ratio	95% CI*	p
Age at diagnosis	1.1	0.9-1.1	0.1
Histologic type	0.8	0.4-1.4	0.3
Grade	1.8	1.1-2.9	0.02
Tumour size	2.0	1.3-3.3	0.004
Lymph node	2.3	1.3-4.4	0.008
AKT3	0.4	0.1-0.8	0.01

Table 17. Results of Cox proportional hazard model of BC specific survival over luminal A patients (N=140, 45 BC specific deaths) (*Confidence Interval, p= level of significance for association).

4.10 AKTs Expression in Cells Lines

The mRNA expression of AKTs in cells revealed a slight higher AKT1 expression in MCF-7 cells (Ratio 3.49) with respect to SKBR-3 and MDA MB-231 (Ratio 0.58 and 0.59 respectively, even if not significant (p=0.2). AKT2 expression was similar among the three cell lines (Ratio 0.85, 0.96, 1.22 for MCF-7, SKBR-3 and MDA MB 231) (p=0.6). AKT3 expression was higher in SKBR-3 (Ratio 5.28) and MDA MB-231 (Ratio 4.10) with respect MCF-7 (Ratio 0.05), but differences were not significant (p=0.1). The histogram representing gene expression of the three AKT isoforms among BC cell lines is reported in Fig. 30.

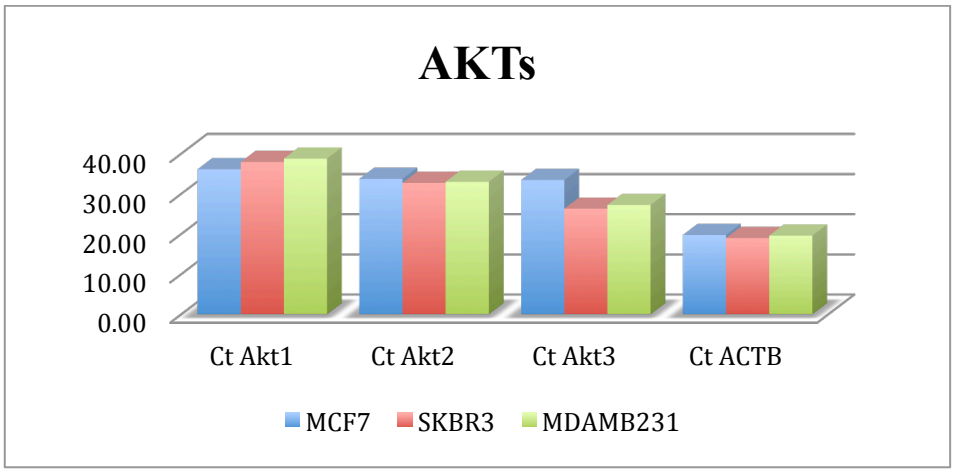


Fig. 30. Gene expression reported as Ct (cycle threshold) of AKT1, AKT2, AKT3 and housekeeping gene β-Actin (ACTB) in the three BC cell lines MCF-7, SKBR-3 and MDA MB 231.

5 DISCUSSION

To date, most of the information on patient's prognosis is based on the combination of pathology and immunohistology, which are closely connected to clinical data. However, patients who present similar clinically and pathologically tumours can show very different clinical outcome with respect to survival and response to therapy. Actually, tumour morphology is not able to reflect completely and explain the different alterations in tumour cells (Andre and Pusztai, 2006; Dietel and Sers, 2006). In recent years, many efforts have been made to deepen understanding the molecular alterations distinctive of each tumour. In this context the molecular biomarkers represent the connection between molecular biology and clinical diagnosis/prognostication.

This thesis aims to investigate on the possibility of combining classical pathological features with new candidate biomarkers for the prognostication of breast cancer. This target was pursued by performing molecular analyses on formalin fixed paraffin embedded tissues specimens in order to directly compare experimental results with the clinical data and follow up.

5.1 Prognostic Role of Clinical and Pathological Features

Treatment decisions after surgery for BC patients are still based largely on clinical and pathological parameters. These parameters include: age at diagnosis, histological type of tumour, tumour grade, tumour size, lymph node involvement and presence of later recurrences (Colombo et al., 2011; Rakha and Ellis, 2011; Soerjomataram et al., 2008). Their prognostic values were assessed in a cohort of 305 BC young patients, age at diagnosis ≤ 55 years, characterized by a long follow-up.

As a first point, we separated patients according to lymph node involvement and compared the clinical pathological parameters between these two groups. It emerged that patients with age at diagnosis ≤ 35 years were more frequent in the LN+ group, according to the fact that very young women exhibited a more aggressive disease and poor survival (Kroman et al., 2000; Soerjomataram et al., 2008). Moreover, some histological types of tumours were significantly associated to the lymph node involvement. For instance, we observed that medullary and tubular tumours belonged mostly to LN- group, according to Ellis et al. who

reported that axillary node metastases do not occur frequently in tubular tumours and are detected only in less than 10% of medullary carcinomas (Ellis et al., 2003). Otherwise, ductal tumours represented the vast majority of LN+ patients. Another variable that resulted strictly associated to the lymph nodes status was the tumour grade. Most low-grade tumours were LN- (G1 rate: 22% in LN- and 3% in LN+), while high-grade tumours were found in LN+ group (G3 rate: 21% and 59% in LN- and LN+ respectively), as previously reported (Galea et al., 1992; Rakha et al., 2008a; Rakha and Ellis, 2011). Tumour size, an important BC prognostic factor, was directly related to the presence of nodal metastasis, as previously reported (Edge et al., 2010a; Rakha and Ellis, 2011). Furthermore, from our analysis it resulted that tumour larger than 5 cm were mainly in LN+ group, while almost three quarter of lymph node negative patients showed tumours less than 2 cm in size. We also found that only one third of LN- patients developed later recurrences, while more than two third of LN+ women recurred, in agreement with Rakha et al. (Rakha et al., 2010).

As expected, the Log-rank test revealed a strong association between those clinical-pathological features and patients' outcome, as already reported by several authors (Elston and Ellis, 1991; Rakha et al., 2010; Soerjomataram et al., 2008; Weigelt et al., 2010b). Histological type of tumour ($p=0.004$), tumour grade ($p<0.001$), tumour size ($p=0.02$), lymph node involvement ($p<0.001$), number of positive nodes ($p=0.01$) and presence of recurrences ($p<0.001$) affected significantly BC specific survival in our cohort of patients. We did not find a significant influence of age at diagnosis on BC specific survival. Nevertheless, multivariate Cox regression analysis confirmed an independent influence on BC specific survival for histological type of tumour ($p=0.04$), tumour grade ($p=0.006$) and lymph node status ($p<0.001$). In particular, the histological type of BC showed a protective role on patient's survival, while positive lymph node involvement and tumour grade were poor prognostic factors. Similar independent prognostic value was recorded for tumour differentiation and lymph node involvement, in agreement with Rakha et al. who reported that in early BC stages the histological grade had a prognostic value equivalent to that of LN status, even higher than tumour size (Rakha et al., 2008a). Therefore, we strongly believe that beyond lymph node status, tumour grade is highly crucial for prognostication especially for those tumours with same stage at diagnosis.

5.2 Prognostic Role of Molecular Classification

Gene expression profiling through microarray technology has led to the establishment of a molecular taxonomy as a tool for supporting the clinical management of patients (Colombo et al., 2011; Perou et al., 2000; Sorlie et al., 2001; Sotiriou and Pusztai, 2009). Accordingly, breast tumours can be grouped into at least four "intrinsic" molecular subtypes: two luminal-like ER-positive subtypes, and two ER-negative subtypes, the HER-2 overexpressing and the basal-like subtypes, which is also defined triple negative (TN) because of negativity to ER, PR and HER2 (Valentin et al., 2012). However, given to the limit of microarray-based molecular classification in clinical practice, immunohistochemical-based classification using the molecular markers ER, PR, HER2, and Ki67 (with a cut-off of 14%) has been proposed as a surrogate of the molecular classification (Goldhirsch et al., 2011; Gruver et al., 2011; Guiu et al., 2012; Muller et al., 2011; Park et al., 2012; Tang et al., 2009; van de Rijn et al., 2002). Nonetheless, molecular classification remains a working model, since a consensus on the definition of molecular classes and on standardization of methodologies has not yet been reached (Geyer et al., 2012; Rakha and Ellis, 2011; Tang et al., 2009).

In this study, to molecularly classify the tumours we used a panel of six biomarkers in IHC assay: ER, PR, HER-2, Ki67, CK8, CK5/6 (Gruver et al., 2011). In addition vimentin expression was analysed in all tumour sample, because it has been recognized as marker for epithelial-to-mesenchymal transition (Satelli and Li, 2011). The distribution of the molecular classes among our cohort of patients was similar to others studies (Dawood et al., 2011; Park et al., 2012; Weigelt et al., 2010a; Wiechmann et al., 2009): luminal tumours representing the majority of cases with 80% (46% and 34% luminal A and luminal B respectively), followed by basal-like (12%) and HER-2+ subtype (8%).

Similarly to other studies we observed significant association between the clinical-pathological features and the molecular subtypes (Gruver et al., 2011; Sorlie et al., 2001; Tang et al., 2009; Weigelt et al., 2010b). In detail, most of low-grade tumours displayed features of luminal A, conversely high-grade were mainly luminal B, HER-2 and basal-like subtypes, in agreement with previous results (Livasy et al., 2006; Tang et al., 2009; Weigelt et al., 2010a). Regarding BC histology, tubular and mucinous carcinomas, which are associated with a favourable prognosis (Ellis et al., 2003), were mostly luminal tumours, as

well as lobular. Conversely, the more aggressive HER-2 subtype and basal-like were mostly ductal, as already reported (Garcia Fernandez et al., 2012; Valentin et al., 2012). We observed that most of the medullary tumours belonged to the basal-like group, in agreement with previous reports (Soerjomataram et al., 2008; Valentin et al., 2012; Weigelt et al., 2010b). Patients younger than 35 years at diagnosis, often associated to more aggressive tumours (Soerjomataram et al., 2008), belong principally to luminal B and HER-2+ subtypes.

The distribution of molecular subtypes between the LN- and LN+ groups revealed significant differences ($p<0.001$). In particular, luminal A type was mainly observed in LN-, while luminal B was the most frequent type in the LN+ group. Luminal B and HER-2+ tumours were significantly associated to positive lymph nodes in comparison to luminal A and HER-2, in agreement with previous result (Wiechmann et al., 2009). Luminal B type, indeed, displays an intermediate rate of positive loco-regional lymph nodes between luminal A and HER-2+ (Wiechmann et al., 2009). Conversely, basal-like were equally distributed between LN- and LN+ groups. Some recent studies reported that the basal-like, despite its poor prognosis, was associated to a lower incidence of positive lymph node involvement (Crabb et al., 2008; Foulkes et al., 2004), indicating that its intrinsic aggressive behaviour is partially independent from the clinical stage.

No significant differences for size of tumour among molecular classes were found, in disagreement with other studies, which showed an association between TN subtype and larger tumour size (Garcia Fernandez et al., 2012; Park et al., 2012; Wiechmann et al., 2009). However, the overall clinical stage well explained the different aggressiveness of molecular classes. Notably, luminal A tumours were mainly of stage I, luminal B of stage II, and HER-2 tumours of stage III. Basal-like type represents a distinct group of tumours, which exhibited heterogeneous biological behaviour that can not be thoroughly explained by means of traditional clinical-pathological features (Rakha et al., 2008b).

Regarding relapses, we observed that more than half luminal A tumours did not exhibited recurrences, while HER-2+ tumours presented the highest percentage of recurrences, as already reported by Park *et al*. (Park et al., 2012). Luminal B and basal-like showed a similar frequency of recurrences. Nonetheless, luminal A and B exhibited comparable frequencies of recurrences if matched by lymph node involvement: in LN- cases luminal A showed 32% and luminal B 27% of recurrences, in LN+ 65% and 66% respectively. This data were also supported by BC specific survival. This result reflects the fact that luminal A

and B cancers differ only for the rate of Ki67 and HER-2 overexpression, both characteristic of higher aggressiveness. In comparison to luminal A, Luminal B tumours were often negative to PR ($p=0.03$), but positive to vimentin ($p=0.003$), both characteristics seem to be associated to worse outcome. In agreement, it was previously reported that patients with high levels of PR within their tumour have a better outcome than low expressors, in patients treated with tamoxifen (Weigel and Dowsett, 2010). Recent studies have reported that vimentin plays a major role in the EMT process of BC, and its knockdown resulted in a decrease in genes linked with BC invasion. Indeed, vimentin expression was prevalently found in high-grade ductal carcinomas (Satelli and Li, 2011).

About the metastatic sites, we observed that basal-like as well as HER2+ subtypes presented higher rate of brain metastasis in comparison to luminal A tumours, in agreement with other reports (Crabb et al., 2008; Rakha et al., 2008b). Conversely, HER2+ tumours were not prone to develop bone metastases.

To our knowledge we are the first to investigate the BC specific survival associated to molecular classes in a population with up to 25-years of follow-up. As expected, we observed that BC specific survival analysis identified two well separate groups of tumours: the ER-positive and ER-negative. Luminal A had the best outcome, luminal B showed an intermediate overall survival, while basal-like and HER-2 types showed the shortest survival ($p<0.001$), as already reported (Garcia Fernandez et al., 2012; Geyer et al., 2012; Park et al., 2012; Perou et al., 2000; Sorlie et al., 2001). Patients of this study were enrolled before the availability of trastuzumab and, as expected, HER2+ cancers exhibited the worst survival.

The longer survival for both luminal A and luminal B tumours was also confirmed by Cox proportional hazard model. Taken together our results, we stress that luminal A and luminal B belong to the same group of tumours, their clinical and pathological differences are mainly related to different tumour stage at diagnosis. This is also confirmed by the fact that for luminal tumours the most important predictor of prognosis remained the lymph node involvement and the histological grade.

Although luminal, or better ER+ tumours are usually referred as less aggressive, they undergo frequently over 10 year long-term recurrences (Park et al., 2012). It is important to underline that referring to the period of diagnosis of our patient's cohort (1983-1993), LN-

patients were not treated with cytotoxic chemotherapy, according to regimens of those years. As a consequence, we observed the real biological history of LN- luminal B tumours. In our case study 10% of patients having luminal tumours died from BC even after 10 year of follow-up. It is possible to hypothesize that those patients characterized by high mitotic index, lymph node negative, vimentin positive, small tumour size and lower grade harbour some specific biological properties that slow down progression of carcinoma and may postpone the BC specific death even after 20 years. This finding could be possibly explained by the dormancy of cancer cells. It is believed that dormant tumour cells are resistant to adjuvant therapies and are responsible for recurrences and cancer death with mechanisms of transition from latency to proliferative status that are, to date, largely unknown (Paez et al., 2012). In the recent St. Gallen guidelines it is reported that adjuvant chemotherapy for women with luminal B tumours should be recommended, while it could be skipped in luminal A (Goldhirsch et al., 2011). We partially disagree with this indication because we observed that, even if at lower rate, also luminal A patients presented long-term recurrences. Therefore, we strongly believe that monitoring and follow-up program could be different for subtype of cancers and, in case of luminal tumours prolonged controls are recommended because of the possibilities of later relapses.

To conclude, we think that the use of the traditional biomarkers (ER, PR, HER-2) is essential for BC characterisation and prognostication in association with clinical pathological features. Molecular classification represent a mere reflection of hormone status, HER-2 and proliferation status, as already reported (Rakha and Ellis, 2011). Moreover, we stress that luminal A and B represents the same group of tumours, but observed at different time of progression. Additionally, we believe on the usefulness of 10% cut-off for Ki67 in substitution of 14% threshold (Cheang et al., 2009) to discriminate between luminal tumours. To support our opinion, also the IMPAKT 2012 working group judged as inadequate the clinical utility of the 14% cut-off for discriminating luminal A /luminal B tumours as a basis to decide the use of adjuvant chemotherapy (Guiu et al., 2012).

5.3 Prognostic Role of Markers in Metastatic Lymph Nodes

Many clinical and pathological factors, as tumour grade, age at diagnosis (Hayes, 2005; Patani et al., 2007), tumour size, as well as hormone receptor status and HER-2

amplification status (Cavalli, 2009) were associated to lymph node metastasis, as shown in this study. However, many of the abovementioned features failed in predicting the development of later recurrences (Cavalli, 2009; Zhou et al., 2012).

To understand the molecular alteration connected to the metastatic progression, we compared the phenotype of primary tumour cells with their corresponding metastatic cells colonizing loco-regional nodes. By comparing the expression of markers ER, PR, HER-2, Ki67, CK8, CK5/6 and vimentin by IHC, we found that HER-2, Ki67, CK8 and vimentin positivity was frequently lost at the lymph node level, but an increment of CK5/6 positivity was observed in the metastatic site. The lower mitotic proliferation index found in the metastatic lesion could be explained by the adverse microenviroment of the lymph node for the replication of cancer cell. In contrast with our result, Park et al. reported an increased Ki67 positivity in metastatic lesion compared to the primary tumour (Park et al., 2007). This discrepancy it could be related to different cut-off level for positivity and different methods of analysis. In our study, indeed, we performed microdissection through TMA technique to isolate and analyse tumour areas from the lymph node, while Park et al. processed the entire slide. Positivity to ER and PR was similar in primary and matched metastatic tumours in our cohort, in agreement with a previous work (Tawfik et al., 2013). However, on average the number of positive nuclei for both ER and PR was higher in the metastatic counterpart: the mean of ER nuclei was 54% and 65% in primary and metastatic lesion, respectively; for PR it was 40% and 53% respectively.

The expression of the markers was combined in order to classify molecularly both primary tumour and the corresponding metastatic lesion. Molecular classification was partially maintained during the process of metastatization, although the expression of CK8 and CK5/6 significantly changed. We could hypothesize that only some cells of the primary tumours are able to colonise lymph nodes, and those cells could likely be characterised by the basal/myoepithelial phenotype (CK5/6+), which represents the proliferating counterpart of breast epithelia(Iyer et al., 2013; Moll et al., 1982).

In order to investigate the prognostication of the analysed markers at the lymph node level, we studied BC specific patients' survival separately in primary and metastatic sites. Surprisingly, CK8 expression assessed in lymph node metastatic specimens resulted to be a marker of good prognosis. The number of analysed specimens allows reaching a power of the study of 0.9985, supporting the strength of this finding. This result was also confirmed by Cox' multivariate hazard regression model ($p<0.001$). To our knowledge, this is the first

study showing that positivity to CK8 in metastatic lymph nodes is significantly associated to longer patients' survival. Although this is a preliminary result, which should be confirmed in a larger cohort of patients, we believe that the assessment of CK8 in regional metastatic lymph nodes by means of IHC is feasible and reliable in clinical diagnostic.

5.4 Prognostic Role of Candidate Genes

Most candidate genes analysed in this study resulted associated with clinical and pathological features, except RAF1. In detail, AKT3 expression was significantly associated to lobular, mucinous and tubular carcinomas, which presented a usually good prognosis (Ellis et al., 2003). HER-2, CDK2 and CK8 were significantly more expressed in ductal tumours, often associated to a worse outcome (Ellis et al., 2003). PI3K, AKT2 and AKT3 expression was significantly associated to low-grade tumours, while CDK2 and AKT1 to high-grade cancers. Regarding tumour size, PI3K and AKT3 expression was significantly related to small cancers, whereas CDK2 and CK8 expression to larger ones.

Regarding the relationship between genes expression and lymph node involvement, significant differences between LN- and LN+ groups was found for all the genes, except RAF1. A higher expression of RB, PI3K, AKT2 and AKT3 was detected in LN- group, even in patients without later recurrences. On the contrary, higher expression of CDK2, HER-2, AKT1 and CK8 in LN+ group was found. Data on CK8 were supported from Brotherick et al., who detected higher CK8 mRNA levels in node positive in comparison to node negative patients (Brotherick et al., 1998). CDK2 high expression was significantly associated to the presence of later recurrences in the entire cohort, in agreement with others authors (Bonin et al., 2008; Kim et al., 2008).

Accordingly, we hypothesise that a high mRNA status for RB, PI3K, AKT2 and AKT3 are associated to favourable clinical and pathological features in BC patients. In addition, expression levels of AKT3 and PI3K were significantly higher in living patients in comparison to those dead of breast cancer. The favourable prognostic role of PI3K agrees with recent reports (Dawson et al., 2013; Kalinsky et al., 2009), although in the present study we did not analyse the mutational status of PI3K. On the other hand, the expression levels of CDK2, HER-2, AKT1 and CK8 associated to unfavourable clinical and pathological features. HER-2, AKT1 and CK8 expression levels were significantly higher in patients who died from breast cancer, compared to living patients.

Stratifying patients into the 4 molecular classes, significant associations were found for 6 out of the 9 genes: CDK2, HER-2, PI3K, AKT2, AKT3 and CK8. Lower expression level of CDK2 was detected in luminal A. HER-2 was mainly expressed in HER-2+ subtype, as expected. PI3K expression as well as AKT3 was higher in luminal A tumours in comparison with the others subtypes. Higher AKT2 level was mainly found in luminal and basal-like subtypes and the expression of CK8 was mostly found in luminal B.

The survival analysis for gene's expression in the entire cohort of patients, revealed that high status of AKT1 and AKT2 expression was significantly associated to worse outcome, although AKT2 was previously associated to favourable clinical BC factors.

When we analysed survival only in the LN- group, we observed that high RB, as previously reported (Bosco and Knudsen, 2007; Musgrove and Sutherland, 2010) and high AKT3 were significantly associated to longer BC specific survival. Any gene did not result significant for BC specific survival in the LN+ group.

Low expression of AKT2 in luminal B and high AKT3 expression in luminal A were significantly associated to longer patients' survival. Data on AKT3 agree with a previous study conducted in our laboratory, which was performed using a different assay, namely the quantitative end point PCR with different primers, and also in a different case study, composed of young and old women (Bonin et al., 2008).

In the present study AKT3 seems to be an independent indicator of good prognosis in luminal A tumours. To our knowledge, only few studies reported on the role of isoform-3 of the AKT gene, especially in BC. In agreement with us, a previous study performed on lung cancers showed that high AKT3 and PI3K protein levels in the stroma were independent positive prognosticators for patients' survival suggesting a paracrine effect of AKT on the tumour cells (Al-Saad et al., 2009). Nevertheless, we believe that the high AKT3 mRNA we detected could be related to cancer cells rather than stroma, because of the accurate microdissection performed in this study to avoid lymphocytes and stroma contamination. In disagreement with our findings, Grabinski et al. proved in lung cancers cells the contribution of AKT3 to proliferation, survival and migration of tumour cells (Grabinski et al., 2011). These discordant results may be due to the different techniques used, immunohistochemistry vs western blot, but also to the sample used for the analysis, FFPE

and cell lines respectively. To confirm this hypothesis we detected different pattern of AKTs expression in FFPE tissues and cell lines, representative of BC molecular classes.

To our knowledge, our study is the first one that investigated the expression of AKT isoforms in a large cohort of human breast cancer samples. AKT1 and AKT2 high expression seem to be associated to worse BC patient's prognosis, in term of higher risk to develop later recurrences (AKT1) and shorter BC survival (AKT1 and AKT2), in agreement with previous results of cell biology (Chin and Toker, 2009; Toker, 2011; Wickenden and Watson, 2010). For AKT3 very few and contrasting results are present in literature. Although recently many studies have sought to understand the role of PI3K/AKT pathway in breast tumourigenesis, the specific functions and the direct substrates of each AKT isoform remain to be identified. It is, indeed, well known that different isoforms of AKT could have different functions depending on the type of cells and of tissues. The main problem to unravel the function of this regulatory system of proteins is the identification of the specific substrates for AKTs, especially AKT3. At the present time more than 200 substrates have been identified but without specific indication on the distribution among the AKTs (Toker, 2011).

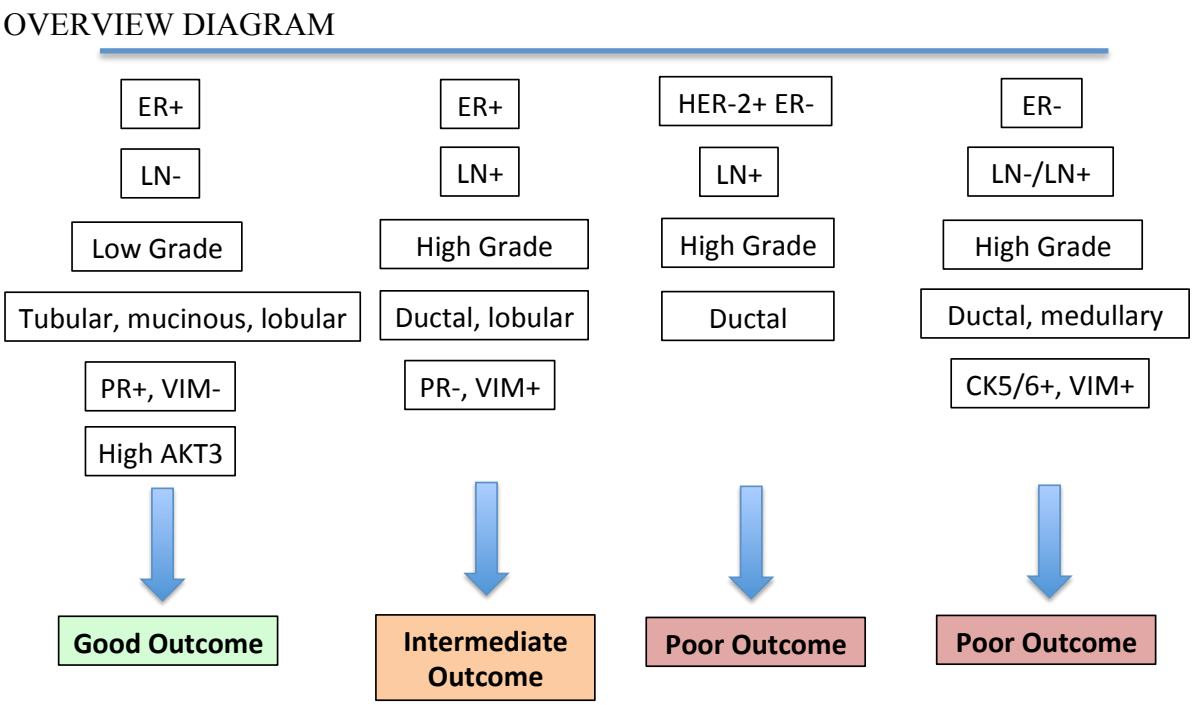


Fig. 31. Overview diagram of our suggestion for BC patients' prognosis.

6 CONCLUSIONS

Traditional morphological approaches are of undeniable importance in breast cancer classification. However, they are not sufficient to properly predict the clinical behaviour of tumours, because BC is characterized by a wide heterogeneity, also at molecular level. These limitations could be overcome by the use of molecular biomarkers in addition to pathological classification, aiming to a more personalized approach in the prognostication and therapeutic decisions.

This thesis focused on the research ad validation of new candidate prognostic molecular markers in association to the classic pathological features in a cohort of 305 young BC patients with long follow-up.

According to our findings we conclude that:

- All the clinical-pathological features analysed revealed a significant prognostic role on BC specific patient's survival: histological type of tumour, tumour grade, tumour size, lymph node involvement, number of positive nodes, and presence of recurrences ($p<0.001$).
- The lymph node involvement is the most important clinical and prognostic factor. LN- group showed a significant better clinical outcome with respect to LN+.
- The most important marker to be assessed is ER, because its expression identifies two different populations of patients with completely opposed clinical outcome: ER+ group showed a better clinical outcome compared to ER- group.
- In the ER+ group of patients, the positive expression of PR represents a favourable prognostic factor, which is associated to longer BC specific survival. Additionally, negativity to vimentin is a good prognostic marker in this group.
- We believe that luminal A and luminal B tumours belong to the same group of ER+ tumour: clinical differences of luminal A and B are the result of different tumour stage at diagnosis, with luminal B presenting the higher one.
- Despite of the better prognosis of luminal tumours, ER+ patients could present very long-term recurrences, even after 10 years. For this reason we believe that longer

monitoring program of follow-up are needed for ER+ patients, even in those cases associated to good prognosis.

- HER-2 is a worse prognostic marker associated to shorter BC survival in ER- group, not in ER+.
- ER- tumours, namely HER-2+ and Basal-like or TN, showed very poor outcome, with a BC specific death within 5 and 10 years of follow-up. However we acknowledge that HER-2+ patients were not submitted to trastuzumab, which is now recommended in cases of HER-2 amplification. Basal-like cancers represent a unique aggressive group of tumours with apparent independence from the traditional pathological features.
- CK5/6 and vimentin markers are important to confirm the basal-like subtype, but they do not seem to affect BC survival over this group of patients.
- HER-2+ and basal-like subtypes showed the higher rate of brain metastasis, in comparison to the others subtypes. These findings could be important to deepen the process of BC metastatization.
- The assessment of the CK8 expression in the metastatic regional lymph node proved to be a novel tool for prognostication of LN+ patients. Positivity to CK8 by IHC in lymph node metastatic tissue resulted a favourable indicator of longer BC survival.
- RB, PI3K, AKT2 and AKT3 were highly expressed in LN-, while CDK2, HER-2, AKT1 and CK8 in LN+.
- CDK2 and AKT1 high expression seem to have an influence on the development of later recurrences in our cohort of BC patients.
- LN- patients with high RB and AKT3 expression seem to have longer survival time.
- AKT3 expression in ER+ HER-2- and low proliferation index patients was associated to longer BC specific survival.

Although the results produced in this PhD project are encouraging and provide some new topic for BC research, further confirmations in larger retrospective and prospective studies are needed.

7 REFERENCES

(2005). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet *365*, 1687-1717.

Al-Saad, S., Donnem, T., Al-Shibli, K., Persson, M., Bremnes, R.M., and Busund, L.T. (2009). Diverse prognostic roles of Akt isoforms, PTEN and PI3K in tumor epithelial cells and stromal compartment in non-small cell lung cancer. Anticancer Res *29*, 4175-4183.

Andre, F., and Pusztai, L. (2006). Molecular classification of breast cancer: implications for selection of adjuvant chemotherapy. Nature clinical practice Oncology *3*, 621-632.

Arriagada, R., Le, M.G., Dunant, A., Tubiana, M., and Contesso, G. (2006). Twenty-five years of follow-up in patients with operable breast carcinoma: correlation between clinicopathologic factors and the risk of death in each 5-year period. Cancer *106*, 743-750.

Beckmann, M.W., Niederacher, D., Schnurch, H.G., Gusterson, B.A., and Bender, H.G. (1997). Multistep carcinogenesis of breast cancer and tumour heterogeneity. J Mol Med (Berl) *75*, 429-439.

Bonin, S., Brunetti, D., Benedetti, E., Dotti, I., Gorji, N., and Stanta, G. (2008). Molecular characterisation of breast cancer patients at high and low recurrence risk. Virchows Arch *452*, 241-250.

Bonin, S., Brunetti, D., Benedetti, E., Gorji, N., and Stanta, G. (2006). Expression of cyclin-dependent kinases and CDC25a phosphatase is related with recurrences and survival in women with peri- and post-menopausal breast cancer. Virchows Arch *448*, 539-544.

Bonin, S., and Stanta, G. (2011). RNA Extracion from Formalin-Fixed Paraffin-Embedded Tissues. In Guidelines for Molecular Analysis in Archive Tissues, G. Stanta, ed. (Berlin, Heidelberg: Springer-Verlag), pp. 57-61.

Bosco, E.E., and Knudsen, E.S. (2007). RB in breast cancer: at the crossroads of tumorigenesis and treatment. Cell Cycle *6*, 667-671.

Brotherick, I., Robson, C.N., Browell, D.A., Shenfine, J., White, M.D., Cunliffe, W.J., Shenton, B.K., Egan, M., Webb, L.A., Lunt, L.G.*, et al.* (1998). Cytokeratin expression in

breast cancer: phenotypic changes associated with disease progression. Cytometry *32*, 301-308.

Cardoso, F., Van't Veer, L., Rutgers, E., Loi, S., Mook, S., and Piccart-Gebhart, M.J. (2008). Clinical application of the 70-gene profile: the MINDACT trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *26*, 729-735.

Cavalli, L.R. (2009). Molecular markers of breast axillary lymph node metastasis. Expert review of molecular diagnostics *9*, 441-454.

Cheang, M.C., Chia, S.K., Voduc, D., Gao, D., Leung, S., Snider, J., Watson, M., Davies, S., Bernard, P.S., Parker, J.S.*, et al.* (2009). Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. Journal of the National Cancer Institute *101*, 736-750.

Chin, Y.R., and Toker, A. (2009). Function of Akt/PKB signaling to cell motility, invasion and the tumor stroma in cancer. Cell Signal *21*, 470-476.

Colombo, P.E., Milanezi, F., Weigelt, B., and Reis-Filho, J.S. (2011). Microarrays in the 2010s: the contribution of microarray-based gene expression profiling to breast cancer classification, prognostication and prediction. Breast cancer research : BCR *13*, 212.

Comanescu, M., Arsene, D., Ardeleanu, C., and Bussolati, G. (2012). The mandate for a proper preservation in histopathological tissues. Romanian journal of morphology and embryology = Revue roumaine de morphologie et embryologie *53*, 233-242.

Crabb, S.J., Cheang, M.C., Leung, S., Immonen, T., Nielsen, T.O., Huntsman, D.D., Bajdik, C.D., and Chia, S.K. (2008). Basal breast cancer molecular subtype predicts for lower incidence of axillary lymph node metastases in primary breast cancer. Clin Breast Cancer *8*, 249-256.

Davidson, T. (2001). Abortion and breast cancer: a hard decision made harder. The lancet oncology *2*, 756-758.

Dawood, S., Hu, R., Homes, M.D., Collins, L.C., Schnitt, S.J., Connolly, J., Colditz, G.A., and Tamimi, R.M. (2011). Defining breast cancer prognosis based on molecular phenotypes: results from a large cohort study. Breast cancer research and treatment *126*, 185-192.

Dawson, S.J., Rueda, O.M., Aparicio, S., and Caldas, C. (2013). A new genome-driven integrated classification of breast cancer and its implications. The EMBO journal *32*, 617-628.

Delpech, Y., Wu, Y., Hess, K.R., Hsu, L., Ayers, M., Natowicz, R., Coutant, C., Rouzier, R., Barranger, E., Hortobagyi, G.N., *et al.* (2012). Ki67 expression in the primary tumor predicts for clinical benefit and time to progression on first-line endocrine therapy in estrogen receptor-positive metastatic breast cancer. Breast cancer research and treatment *135*, 619-627.

Dietel, M., and Sers, C. (2006). Personalized medicine and development of targeted therapies: The upcoming challenge for diagnostic molecular pathology. A review. Virchows Arch *448*, 744-755.

Dotti, I., and Bonin, S. (2011). DNase Treatment of RNA. In Guidelines for Molecular Analysis in Archive Tissues, G. Stanta, ed. (Berlin, Heidelberg: Speringer-Verlag), pp. 87-90.

Downs-Holmes, C., and Silverman, P. (2011). Breast cancer: overview & updates. The Nurse practitioner *36*, 20-26; quiz 27.

Dowsett, M., Houghton, J., Iden, C., Salter, J., Farndon, J., A'Hern, R., Sainsbury, R., and Baum, M. (2006). Benefit from adjuvant tamoxifen therapy in primary breast cancer patients according oestrogen receptor, progesterone receptor, EGF receptor and HER2 status. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO *17*, 818-826.

Edge, S.B., Byrd, D.R., Compton, C.C., Fritz, A.G., Greene, F.L., and Trotti, A. (2010a). AJCC Cancer Staging Manual, seventh edn (New York: Springer).

Edge, S.B., Byrd, D.R., Compton, C.C., Fritz, A.G., Greene, F.L., and Trotti, A. (2010b). AJCC Cancer Staging Manual, Seventh Edition.

Elledge, R.M., Green, S., Pugh, R., Allred, D.C., Clark, G.M., Hill, J., Ravdin, P., Martino, S., and Osborne, C.K. (2000). Estrogen receptor (ER) and progesterone receptor (PgR), by ligand-binding assay compared with ER, PgR and pS2, by immuno-histochemistry in predicting response to tamoxifen in metastatic breast cancer: a Southwest Oncology Group Study. International journal of cancer Journal international du cancer *89*, 111-117.

Ellis, I.O., Schnitt, S.J., Sastre-Garau, X., Bussolati, G., Tavassoli, F.A., Eusebi, V., Peterse, J.L., Mukai, K., Tabar, L., Jacquemier, J., *et al.* (2003). Invasive breast carcinoma. In World Health Organization Classification of Tumours Pathology and Genetics of Tumours of the

Breast and Female Genital Organs, F.A. Tavassoli, and P. Devilee, eds. (Lyon: IARC Press), pp. 13-59.

Elston, C.W., and Ellis, I.O. (1991). Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology *19*, 403-410.

Faoro, V., and Sapino, A. (2011). Tissue Microarray (TMA). In Guidelines for Molecular Analysis in Archive Tissues, G. Stanta, ed. (Berlin, Heidelberg: Springer Verlag), pp. 23-26.

Faoro, V., and Stanta, G. (2011). Manual Microdissection. In Guidelines for Molecular Analysis in archive tissues G. Stanta, ed. (Berlin, Heidelberg: Springer-Verlag), pp. 21-22.

Ferlay, J., Shin, H.R., Bray, F., Forman, D., Mathers, C., and Parkin, D.M. (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. International journal of cancer Journal international du cancer *127*, 2893-2917.

Foulkes, W.D., Brunet, J.S., Stefansson, I.M., Straume, O., Chappuis, P.O., Begin, L.R., Hamel, N., Goffin, J.R., Wong, N., Trudel, M.*, et al.* (2004). The prognostic implication of the basal-like (cyclin E high/p27 low/p53+/glomeruloid-microvascular-proliferation+) phenotype of BRCA1-related breast cancer. Cancer Res *64*, 830-835.

Galea, M.H., Blamey, R.W., Elston, C.E., and Ellis, I.O. (1992). The Nottingham Prognostic Index in primary breast cancer. Breast cancer research and treatment *22*, 207-219.

Garcia Fernandez, A., Gimenez, N., Fraile, M., Gonzalez, S., Chabrera, C., Torras, M., Gonzalez, C., Salas, A., Barco, I., Cirera, L.*, et al.* (2012). Survival and clinicopathological characteristics of breast cancer patient according to different tumour subtypes as determined by hormone receptor and Her2 immunohistochemistry. a single institution survey spanning 1998 to 2010. Breast *21*, 366-373.

Geyer, F.C., Rodrigues, D.N., Weigelt, B., and Reis-Filho, J.S. (2012). Molecular classification of estrogen receptor-positive/luminal breast cancers. Adv Anat Pathol *19*, 39-53.

Gianni, L., Norton, L., Wolmark, N., Suter, T.M., Bonadonna, G., and Hortobagyi, G.N. (2009). Role of anthracyclines in the treatment of early breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *27*, 4798-4808.

Gnant, M., Harbeck, N., and Thomssen, C. (2011). St. Gallen 2011: Summary of the Consensus Discussion. Breast Care (Basel) *6*, 136-141.

Goldhirsch, A., Ingle, J.N., Gelber, R.D., Coates, A.S., Thurlimann, B., and Senn, H.J. (2009). Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2009. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO *20*, 1319-1329.

Goldhirsch, A., Wood, W.C., Coates, A.S., Gelber, R.D., Thurlimann, B., and Senn, H.J. (2011). Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO *22*, 1736-1747.

Grabinski, N., Bartkowiak, K., Grupp, K., Brandt, B., Pantel, K., and Jucker, M. (2011). Distinct functional roles of Akt isoforms for proliferation, survival, migration and EGF-mediated signalling in lung cancer derived disseminated tumor cells. Cell Signal *23*, 1952-1960.

Gruver, A.M., Portier, B.P., and Tubbs, R.R. (2011). Molecular pathology of breast cancer: the journey from traditional practice toward embracing the complexity of a molecular classification. Arch Pathol Lab Med *135*, 544-557.

Guiu, S., Michiels, S., Andre, F., Cortes, J., Denkert, C., Di Leo, A., Hennessy, B.T., Sorlie, T., Sotiriou, C., Turner, N.*, et al.* (2012). Molecular subclasses of breast cancer: how do we define them? The IMPAKT 2012 Working Group Statement. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO *23*, 2997-3006.

Hammond, M.E., Hayes, D.F., Dowsett, M., Allred, D.C., Hagerty, K.L., Badve, S., Fitzgibbons, P.L., Francis, G., Goldstein, N.S., Hayes, M.*, et al.* (2010). American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *28*, 2784-2795.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57-70.

Hayes, D.F. (2005). Prognostic and predictive factors revisited. Breast *14*, 493-499.

http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-treating-general-info.

http://www.lgcstandards-atcc.org/ATCCCulturesandProducts/CellBiology/CellLinesandHybridomas/tabid/981/Default.aspx.

http://www.nccn.org/professionals/physician_gls/f_guidelines.asp.

http://www.registri-tumori.it/cms/?q=sede_mammella.

Iyer, S.V., Dange, P.P., Alam, H., Sawant, S.S., Ingle, A.D., Borges, A.M., Shirsat, N.V., Dalal, S.N., and Vaidya, M.M. (2013). Understanding the role of keratins 8 and 18 in neoplastic potential of breast cancer derived cell lines. PloS one *8*, e53532.

Jares, P., Rey, M.J., Fernandez, P.L., Campo, E., Nadal, A., Munoz, M., Mallofre, C., Muntane, J., Nayach, I., Estape, J.*, et al.* (1997). Cyclin D1 and retinoblastoma gene expression in human breast carcinoma: correlation with tumour proliferation and oestrogen receptor status. J Pathol *182*, 160-166.

Kalinsky, K., Jacks, L.M., Heguy, A., Patil, S., Drobnjak, M., Bhanot, U.K., Hedvat, C.V., Traina, T.A., Solit, D., Gerald, W.*, et al.* (2009). PIK3CA mutation associates with improved outcome in breast cancer. Clinical cancer research : an official journal of the American Association for Cancer Research *15*, 5049-5059.

Kim, S.J., Nakayama, S., Miyoshi, Y., Taguchi, T., Tamaki, Y., Matsushima, T., Torikoshi, Y., Tanaka, S., Yoshida, T., Ishihara, H.*, et al.* (2008). Determination of the specific activity of CDK1 and CDK2 as a novel prognostic indicator for early breast cancer. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO *19*, 68-72.

Kollias, J., Elston, C.W., Ellis, I.O., Robertson, J.F., and Blamey, R.W. (1997). Early-onset breast cancer--histopathological and prognostic considerations. British journal of cancer *75*, 1318-1323.

Kroman, N., Jensen, M.B., Wohlfahrt, J., Mouridsen, H.T., Andersen, P.K., and Melbye, M. (2000). Factors influencing the effect of age on prognosis in breast cancer: population based study. BMJ *320*, 474-478.

Kumle, M. (2008). Declining breast cancer incidence and decreased HRT use. Lancet *372*, 608-610.

Lai, D., Visser-Grieve, S., and Yang, X. (2012). Tumour suppressor genes in chemotherapeutic drug response. Bioscience reports *32*, 361-374.

Lester, S.C. (2010). La mammella. In Robbins e Cotran Le basi patologiche delle Malattie, Elsevier, ed. (Kumar, V.

Abbas, A.K.

Fausto, N.

Aster, J.C.), pp. 1119-1154.

Li, X.R., Liu, M., Zhang, Y.J., Wang, J.D., Zheng, Y.Q., Li, J., Ma, B., and Song, X. (2011). Evaluation of ER, PgR, HER-2, Ki-67, cyclin D1, and nm23-H1 as predictors of pathological complete response to neoadjuvant chemotherapy for locally advanced breast cancer. Med Oncol *28 Suppl 1*, S31-38.

Livasy, C.A., Karaca, G., Nanda, R., Tretiakova, M.S., Olopade, O.I., Moore, D.T., and Perou, C.M. (2006). Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. Mod Pathol *19*, 264-271.

Manning, A.L., and Dyson, N.J. (2012). RB: mitotic implications of a tumour suppressor. Nature reviews Cancer *12*, 220-226.

Melis, M. (1975). Compendio di tecnica istopatologica (Roma: Lombardo Editore).

Miller, T.W., Rexer, B.N., Garrett, J.T., and Arteaga, C.L. (2011). Mutations in the phosphatidylinositol 3-kinase pathway: role in tumor progression and therapeutic implications in breast cancer. Breast cancer research : BCR *13*, 224.

Moll, R., Franke, W.W., Schiller, D.L., Geiger, B., and Krepler, R. (1982). The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. Cell *31*, 11-24.

Mook, S., Schmidt, M.K., Rutgers, E.J., van de Velde, A.O., Visser, O., Rutgers, S.M., Armstrong, N., van't Veer, L.J., and Ravdin, P.M. (2009). Calibration and discriminatory accuracy of prognosis calculation for breast cancer with the online Adjuvant! program: a hospital-based retrospective cohort study. The lancet oncology *10*, 1070-1076.

Muller, B.M., Kronenwett, R., Hennig, G., Euting, H., Weber, K., Bohmann, K., Weichert, W., Altmann, G., Roth, C., Winzer, K.J.*, et al.* (2011). Quantitative determination of estrogen receptor, progesterone receptor, and HER2 mRNA in formalin-fixed paraffin-

embedded tissue--a new option for predictive biomarker assessment in breast cancer. Diagnostic molecular pathology : the American journal of surgical pathology, part B *20*, 1-10.

Musgrove, E.A., and Sutherland, R.L. (2010). RB in breast cancer: differential effects in estrogen receptor-positive and estrogen receptor-negative disease. Cell Cycle *9*, 4607.

Muss, H.B., Thor, A.D., Berry, D.A., Kute, T., Liu, E.T., Koerner, F., Cirrincione, C.T., Budman, D.R., Wood, W.C., Barcos, M.*, et al.* (1994). c-erbB-2 expression and response to adjuvant therapy in women with node-positive early breast cancer. The New England journal of medicine *330*, 1260-1266.

Nardon, E., Donada, M., Bonin, S., Dotti, I., and Stanta, G. (2009). Higher random oligo concentration improves reverse transcription yield of cDNA from bioptic tissues and quantitative RT-PCR reliability. Exp Mol Pathol *87*, 146-151.

Nielsen, T.O., Hsu, F.D., Jensen, K., Cheang, M., Karaca, G., Hu, Z., Hernandez-Boussard, T., Livasy, C., Cowan, D., Dressler, L.*, et al.* (2004). Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clinical cancer research : an official journal of the American Association for Cancer Research *10*, 5367-5374.

Paez, D., Labonte, M.J., Bohanes, P., Zhang, W., Benhanim, L., Ning, Y., Wakatsuki, T., Loupakis, F., and Lenz, H.J. (2012). Cancer dormancy: a model of early dissemination and late cancer recurrence. Clin Cancer Res *18*, 645-653.

Park, D., Karesen, R., Noren, T., and Sauer, T. (2007). Ki-67 expression in primary breast carcinomas and their axillary lymph node metastases: clinical implications. Virchows Arch *451*, 11-18.

Park, S., Koo, J.S., Kim, M.S., Park, H.S., Lee, J.S., Lee, J.S., Kim, S.I., and Park, B.W. (2012). Characteristics and outcomes according to molecular subtypes of breast cancer as classified by a panel of four biomarkers using immunohistochemistry. Breast *21*, 50-57.

Parker, J.S., Mullins, M., Cheang, M.C., Leung, S., Voduc, D., Vickery, T., Davies, S., Fauron, C., He, X., Hu, Z.*, et al.* (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *27*, 1160-1167.

Patani, N.R., Dwek, M.V., and Douek, M. (2007). Predictors of axillary lymph node metastasis in breast cancer: a systematic review. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology *33*, 409-419.

Pegram, M.D., Konecny, G., and Slamon, D.J. (2000). The molecular and cellular biology of HER2/neu gene amplification/overexpression and the clinical development of herceptin (trastuzumab) therapy for breast cancer. Cancer treatment and research *103*, 57-75.

Perou, C.M., Sorlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A.*, et al.* (2000). Molecular portraits of human breast tumours. Nature *406*, 747-752.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res *29*, e45.

Pfeffer, U., Romeo, F., Noonan, D.M., and Albini, A. (2009). Prediction of breast cancer metastasis by genomic profiling: where do we stand? Clinical & experimental metastasis *26*, 547-558.

Porter, P. (2008). "Westernizing" women's risks? Breast cancer in lower-income countries. The New England journal of medicine *358*, 213-216.

Rakha, E.A., El-Sayed, M.E., Lee, A.H., Elston, C.W., Grainge, M.J., Hodi, Z., Blamey, R.W., and Ellis, I.O. (2008a). Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *26*, 3153-3158.

Rakha, E.A., and Ellis, I.O. (2011). Modern classification of breast cancer: should we stick with morphology or convert to molecular profile characteristics. Adv Anat Pathol *18*, 255-267.

Rakha, E.A., Reis-Filho, J.S., and Ellis, I.O. (2008b). Basal-like breast cancer: a critical review. J Clin Oncol *26*, 2568-2581.

Rakha, E.A., Reis-Filho, J.S., and Ellis, I.O. (2010). Combinatorial biomarker expression in breast cancer. Breast cancer research and treatment *120*, 293-308.

Richie, R.C., and Swanson, J.O. (2003). Breast cancer: a review of the literature. J Insur Med *35*, 85-101.

Romond, E.H., Perez, E.A., Bryant, J., Suman, V.J., Geyer, C.E., Jr., Davidson, N.E., Tan-Chiu, E., Martino, S., Paik, S., Kaufman, P.A*., et al.* (2005). Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. The New England journal of medicine *353*, 1673-1684.

Rossouw, J.E., Anderson, G.L., Prentice, R.L., LaCroix, A.Z., Kooperberg, C., Stefanick, M.L., Jackson, R.D., Beresford, S.A., Howard, B.V., Johnson, K.C*., et al.* (2002). Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. JAMA *288*, 321-333.

Sarode, V.R., Han, J.S., Morris, D.H., Peng, Y., and Rao, R. (2011). A Comparative Analysis of Biomarker Expression and Molecular Subtypes of Pure Ductal Carcinoma In Situ and Invasive Breast Carcinoma by Image Analysis: Relationship of the Subtypes with Histologic Grade, Ki67, p53 Overexpression, and DNA Ploidy. International journal of breast cancer *2011*, 217060.

Satelli, A., and Li, S. (2011). Vimentin in cancer and its potential as a molecular target for cancer therapy. Cellular and molecular life sciences : CMLS *68*, 3033-3046.

Schnitt, S.J. (2003). Benign breast disease and breast cancer risk: morphology and beyond. Am J Surg Pathol *27*, 836-841.

Shao, M.M., Chan, S.K., Yu, A.M., Lam, C.C., Tsang, J.Y., Lui, P.C., Law, B.K., Tan, P.H., and Tse, G.M. (2012). Keratin expression in breast cancers. Virchows Archiv : an international journal of pathology *461*, 313-322.

Shapiro, G.I. (2006). Cyclin-dependent kinase pathways as targets for cancer treatment. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *24*, 1770-1783.

Slamon, D., and Pegram, M. (2001). Rationale for trastuzumab (Herceptin) in adjuvant breast cancer trials. Seminars in oncology *28*, 13-19.

Slamon, D.J., Clark, G.M., Wong, S.G., Levin, W.J., Ullrich, A., and McGuire, W.L. (1987). Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science *235*, 177-182.

Slamon, D.J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., Fleming, T., Eiermann, W., Wolter, J., Pegram, M*., et al.* (2001). Use of chemotherapy plus a

monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. The New England journal of medicine *344*, 783-792.

Soerjomataram, I., Louwman, M.W., Ribot, J.G., Roukema, J.A., and Coebergh, J.W. (2008). An overview of prognostic factors for long-term survivors of breast cancer. Breast cancer research and treatment *107*, 309-330.

Sorlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S.*, et al.* (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A *98*, 10869-10874.

Sotiriou, C., and Pusztai, L. (2009). Gene-expression signatures in breast cancer. The New England journal of medicine *360*, 790-800.

Soule, H.D., Vazguez, J., Long, A., Albert, S., and Brennan, M. (1973). A human cell line from a pleural effusion derived from a breast carcinoma. Journal of the National Cancer Institute *51*, 1409-1416.

Sparano, J.A., and Paik, S. (2008). Development of the 21-gene assay and its application in clinical practice and clinical trials. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *26*, 721-728.

Stanta, G. (2011). Guidelines for Molecular Analysis in Archive Tissues, First edn (Berlin Heidelberg: Springer-Verlag GmbH).

Stanta, G., Bonin, S., and Perin, R. (1998). RNA extraction from formalin-fixed and paraffin-embedded tissues. Methods Mol Biol *86*, 23-26.

Stuart-Harris, R., Caldas, C., Pinder, S.E., and Pharoah, P. (2008). Proliferation markers and survival in early breast cancer: a systematic review and meta-analysis of 85 studies in 32,825 patients. Breast *17*, 323-334.

Tang, P., Skinner, K.A., and Hicks, D.G. (2009). Molecular classification of breast carcinomas by immunohistochemical analysis: are we ready? Diagn Mol Pathol *18*, 125-132.

Tawfik, K., Kimler, B.F., Davis, M.K., Fan, F., and Tawfik, O. (2013). Ki-67 expression in axillary lymph node metastases in breast cancer is prognostically significant. Human pathology *44*, 39-46.

Toker, A. (2011). Achieving specificity in Akt signaling in cancer. Advances in enzyme regulation.

Truong, P.T., Berthelet, E., Lee, J., Kader, H.A., and Olivotto, I.A. (2005). The prognostic significance of the percentage of positive/dissected axillary lymph nodes in breast cancer recurrence and survival in patients with one to three positive axillary lymph nodes. Cancer *103*, 2006-2014.

Valentin, M.D., da Silva, S.D., Privat, M., Alaoui-Jamali, M., and Bignon, Y.J. (2012). Molecular insights on basal-like breast cancer. Breast cancer research and treatment *134*, 21-30.

van 't Veer, L.J., Dai, H., van de Vijver, M.J., He, Y.D., Hart, A.A., Mao, M., Peterse, H.L., van der Kooy, K., Marton, M.J., Witteveen, A.T.*, et al.* (2002). Gene expression profiling predicts clinical outcome of breast cancer. Nature *415*, 530-536.

van de Rijn, M., Perou, C.M., Tibshirani, R., Haas, P., Kallioniemi, O., Kononen, J., Torhorst, J., Sauter, G., Zuber, M., Kochli, O.R.*, et al.* (2002). Expression of cytokeratins 17 and 5 identifies a group of breast carcinomas with poor clinical outcome. Am J Pathol *161*, 1991-1996.

van der Groep, P., van der Wall, E., and van Diest, P.J. (2011). Pathology of hereditary breast cancer. Cell Oncol (Dordr) *34*, 71-88.

Venkitaraman, A.R. (2002). Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell *108*, 171-182.

Veronesi, U., Cascinelli, N., Mariani, L., Greco, M., Saccozzi, R., Luini, A., Aguilar, M., and Marubini, E. (2002). Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer. The New England journal of medicine *347*, 1227-1232.

Veronesi, U., Paganelli, G., Viale, G., Luini, A., Zurrida, S., Galimberti, V., Intra, M., Veronesi, P., Robertson, C., Maisonneuve, P.*, et al.* (2003). A randomized comparison of sentinel-node biopsy with routine axillary dissection in breast cancer. The New England journal of medicine *349*, 546-553.

Vetto, J.T., Luoh, S.W., and Naik, A. (2009). Breast cancer in premenopausal women. Current problems in surgery *46*, 944-1004.

Weigel, M.T., and Dowsett, M. (2010). Current and emerging biomarkers in breast cancer: prognosis and prediction. Endocrine-related cancer *17*, R245-262.

Weigelt, B., Baehner, F.L., and Reis-Filho, J.S. (2010a). The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. J Pathol *220*, 263-280.

Weigelt, B., Geyer, F.C., and Reis-Filho, J.S. (2010b). Histological types of breast cancer: how special are they? Mol Oncol *4*, 192-208.

Wellings, S.R., and Jensen, H.M. (1973). On the origin and progression of ductal carcinoma in the human breast. Journal of the National Cancer Institute *50*, 1111-1118.

Wickenden, J.A., and Watson, C.J. (2010). Key signalling nodes in mammary gland development and cancer. Signalling downstream of PI3 kinase in mammary epithelium: a play in 3 Akts. Breast cancer research : BCR *12*, 202.

Wiechmann, L., Sampson, M., Stempel, M., Jacks, L.M., Patil, S.M., King, T., and Morrow, M. (2009). Presenting features of breast cancer differ by molecular subtype. Ann Surg Oncol *16*, 2705-2710.

Wolff, A.C., Hammond, M.E., Schwartz, J.N., Hagerty, K.L., Allred, D.C., Cote, R.J., Dowsett, M., Fitzgibbons, P.L., Hanna, W.M., Langer, A.*, et al.* (2007). American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *25*, 118-145.

Yang, Y., Hao, J., Liu, X., Dalkin, B., and Nagle, R.B. (1997). Differential expression of cytokeratin mRNA and protein in normal prostate, prostatic intraepithelial neoplasia, and invasive carcinoma. The American journal of pathology *150*, 693-704.

Youlden, D.R., Cramb, S.M., Dunn, N.A., Muller, J.M., Pyke, C.M., and Baade, P.D. (2012). The descriptive epidemiology of female breast cancer: an international comparison of screening, incidence, survival and mortality. Cancer Epidemiol *36*, 237-248.

Zhou, W., He, Z., Xue, J., Wang, M., Zha, X., Ling, L., Chen, L., Wang, S., and Liu, X. (2012). Molecular subtype classification is a determinant of non-sentinel lymph node metastasis in breast cancer patients with positive sentinel lymph nodes. PloS one *7*, e35881.

8 LIST OF PUBLICATIONS

PUBLICATIONS

- Isabella Dotti, Ermanno Nardon, Danae Pracella, and Serena Bonin. Quantitative Real-Time RT-PCR. In G. Stanta Ed. Guidelines for Molecular Analysis in Archive Tissues. Heidelberg, Berlin: Springer-Verlag GmbH, 2011: 121-132.
- Pracella D, Bonin S, Barbazza R, Sapino A and Stanta G. Are Breast Cancer Molecular Classes Predictive of Survival in Patients With Long Follow-up? 2013, submitted.

POSTERS AND PRESENTATIONS

- Danae Pracella, Serena Bonin and Giorgio Stanta. *Other Biomarkers in Breast Cancer Molecular Characterization*. Oral presentation at 25th International Meeting Adriatic Society of Pathology, Duino 2011.
- Danae Pracella, Bonin Serena and Giorgio Stanta. *Breast Cancer: Molecular Classification and New Biomarkers Assessment In FFPE Tissues*. Poster session at 14th Breast cancer Conference, Milan 2012.
- Giorgio Stanta, Renzo Barbazza, Enrico Roggero, Federica Marchesin, Danae Pracella and Serena Bonin. *HPV E7 monoclonal antibody may help to define better prognosis of prostate cancer.* Oral presentation at Global Congress on Prostate Cancer, Bruxelles 2012.
- Danae Pracella, Bonin Serena and Giorgio Stanta. *New Biomarker For Breast Cancer Molecular Subtyping.* Poster session at 2nd Regional Summer School of Biomedicine, Udine 2012.
- S. Bonin, D. Pracella, R. Barbazza, F. Marchesin, C. Puiatti, G. Stanta. *Detection of HPV E7 oncoviral protein in prostate cancer tissues.* Oral presentation at 54th Symposium of the Society for Histochemistry, Vienna 2012.

9 ACKNOWLEDGMENTS

I would like to thank my supervisor Prof. Giorgio Stanta for giving me the opportunity to work in his laboratory, for allowing me discover this fascinating field of research and for his constant guidance.

My sincere gratitude goes to Dr. Serena Bonin for teaching me and giving me precious advice during all the phases of the research.

I wish to thank Dr. Renzo Barbazza, Dr Sandro Sulfaro and Dr. Maria Malagoli for their knowledge and for providing thorough clinical explanations.

I wish to thank my external and internal advisors Prof. Ambrogio Fassina and Prof. Giuliana Decorti.

I thank my lab partners Ermanno, Cinzia, Marisa, Valentina, Federica, Melita for sharing this important experience with me, for their support, the stimulating discussion and, most important, their friendship.

Finally I would thank, and I will never stop doing, my parents and my love Simone for their continuous support, their advice and the immense patience. To them I want to express all my affection and love.