Investigating Male Breast Cancer Using Transcriptomics and Immunohistochemistry

Sreekumar Sundara Rajan

Submitted in accordance with the requirements for the degree of

Doctor of Medicine

University of Leeds

School of Medicine

May 2016

Intellectual property and publication statement

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The findings of my research on External validation of ImmunoRatio™ image analysis application was published in the following article:

Sundara Rajan S, Horgan K, Speirs V, Hanby AM. External validation of the ImmunoRatio image analysis application for ERα determination in breast cancer. J Clin Pathol. 2014; 67(1):72-5.

This work forms the ERα part of the methodology, results and discussions sections of Chapter 3 of this thesis. The experiments and immunohistochemistry scoring in male breast cancer samples was performed solely by me. The scientific aspects of the study were reviewed by Prof. V Speirs, Prof. A M Hanby and Mr K Horgan.

Acknowledgments

I would like to first and foremost express my gratitude to Professor Valerie Speirs for her guidance and patience during a rather prolonged part-time research due to personal reasons. I would like to thank my academic supervisor, Professor Andrew M Hanby, who has been always supportive and for all our invigorating discussions.

I am obliged to my clinical supervisor, Mr Kieran Horgan, who has been instrumental in guiding me through my difficult four years during research. Indubitably, without the constant support of Professor V Speirs and Mr Horgan, completion of this project would have been impossible

I would like to acknowledge my mentors who have helped me during various stages of my research. First of all, I would like to thank Mrs Jennifer Pollard, Dr Philomena Esteves, Mrs Barbara Ozloz, Mr Steven Pollock and Mr Mike Shires for providing technical assistance. I am thankful to Dr Helene Thygesen for providing statistical advice. I am grateful to Dr Alastair Droop for the invaluable input while assessing the bioinformatics data, without which it would have been impossible to analyse the gene expression profiling results.

A special thanks to Dr Eldo Verghese, for being a great friend and brainstorming partner during my research years.

On a personal note, I would like to thank my wife, Rashmi, for all her love and support and patience in looking after my lovely twins, Arnav and Pranav, while I was away writing the thesis. I would like to thank my mother and sister for their endurance for tolerating me as a son/little brother.

Finally, I would like to dedicate this work to my late wife, Roopa Chandran and to my father, Late Dr M Sundara Rajan, whom I am forever indebted for teaching me the true essence of life.

Abstract

Background: The rare nature of male breast cancer (MBC) has led to its management being guided by the extensive research conducted in the field of female breast cancer (FBC). The aim of this study was to evaluate MBC at both protein and molecular level to improve understanding of its pathology.

Methodology: Immunohistochemistry analysis was performed in MBC (n=428) TMAs for 18 biomarkers (ERα, ERβ1, ERβ2, ERβ5, Total PR, AR, CK5/6, CK14, CK18, CK19, p53, Bcl-2, Her2, E-cadherin, Ki67, Survivin, Prolactin and FOXA1). The manual scoring of ERα and Ki67 was correlated with a fully automated immunohistochemistry image analysis system (ImmunoRatio™). Finally gene expression profiling (GEP) was undertaken in matched MBC (n=15) and FBC (n=10) samples.

Results: There was poor 5 year overall survival (OS) in CK18 and CK19 negative patients (p= 0.05; p= 0.003), as well as poor 10 year OS in CK19 negative patients (p= 0.002). Age (p= 0.001) and nodal status (p= 0.04) was found to be independent predictors of OS at 5 years.

There was significant correlations between manual and ImmunoRatio™ ERα (ρ= 0.872; p= 0.000) and Ki67 (r= 0.675; p= 0.000) scores. However due to a low measure of agreement it was not possible to validate Ki67 scoring using ImmunoRatio™.

The functional enrichment analysis of GEP data using less stringent criteria ($p < 0.05$) identified 735 differentially expressed genes. The data analysis showed up-regulation of genes involved in ECM synthesis, degradation and re-modelling in MBC. The end product of one of the up-regulated genes (Fibronectin (FN1)) was validated in the MBC cohort with high fibronectin expression (60%) being positively associated with nodal status and showed a trend towards poor 5 year OS (p= 0.06).

Conclusion: In MBC, epithelial cytokeratins, especially CK19 was found to be of prognostic significance. The extracellular matrix remodelling associated genes were found to be up-regulated in MBC. Fibronectin, end product of one of the up-regulated gene was found to have prognostic significance in MBC.

Table of Contents

List of Abbreviations

Abbreviation	Description
aCGH	Array Comparative Genomic Hybridisation
AI	Aromatase Inhibitors
AMH	Prof. Andrew Hanby
ANC	Axillary Node Clearance
anti-FN1	anti-Fibronectin 1 antibody
AR	Androgen Receptor
Bcl-2	B-cell lymphoma 2
BCSS	Breast Cancer Specific Survival
CB	Core Biopsy
CBD	Cellular Binding Domain
CGH	Comparative Genomic Hybridisation
CI	Confidence Interval
CK	Cytokeratin
DAB	3, 3'- diaminobenzidine
DCIS	Ductal Carcinoma in Situ
DFS	Disease Free Survival
ECM	Extracellular Matrix
EDA	Extra Domain A
EDB	Extra Domain B
EDTA	Ethylene Di-amine Tetra-acetic Acid
EGFR	Epidermal Growth Factor Receptor
EIF4	Eukaryotic Translation Initiation Factor 4
EMF	Electromagnetic Fields
EMT	Epithelial Mesenchymal Transition
ER	Oestrogen Receptor
ERRB2	erb-b2 receptor tyrosine kinase 2
ERβ1	Oestrogen Receptor β1
ERβ2	Oestrogen Receptor β2
ERβ5	Oestrogen Receptor β5
FBC	Female Breast Cancer
FDR	False Discovery Rate
FFPE	Formalin Fixed Paraffin Embedded
FISH	Fluorescent In Situ Hybridisation
FN	Fibronectin
FNAC	Fine Needle Aspiration Cytology
FOXA1	Forkhead Box Protein A1
GEO	Gene Expression Omnibus
GEP	Gene Expression Profiling
GO	Gene Ontology

H&E	Haematoxylin and Eosin
H2O2	Hydrogen Peroxide
Her2	Human Epidermal Growth Factor receptor
HRP	Horseradish Peroxidase
IAP	Inhibitor of Apoptosis family
IHC	Immunohistochemistry
III CS	Type III homology Connecting Segment
IPA	Ingenuity Pathway Analysis
IQR	Interquartile Range
LDS	Lithium Dodecyl Sulphate
LI	Labelling Index
LTHT	Leeds Teaching Hospitals NHS Trust
LVI	Lympho-vascular Invasion
MBC	Male Breast Cancer
MFS	Metastasis Free Survival
MOPS	3-(N-Morpholino) Propane Sulphuric Acid
NCBI	National Centre for Biotechnology Information
OR	Odds Ratio
OS	Overall Survival
PBS	Phosphate Buffered Saline
PBST	Phosphate Buffered Saline with Tween
PGR	Gene encoding for progesterone receptor
PR	Progesterone Receptor
PRL	Prolactin
PRLR	Prolactin receptor
PVDF	Hybond-P Polyvinylidene Difluoride
RMA	Robust Multichip Algorithm
SAB	Streptavidin-Biotin
SDS	Sodium Dodecyl Sulphate
SEER	The Surveillance, Epidemiology and End Result
SLNB	Sentinel Lymph Node Biopsy
SPSS	Statistical Package for Social Sciences
Stat3	Signal Transducer and Activator of Transcription 3
TBS	Tris Buffered Saline
TBST	Tris Buffered Saline Tween
TDLU	Terminal Ductal Lobular Units
TGFβ1	Transforming Growth Factor β1
TMA	Tissue Micro Array
Tris-HCl	Tris-Hydrochloric acid
UK	United Kingdom
USA	United States of America

List of Figures

List of Tables

CHAPTER 1

Male Breast Cancer Overview

1.1 INTRODUCTION

1.1.1 Male Breast Cancer

Male breast cancer (MBC) is uncommon and accounts for approximately 1% of the breast cancers diagnosed (Callari et al., 2011, Shaaban et al., 2012, White et al., 2011). MBC accounts for only 0.1% of male cancer deaths (Jemal et al., 2008) but its annual incidence continues to increase by 1.1% (Giordano et al., 2002a). However, the improvements in overall survival seen over the years in female breast cancer (FBC) have not been observed in MBC primarily due to the poor understanding of the pathology due to the lack of large prospective trials. Hence, the management of MBC is currently based on the information extrapolated from the immense research undertaken in the field of FBC.

1.1.2 Epidemiology

The worldwide distribution of MBC incidence resembles that of FBC with the highest incidence seen in North America and Europe and lower rates in Asia (Ravandi-Kashani and Hayes, 1998). The incidence of MBC is increasing but at a slower rate than FBC (Ewertz, 1996, Shavers et al., 2003). The National Cancer Institute data from the United States of America (USA) shows that the incidence of MBC has risen to 1.08 from 0.86 per 100,000 men (Giordano et al., 2004). The incidence in the United Kingdom (UK) is similar to that in the USA with just under 1% (n=334) diagnosed with MBC in 2012 (CRUK, 2012). The incidence of MBC is considerably higher (15%) in sub-Saharan Africa, which has been mainly attributed to the hyper-oestrogenism from liver damage due to the prevalence of endemic infectious diseases like Schistosomiasis (Andersen

and Gram, 1982, Carlsson et al., 1981). The prevalence of MBC increases with age but unlike FBC it has a unimodal distribution with the peak seen at late sixties and early seventies (Anderson et al., 2004, Giordano et al., 2004). Hence the average age at diagnosis for MBC is at least 10 years older than that for FBC (Devesa et al., 1995, Hill et al., 2005). Moreover, there have not been any screening programmes for MBC which can distort the pattern of age at diagnosis as seen in FBC.

The racial predilection of MBC is unique and different compared to that of FBC. There is a relative increase in the incidence of MBC compared to FBC amongst people of African-American origin (Crew et al., 2007, Nahleh et al., 2007). This is in sharp contrast to the consistently higher incidence of FBC seen amongst post-menopausal Caucasian women compared to African-American women (Anderson et al., 2004). A higher risk for developing MBC has been shown amongst Jews, with younger age at presentation for the Sephardic Jews and an increased life time risk for the Ashkenazi Jews (Brenner et al., 2002, Steinitz et al., 1981).

1.1.3 Aetiology and Risk Factors

The unimodal age-frequency distribution of breast cancer incidence suggests a hormone independent epithelial carcinogenesis in MBC (Pike et al., 1983). However, hormonal imbalance due to liver disease, Klinefelter's syndrome and obesity have been implicated in the causation of MBC. It has been established through previous published studies that, hormonal imbalances through change in the oestrogen-testosterone ratio can predispose to MBC (Ballerini et al.,

1990, Casagrande et al., 1988). An elevated oestradiol level was observed in MBC patients compared to controls in the published literature (Sasco et al., 1993, Brinton et al., 2015). Moreover, the recent evidence from the Male Breast Cancer Pooling Project reveals obesity increases the risk of MBC (Brinton et al., 2014, Humphries et al., 2015). The elevated circulating level of oestrogen in obese men is thought to be due to the aromatization of androgens with conversion of testosterone to oestradiol and androstenedione to oestrone in peripheral adipose tissue (Hsing et al., 1998). Hormonal imbalance can be worsened due to a decrease in the sex hormone binding globulin, leading to a greater amount of bio-available oestrogen in obese men (Casagrande et al., 1988).

The hormonal imbalance seen in patients with gynaecomastia and Klinefelter's syndrome has drawn the attention of researchers to explore whether there is a causative relationship to MBC. The most recent evidence from the Male Breast Cancer Pooling Project has shown increased risk for the development of MBC associated with both gynaecomastia and Klinefelter's syndrome (Brinton et al., 2014). Elevated levels of oestrogen due to chronic liver disease (cirrhosis), exogenous administration of oestrogen (transsexuals) or iatrogenic oestrogen therapy (for prostate cancer) have been implicated as risk factors for MBC (Contractor et al., 2008). Androgen deficiency due to underlying testicular conditions such as mumps orchitis, undescended testes and testicular injury has also been linked to MBC (Mabuchi et al., 1985, Thomas et al., 1992). However the strongest proven relationship is between MBC and Klinefelter's syndrome due to testosterone deficiency (Hultborn et al., 1997).

Other risk factors for MBC are environmental/occupational factors such as exposure to ionizing radiation and chronic heat (Brinton et al., 2008, Ottini et al., 2009). It has been proposed that, the environmental and occupational factors either directly or indirectly lead to testicular damage resulting in the causation of MBC (Rosenbaum et al., 1994, Stenlund and Floderus, 1997). The individual studies conducted to evaluate the association between electromagnetic fields (EMF) and MBC had inherent biases with type II error, selection bias (miss-classification of the EMF exposure) and recall bias (Weiss et al., 2005). However, a subsequent meta-analysis showed that the exposure of EMF increases the risk of breast cancer by 1.37 times (CI, 1.11-1.71) in men (Erren, 2001). Unlike FBC, there is little reliable evidence to conclude any association between dietary intake and MBC (Weiss et al., 2005).

Category	Known risk factors	Inconclusive evidence
Genetic factors	BRCA2 Family history –Ashkenazi Jews Klinefelter's syndrome	AR gene CHECK2 PTEN
Hormonal imbalance	Obesity Gynaecomastia Cirrhosis Testicular disorders Cryptorchidism Mumps orchitis Orchidectomy	Prostate cancer
Environmental	Ionizing radiation	EMFs Occupational exposure High temperature Diet and alcohol
Iatrogenic	Exogenous oestrogen Radiotherapy Fluoroscopy	Prostate cancer treatment

Table 1: Risk factors for Male Breast Cancer

1.1.4 Genetics

A positive family history has been associated with an increased risk for MBC but unlike FBC the risk cannot be quantified primarily due to the rarity of the disease resulting in a lack of information to calculate risk estimates. However, it has been estimated that a positive family history of male or female breast cancer among first degree relatives increases MBC risk by 2 or 3 folds (Weiss et al., 2005). Similarly, the diagnosis of MBC in a first degree relative is a substantial risk factor for development of future breast cancer in women. It is well established that as with other solid tumours, the development of MBC is a multi-step process and genomic DNA alteration playing an important role (Tommasi et al., 2010). The DNA alteration by amplification or deletions leads to changes in the gene expression, which alter key cellular processes downstream for the development of breast cancer.

In comparison to FBC, where there is a strong inherited susceptibility (30-86%) for rare high penetrance mutations (Ford et al., 1998, Tischkowitz et al., 2002), such mutations are observed only in 4 to 40% of MBC (Friedman et al., 1997, Haraldsson et al., 1998, Thorlacius et al., 1997). BRCA 2 gene mutation is frequent and accounts for 4 to 40% of patients with MBC (Callari et al., 2011). In comparison to FBC, BRCA2 mutations confer 6 to 9% life time relative risk of developing breast cancer in men in the Western world (Evans et al., 2010). In the Cambridge series, 8% of the patients with MBC were found to have BRCA2 mutations in comparison to 40% in the Icelandic series, which is the largest incidence reported so far in the literature (Basham et al., 2002, Thorlacius et al., 1996). The high incidence seen in the Icelandic series has been attributed to

the founder mutation of 999del5 in the BRCA2 gene (Thorlacius et al., 1996). The wide variation seen amongst various studies may suggest broad genetic variation across the population. However caution needs to be exercised while interpreting these results due to the small sample size. There is evidence to suggest 50-92% of familial breast cancer arises from BRCAX families, however the underlying mechanism remains uncertain (Besic et al., 2008, Ahn et al., 2004)

Klinefelter's syndrome is a rare chromosomal abnormality of 47 XXY karyotype and occurs in 1:1000 men (Evans and Crichlow, 1987, Lynch et al., 1999). In Klinefelter's syndrome, proliferation of the mammary ductal epithelium due to alteration in the androgen-oestrogen ratio has been implicated in the causation of cancer (Newman, 1997, Thomas, 1993). The reported incidence of Klinefelter's syndrome in MBC ranges from 3 to 4% (Lynch et al., 1999) with a 20 fold increased risk of developing breast cancer in comparison to males without this condition (Newman, 1997, Thomas, 1993).

Even though there has not been any causal relationship established, the commonest genetic mutations that have been associated with MBC are the polymorphism in the CYP17 gene involved in steroid synthesis and mutation of the CHEK2 (Falchetti et al., 2008) and PTEN tumour suppressor genes (Cowden syndrome) (Yang et al., 2010). CYP17 gene encodes for the P450c17α enzyme involved in the synthesis of oestrogen and androgens. The increased transcriptional activity due to the T-to-C polymorphism at the 5'untranslated region has been hypothesised to enhance the steroid hormone production and lead to alteration in the oestrogen-androgen ratio leading to an

increased risk of developing cancer (Carey et al., 1994). The level of evidence to support this association in MBC is weak and clearly more evidence is needed to elucidate this causative relationship (Dunning et al., 1998, Gudmundsdottir et al., 2003).

Cowden syndrome is an autosomal dominant disease characterized by multiple hamartomas and associated with germ line mutations in the PTEN tumour suppressor gene (Fackenthal et al., 2001). In women, Cowden syndrome has been implicated with an increased risk for breast and thyroid cancers as well as many non-cancerous lesions (Weiss et al., 2005). Similarly, it is proposed that the PTEN mutation contributes to MBC and an association has been drawn with earlier onset of cancer (Fackenthal et al., 2001).

The CHEK2 gene encodes for the cell cycle check point kinase involved in DNA repair process involving BRCA1 and P53 (Meijers-Heijboer et al., 2002, Weiss et al., 2005). Mutation of the CHEK2*1100delC has been associated with a 10 fold increased risk of developing MBC in families who do not harbour the BRCA1/2 mutations (Meijers-Heijboer et al., 2002). The Breast Cancer Consortium study has proposed that as much as 9% of MBC can be attributed to CHEK2 mutation (Meijers-Heijboer et al., 2002). However more recent studies failed to reproduce this finding with a far lower incidence of CHEK2 mutation noted (0 to 1.8%) in MBC patients (Neuhausen et al., 2004, Ohayon et al., 2004, Syrjakoski et al., 2004). Hence there is a school of thought that, the higher incidence noted in the Dutch studies (Meijers-Heijboer et al., 2002, Wasielewski et al., 2009) may be inherent to the population studied and may not be widely representative.

A germline mutation of the AR gene was explored in detail by various researchers, with some implicating long polyglutamine repeats of the AR receptor in MBC (Young et al., 2000) but not others (Friedman et al., 1997, Syrjakoski et al., 2003). Imbalance of the androgen and oestrogen ratio due to mutation of AR gene and binding of the AR promoter to oestrogen response elements (Haraldsson et al., 1998, Young et al., 2000) has also been implicated in the development of MBC.

There is more evidence emerging from high throughput studies on MBC genomic landscape. In a recent study, recurrent mutations of PIK3CA, GATA3, TP53 and MAP3K1 genes were observed in MBC (Piscuoglio et al., 2016). However in comparison to ER+/Her2- FBC, PIK3CA and TP53 mutations were less frequent in MBC with more frequent mutations of genes associated with DNA repair (PALB2 and FANCM) (Piscuoglio et al., 2016).

1.1.5 Clinical presentation

Male breast cancer presents typically as a hard eccentric non-tender breast mass in almost 75 to 95% of the patients (Gennari et al., 2004). The central sub-areolar region is the most common site for cancer in men (Goss et al., 1999, Stierer et al., 1995) compared to the upper outer quadrant in women. The Surveillance, Epidemiology and End Result (SEER) data confirms that there is no evidence of lateralisation amongst MBC (Weiss et al., 1996)]. The mean size of the MBC at diagnosis is larger than that of FBC (Giordano et al., 2004). Involvement of the nipple areolar complex is quite common (40 to 50%) as a reflection of the small amount of breast tissue present leading to skin retraction,

fixation, nipple discharge and areolar changes (Contractor et al., 2008). The presence of skin ulceration is a more common sign at presentation amongst MBC than FBC (Wagner et al., 1995). The central location of the tumour along with small amount of breast tissue results in early ulceration in MBC. Unlike FBC, axillary nodal involvement has been reported in as many as 40 to 55% of the patients reflecting the overall increased stage at diagnosis (Lefor and Numann, 1988). In a population based study it was shown that, axillary nodal metastases were found in 37.7% with MBC and 29.2% with FBC (Gentilini et al., 2007). When interpreting these findings, it is important to consider that the lower axillary nodal burden amongst women may have been influenced by the earlier detection of breast cancer through breast screening programmes.

1.1.6 Pathological features

Invasive ductal carcinoma is the most common histological type among MBC (64-93%) followed by papillary carcinoma (2.6 to 5%) (Cutuli et al., 1995, Donegan et al., 1998, Shaaban et al., 2012). Invasive lobular carcinoma is rare amongst males (Giordano et al., 2004, Tischkowitz et al., 2002). The rest of the histological types including medullary, tubular and mucinous cancer constitute less than 15% of MBC (Contractor et al., 2008). Ductal carcinoma in situ (DCIS) constitutes around 5% of MBC and in almost 75% of the cases it is of the papillary subtype which is a markedly higher proportion than that observed in women (Hittmair et al., 1998). Studies to date have not conclusively established the prognostic significance of grade in MBC (Giordano et al., 2002a, Ravandi-Kashani and Hayes, 1998).

1.1.7 Molecular classification

Genome-wide profiling techniques have been used to study FBC extensively in the last decade. Perou et al (2000) described distinct intrinsic subtypes (Luminal A, luminal B, Her2 enriched, Basal or triple negative) using gene expression profiling. Since then it has been established that, the Luminal A subtype has got the best outcome and worst outcome for Her2 and Basal subtypes (Sorlie et al., 2003). Since then the prognostic significance of these subtypes has been externally validated and refined (Hu et al., 2006). It has been also possible to use biomarkers as surrogate to replicate the prognostic significance of intrinsic subtypes using immunohistochemistry (Carey et al., 2006).

The genomic aberrations has also been used to classify FBC in to molecular subgroups. Fridlyand et al (2006) classified FBC in to 3 subgroups of clinical significance, simple, mixed amplifier and complex. These subgroups has been externally validated by Chin et al (2006). The high resolution array comparative genomic hybridisation (aCGH) method was used by Jonsson et al (2010) more recently to classify FBC in to six different subgroups. These were Luminal simple, Luminal complex, Basal complex, Amplifier, Mixed and 17q12 subgroups. The Amplifier and Mixed subgroups contained all FBC intrinsic subtypes, whereas Luminal complex subgroup consisted of Luminal B and majority of BRCA2. The Luminal A subtype was present in Luminal simple and complex groups. The Luminal simple subgroup was found to have the best prognosis in the group. The Basal complex group consisted of majority of BRCA1 patients. The 17q12 subgroup consists of Her2 enriched subtype and

had the worst prognosis. The findings from these studies further illustrate and consolidate the heterogeneous nature of FBC.

The rare nature of MBC resulted in a paucity of studies conducted using genome-wide profiling methods. Earlier studies using comparative genomic hybridisation (CGH) showed similarities amongst male and female breast cancer in chromosomal gains and losses (Rudlowski et al., 2006). However it was only recently more advanced aCGH methods were used in an attempt to classify MBC in to molecular subtypes (Johansson et al., 2011, Tommasi et al., 2010). Based on the results of the aCGH, Johansson et al (2011) managed to classify MBC in to, male simple and complex groups. There was some similarity between the male complex and female complex subgroup. Whereas, male simple group was less aggressive and was associated with better outcome.

There was attempt to further understand the similarities and differences between male and female breast cancer by classifying MBC using gene expression profiling (Johansson et al., 2012). They identified 2 unique subgroups, luminal M1 and luminal M2 which correlated well with male complex and simple subgroups respectively. However they were different from the well-established FBC intrinsic subtypes. The luminal M1 subgroup was more frequent and aggressive phenotype compared to luminal M2, which was associated with better prognosis.

1.1.8 Clinical management

The diagnostic methods used for evaluating men presenting with a breast symptoms are similar to those used in women. Patients who present with breast

lumps are evaluated using triple assessment. A careful history and clinical examination are critical in the evaluation of men presenting with breast symptoms. The paucity of breast tissue makes the radiological evaluation using mammography more challenging and hence traditionally assessed initially using an ultrasound. However, mammography has been shown to have good sensitivity (92%) and specificity (90%) for detecting MBC in some series (Evans et al., 2001). Conventionally tissue samples are obtained for histological diagnosis either using fine needle aspiration cytology (FNAC) or preferentially using core biopsy (CB).

The treatment of MBC is largely based on the evidence and experience gained from managing FBC. Similar to FBC, there has been a shift in the surgical management of MBC from radical mastectomy to simple mastectomy. This was supported by various studies which failed to show any overall or disease free survival advantage for MBC patients treated with radical procedures (El-Tamer et al., 2004, Margaria et al., 2000). The relatively small amount of breast tissue in men (1/100 of that of women) have led to fewer breast conservation surgeries to treat MBC (Goss et al., 1999). The indications for adjuvant radiotherapy in MBC remain controversial due to inconclusive and inconsistent results published in the literature. Some suggest that adjuvant radiotherapy should be given to all the MBC patients due to the lack of sufficient breast tissue preventing adequate surgical clearance (Vetto et al., 1999). However, the low recurrence rates achieved after surgery alone without radiotherapy have challenged this approach (Chakravarthy and Kim, 2002, Chung et al., 1990). Similar to the role of radiotherapy, radiotherapy field is also controversial as there is limited data to suggest which MBC patients should receive radiotherapy

to chest wall or axilla or supraclavicular fossa. However, there is general consensus to advice adjuvant radiotherapy to those patients with a high risk of local recurrence as identified in FBC trials such as those with large tumours, four or more involved axillary lymph nodes, extra-capsular nodal extension and those with close surgical margins (Contractor et al., 2008, Volm, 2003).

The management of axillary disease has evolved and changed over the past few decades. Prior to the introduction of sentinel lymph node biopsy (SLNB) in the late 90's, the standard management of axillary disease in men and women was axillary node clearance (ANC). However, since then the wide acceptance of SLNB as an accurate staging and perhaps therapeutic procedure in FBC made it the standard procedure in MBC patients without any radiologically or histologically proven axillary lymph node involvement (Cimmino et al., 2004, Schuchardt et al., 1996). Currently, SLNB is considered as the standard for managing patients with early primary MBC (Albo et al., 2003, Cimmino et al., 2004).

Adjuvant hormonal therapy has also evolved over the years from surgical anti-hormonal ablative procedures (Orchidectomy, Adrenalectomy and Hypophysectomy) through to hormonal manipulation using oral/injectable medications. Amongst the various agents tested, mild to moderate success has been reported for diethylstilboestrol (Ribeiro, 1976, Lopez et al., 1985), cyproterone acetate (Pannuti et al., 1982), androgens (Kantarjian et al., 1983, Ribeiro, 1976), luteinizing hormone releasing agonist alone or in combination with anti-androgens (Doberauer et al., 1988, Vorobiof and Falkson, 1987). There are not any prospective or randomized studies undertaken to establish

the role of adjuvant hormonal therapy in MBC. The current clinical practice of administering tamoxifen to MBC patients has emerged from its well established role in oestrogen receptor (ER) positive FBC and the established high ER positivity (75-80%) in MBC. The proven benefit of aromatase inhibitors (AI) in post-menopausal women with breast cancer has led to studies exploring its role in MBC. The first generation of AI such as aminoglutethimide were shown to be effective in MBC patients who had undergone orchidectomy (Harris et al., 1986, Patterson et al., 1980). In orchidectomised patients, the lack of additional oestrogen production from the testes leads to a more effective action of the AIs to prevent the conversion of androgen to oestrogen. The third generation AIs (Anastrazole, Exemestane and Letrozole) are more specific and are used in the management of post-menopausal women with FBC. However there are some theoretical reservations about aromatase activity in men and there is a scarcity of evidence within the literature especially randomized controlled trails regarding the role of third generation AIs in the adjuvant settings for MBC (Giordano et al., 2002b, Italiano et al., 2004).

Similarly, even though there is no conclusive evidence regarding the role of adjuvant chemotherapy in MBC patients, its beneficial role in high risk patients (locally advanced and/or node positive) has been established through large central reviews (Contractor et al., 2008) and prospective studies (El-Tamer et al., 2004, Patel et al., 1989). Various regimens have been tried in MBC, with historically the vast majority of the studies using CMF (Cyclophosphamide, Methotrexate and 5-fluorouracil) (Bagley et al., 1987, Patel et al., 1989), and more recently anthracycline and taxane based regimens (Giordano et al., 2005, Patel et al., 1989).

1.1.9 Prognosis and prognostic markers

In contrast to FBC, there has been little improvement in the survival of patients diagnosed with MBC (Johansson et al., 2012). The incidence of local recurrence is higher amongst MBC patients possibly due to the smaller volume of the breast tissue enhancing lymphatic infiltration and invasion of the underlying muscles (Wagner et al., 1995, Winchester, 1996). The reported 5 year survival rates for MBC ranges from 40 to 65% (Cutuli et al., 1995, Wagner et al., 1995). It has been shown that the worse prognosis seen amongst MBC patients in comparison to FBC disappear once it has been corrected for the age and stage at diagnosis (Evans et al., 2001, Adami et al., 1985, Willsher et al., 1997b). However worse survival is observed in male patients with stage III and IV disease compared to their female counterpart, whereas similar 5 year survival rates were observed for stage 0, I and II disease (Scott-Conner et al., 1999). The relative prognosis of MBC therefore remains somewhat uncertain.

Tumour stage (Giordano et al., 2004, La Vecchia et al., 1992, Fonseca et al., 2006) and axillary nodal status on its own (Guinee et al., 1993, Pich et al., 1999) has been shown over the years by various researchers to be the most consistent and reliable independent prognostic factor influencing overall survival (OS) in MBC patients. The role of tumour grade as a prognostic factor in MBC is controversial due to the lack of consistency in the published results (Salvadori et al., 1994). Smaller studies have shown a relative higher grade in MBC compared to FBC (Muir et al., 2003), whereas the larger SEER database analysis didn't reveal any significant difference (Giordano et al., 2004, Giordano et al., 2002a). Tumour size was also suggested to be an independent predictor

of overall survival by various researchers (Anderson et al., 2004, Giordano et al., 2004).

The prognostic role of various biomarkers has been explored in MBC without any biomarker identified conclusively as a prognosticator due to the inconsistencies in the published literature. Even though ER and PR are highly expressed in MBC, the prognostic role of both these biomarkers was insignificant after accounting for other clinical and pathological prognostic factors (Giordano et al., 2004, Goss et al., 1999, Truong et al., 2005, Wang-Rodriguez et al., 2002). Individual studies have shown better overall survival amongst ER positive MBC (Donegan et al., 1998, Wang-Rodriguez et al., 2002), whereas ER positivity was not found to be a significant predictor of overall survival on multivariate analysis (Goss et al., 1999). The roles of AR (Pich et al., 1999, Kwiatkowska et al., 2003), c-erbB2 (Joshi et al., 1996, Willsher et al., 1997a) and p53 (Wang-Rodriguez et al., 2002, Willsher et al., 1997a) have been similarly explored without reaching any consensus on their prognostic significance in MBC.

CHAPTER 2

Protein Biomarker Analysis in Male Breast Cancer Cohort

2.1 INTRODUCTION

Male breast cancer arises in a different hormonal milieu than that of FBC but it remains uncertain whether they behave similarly or not. Researchers over the years have tried to evaluate the level of expression and the interactions of various biomarkers in MBC that have been previously evaluated in FBC. However, the low incidence and prevalence of MBC led to small sample size and hence a lack of power of the observed findings.

High expression of steroid receptors, especially oestrogen (ER) and progesterone receptor (PR) has been observed in MBC compared to FBC (Contractor et al., 2008). It has been hypothesised that the up-regulation of the steroid receptor in an oestrogen depleted environment as seen in post-menopausal women may be responsible for this (Muir et al., 2003). However unlike the post-menopausal FBC, higher grade and proliferation rates were observed in ER positive MBC (Muir et al., 2003, Munoz de Toro et al., 1998). It was also shown that, grade 3 MBC were more frequently ER and PR positive and negative for p53 and ERBB2 compared to grade 3 FBC (Muir et al., 2003). This indirectly indicates that up-regulation of ER leading to activation of downstream targets such as p53 and/or ERBB2 may not persist in MBC. Moreover the oestrogen modulated proteins (Cathepsin D, protein S2 (pS2) and heat shock protein 27 (hsp27)) were also differentially expressed in male and female breast cancer (Muir et al., 2003).

It was previously established that proteins (Alpha-2-Zn-glycoprotein and Apolipoprotein D) under androgen control are differentially expressed in MBC

compared to FBC, even though their prognostic roles were uncertain (Muir et al., 2003). The functional studies performed have shown a greater role for androgen receptor (AR) in MBC compared to ER (Weber-Chappuis et al., 1996a). However the expression levels of AR in MBC have been variably reported within the literature, ranging from minimal up to 95% (Contractor et al., 2008). Hence it shouldn't be unreasonable to suggest that the ER function in MBC is different compared to FBC.

In an oestrogen depleted environment, up-regulation of ER can lead to an increased response to oestrogen targets such as the Bcl-2 (Nahleh and Girnius, 2006). Bcl-2 is an inhibitor of apoptosis and when up-regulated can lead to malignant transformation through genomic modification and downstream target activation (Muir et al., 2003). Bcl-2 was shown to be over expressed in MBC compared to FBC (Rayson et al., 1998, Weber-Chappuis et al., 1996a, Muir et al., 2003). It was proposed that, the high expression of Bcl-2 in MBC through stimulation of growth factors (Leek et al., 1994) can lead to inhibition of apoptotic cell death and therefore can influence carcinogenesis (Rayson et al., 1998). However, unlike in FBC, the prognostic role of Bcl-2 in MBC remains uncertain.

Cytokeratins (CK) were shown to play an important role in the carcinogenesis of FBC (van de Rijn et al., 2002, Korsching et al., 2002, Malzahn et al., 1998). Cytokeratin expression profiles vary between different breast epithelial compartments. Some cytokeratins (CK5/6, CK14 and CK17) are principally expressed in the basal/myoepithelial cell phenotype, whereas others (CK8, 18 and 19) represent luminal differentiation (Ciocca et al., 2006). Female breast

cancer was classified based on molecular profiling in to different subgroups. The CK, ER, PR and Her2 (Human Epidermal Growth Factor receptor -2) immunohistochemistry (IHC) expression was used as a surrogate for the earlier gene expression profiling described by Perou et al (2015). For example basal like subtype would be ER-, PR-, HER2- and CK5/6+; this surrogate for molecular subtype permits the use of IHC for the evaluation of breast cancers (Carey et al., 2006, Nielsen et al., 2004). The incidence of CK5/6 and CK14 positive tumours in FBC is approximately 10-15%, with higher incidences noted in African-Americans (van de Rijn et al., 2002) and similar observations are noted in MBC (Ciocca et al., 2006). However, unlike in FBC, the clinical significance or prognostic role of any of the cytokeratins in MBC has not been explored.

Human Epidermal Growth Factor receptor - 2 (Her2) over expression patterns and frequency in MBC is controversial with reported values ranging from 0-95% (Rayson et al., 1998, Ravandi-Kashani and Hayes, 1998). Unlike in FBC, there is still dispute about its prognostic significance in MBC. Over expression of Her2 has been associated with poor prognosis by some (Bloom et al., 2001, Reed et al., 2000), whereas others have failed to identify any prognostic significance in MBC (Tischkowitz et al., 2002, Muir et al., 2003).

The tumour suppressor gene p53 regulates the progression of the cell cycle in the presence of DNA damage. Over expression of p53 in FBC has been associated with a poor prognosis (Shpitz et al., 2000, Willsher et al., 1997b, Plesan et al., 2010). Conversely, p53 was over expressed only in < 10% of MBC with some researchers showing poor prognosis (Anelli et al., 1995, Joshi

et al., 1996, Pich et al., 1996, Weber-Chappuis et al., 1996a, Willsher et al., 1997b), whereas others showed no correlation (Clark et al., 2000). It has been proposed that, a different p53 mutational profile in males compared to females may result in lack of over expression of p53 leading to the inability of the immunohistochemistry methods to detect the protein (Muir et al., 2003).

Ki-67 is a known proliferation marker expressed in all cells that are not in the resting phase (G0) of the cell cycle and has been associated with poor survival in FBC (Pinder et al., 1995). Similar to FBC, correlation has been established between Ki67 and poor progression free survival in MBC (Rayson et al., 1998). It is estimated that, around 20 to 40% of the MBC show high expression of Ki67 (Wang-Rodriguez et al., 2002). However, majority of the studies in the past used varying cut-off for determining Ki67 expression. In MBC, some correlation was drawn between positive Ki67 expression and lymph nodal metastasis (Anderson et al., 2002). However, there is still lack of good quality evidence within the literature to substantiate the role of Ki67 as a prognostic biomarker in MBC.

Survivin is a member of the inhibitor of apoptosis family (IAP) and it regulates cell division, inhibits apoptosis and enhances angiogenesis (Velculescu et al., 1999, Tran et al., 1999). Due to its anti-apoptotic property, high survivin expression has been generally correlated with poor prognosis (Ryan et al., 2006). In FBC, survivin expression has been correlated with known poor prognostic markers by some (Brennan et al., 2008, Al-Joudi et al., 2007, Hinnis et al., 2007), whereas others have found no such correlation (Kennedy et al.,

2003). However, the only study conducted in a small cohort of MBC patients failed to establish any prognostic significance for survivin (Younis et al., 2009).

Epidemiological studies have shown high serum prolactin (PRL) to be associated with an increased risk of developing breast cancer in pre-menopausal (40%) and post-menopausal women (30%) (Tworoger and Hankinson, 2008). Unlike in the female breast, PRL does not influence the growth of normal male breast tissue. There is lack of evidence about the role of PRL or its receptor (PRLR) in MBC and to date there is only a single study within the literature which evaluated the biomarker expression of PRLR in MBC and in patients with gynaecomastia (Ferreira et al., 2008). They observed higher expression of PRLR in MBC compared to gynaecomastia tissue but there was no correlation with ER, PR, AR, grade or stage of the disease.

The evidence available within the literature suggests potential differences in biomarker expression between male and female breast cancer. However due to the scarcity of studies in MBC compared to FBC, there is lack of information about the clinical significance and prognostic role of biomarkers in MBC. Hence, it would be valuable to evaluate the expression, interactions and prognostic roles of various biomarkers (selected based on their role in FBC) in a large cohort of MBC.

2.2 HYPOTHESIS

We hypothesised that, there will be fundamental differences at protein level between male and female breast cancer that could explain MBC development and progression. In order to evaluate this, various known biomarkers with established role in FBC (ERα, ERβ1, ERβ2, ERβ5, Total PR, AR, CK5/6, CK14, CK18, CK19, p53, Bcl-2, Her2, E-cadherin, Ki67, Survivin, Prolactin and FOXA1) were studied immunohistochemically in MBC tissue microarrays.

2.2.1 Aims

1) Evaluate the expression levels of various biomarkers and correlate their expression with known clinical or pathological prognostic variables in a large cohort of MBC

2) Identify any prognostic or predictive role for selected biomarkers in MBC

2.3 METHODOLOGY

Immunohistochemistry analysis was performed in tissue micro arrays (TMAs) constructed from formalin fixed paraffin embedded (FFPE) tissue blocks from 428 male breast cancers. These blocks were collected through collaboration by Prof. Speirs with various pathologists across the United Kingdom (n=243), Canada (n= 53), Italy (n = 50), Hungary (n=42), Poland (n=31), and Nigeria (n=9). MBC tissues were incorporated in to 8 TMA blocks to standardise and maximise the information gathered using IHC analysis. The ethical approval for the study was obtained from the Leeds (West) Research Ethics Committee (06/Q1205/156) (Appendix 1, Page no. 177).

The clinical and pathological prognostic parameters were provided anonymously by the institution providing the MBC tissue blocks (where available) and held on a secure password-protected electronic database at the University of Leeds. The electronic data base (Patient Pathway Manager) held at Leeds Teaching Hospitals NHS Trust (LTHT) was searched retrospectively for tissue blocks originating locally. The variables collected included, age and date of diagnosis, type of surgery, adjuvant therapies, development of loco-regional recurrence, distant metastasis and survival data. Histological features including, type of cancer, tumour size, grade, nodal status, lympho-vascular invasion (LVI), presence of DCIS, hormonal receptor status and TNM staging information were collected from the pathological data base (CoPath) held at LTHT. The data was anonymised through assigning a unique MBC identification number to each tissue block which was stored in the central database.

2.3.1 Microtomy

The FFPE tissue was sectioned using the manual rotary microtome (Leica^{TM} rm2235). Disposable blades were used for sectioning with a clearance angle of 3º to 4º with the angle of the slope pre-set at 40º. The TMA block was fixed onto the microtome and trimmed using an old blade to remove any excess wax until a suitable tissue section was exposed. Subsequently the TMA block was placed on melting ice for 15 to 30 minutes to facilitate sectioning by cooling of the wax and the slight expansion of the tissue caused by imbibing. The tissue block was then sectioned at 5µm thickness. A gentle exhalation onto the block while cutting helped to alleviate the difficulty encountered in cutting a smooth flat section by expanding the block to give thicker sections. When a ribbon of 6 to 8 sections were cut it was transferred to a preheated water bath at 37º C by holding the first section using a forceps and the last section was gently lifted away from the blade using the back of a small squirrel hair brush.

Once floated in the heated water bath, the sections were separated from each other gently by using a fine pointed or curved forceps. The folds in the sections were rectified by gently teasing using a forceps and if unsuccessful such sections were discarded. Once separated, the sections were drawn on to a SuperFrost plus slide (SuperFrost, VWR) and were consecutively numbered. The slides were then held in a slide rack to drain the water and kept at 37º C for 72 hours in an incubator.

2.3.2 Tissue micro array construction

Most of the TMAs used (n =8) in the IHC analysis were already constructed (TMA 1 to TMA 7) (Shaaban et al., 2012) and I constructed one TMA as part of my research from FFPE blocks (n=18) received from collaborators in Portsmouth (TMA 8). FFPE blocks were sectioned at 5µm thickness and stained using haematoxylin and eosin method (H&E). The H&E stained sections were examined under the microscope (Olympus BX41) and tumour rich areas were marked using a fibre tipped permanent marker pen (Figure 1).

Recipient paraffin TMA blocks were constructed using melted paraffin wax (Surgipath Formula R). The melted wax was poured on to plastic moulds (Dispomoulds, Cellpath) with the cassette inserted on top of the mould and was left at room temperature overnight to solidify. Once solidified, all recipient TMA blocks were routinely X-rayed and only those without any cracks or bubbles were used.

A manual TMA instrument (MTA1, Beecher Instruments, USA) (Figure 2) was used for constructing the TMA. The small punch of the instrument was used to construct a hole in the recipient TMA block and 0.6mm diameter cores were punched out from the donor blocks in triplicate to reduce sampling errors using the large punch. These cores were then inserted in to pre- made holes in the recipient TMA blocks. For thin donor blocks, multiple cores were stacked in to each recipient hole to create deeper cores. The donor cores were arranged in the recipient TMA block precisely according to a pre-determined TMA plan. For orientation purposes, we incorporated marker cores using miscellaneous

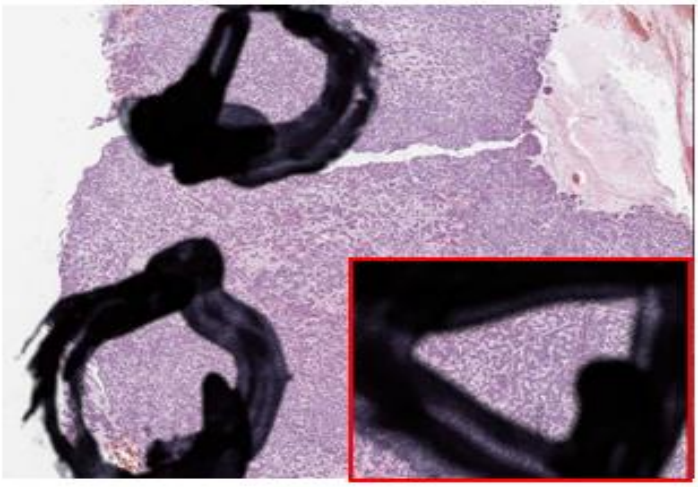


Figure 1: H&E stained slide marked with tumour rich areas

The inset represents magnified image of the marked area with breast cancer from where TMA core was taken subsequently.

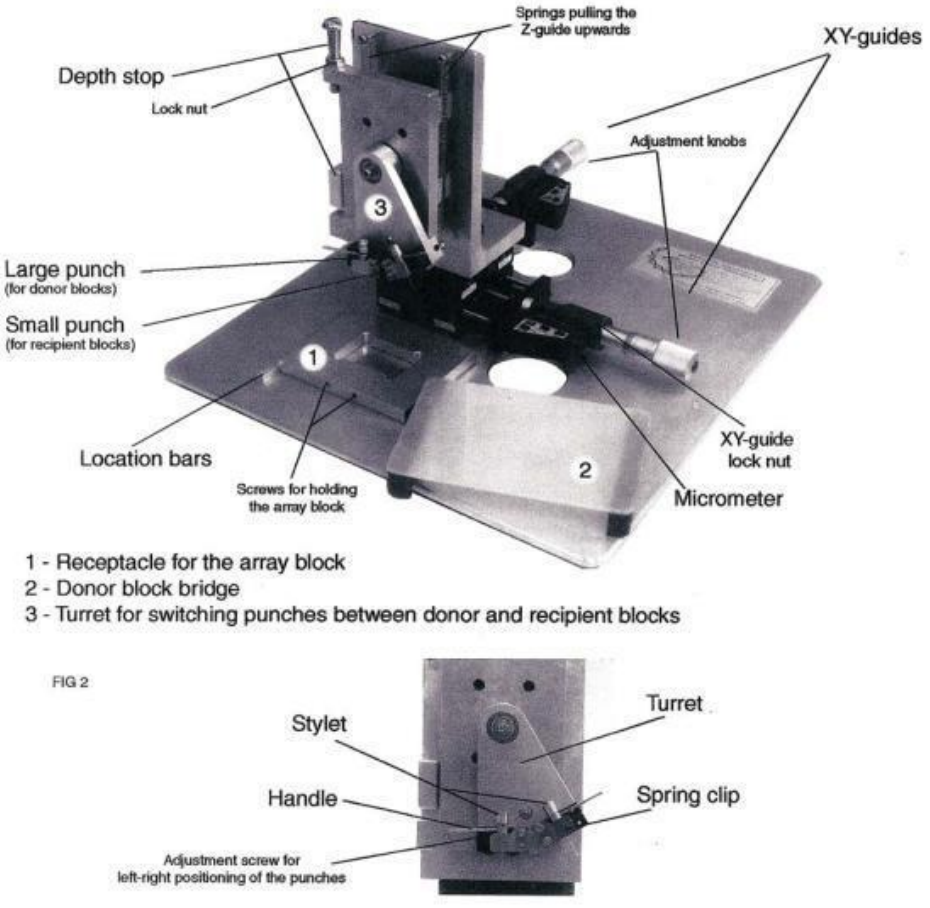


Figure 2: Manual Tissue Microarray instrument

tissues (sheep liver, lung, brain) in the TMA plan. After construction, TMA block was kept in the incubator at 37ºC overnight to facilitate embedding of the grafted donor tissue cores into the recipient block. Once removed from the incubator TMAs were stored at 4ºC. TMA blocks were sectioned using a manual microtome (Leica^{TM} RM2235) and the cellularity of the TMA sections were measured using H&E staining of the first and last sections. The sections were placed on Superfrost plus slides (Superfrost, VWR) and serially numbered. The sections were then dried in an incubator at 37ºC overnight. In order to minimise potential loss of antigenicity, it was a routine practice to apply a layer of wax to cover the sections for storage at 4ºC. Sections stored under the above conditions were generally used for IHC within 90 to 120 days.

2.3.3 Immunohistochemistry

Immunohistochemistry analysis was performed in the MBC TMAs (TMA 1 to 8) for 18 biomarkers (Table 2). The IHC staining and scoring for ERα, Total PR, CK5/6, CK14, CK18, CK19, Bcl2, AR, ERβ1, ERβ2, ERβ5 and Her2 were performed by my predecessors (Shaaban et al., 2012) in TMAs 1 to 4, 6 and 7. Similarly the IHC analysis for E-cadherin and p53 were already performed in TMAs 1 to 5. During the conduct of the research project, I have completed the IHC analysis in the remaining TMAs and scored them for the above biomarkers. In addition, I performed the IHC and scoring for Survivin, Prolactin, FOXA1 and Ki67 in all the MBC TMA blocks.

2.3.3.1 General steps

TMA blocks were sectioned at 5 µm thickness and placed on Superfrost plus slides (Superfrost, VWR) prior to incubating at 37ºC overnight. The slides were then de-waxed by passing through a series of xylene (3x5 minutes), rehydrated through graded ethanol 100% (2x3 minutes), 90% (1x3 minutes) and 70% (1x3 minutes) before washing in running tap water for 2 minutes. The slides were subjected to peroxidase block to inhibit endogenous peroxidase activity by immersing in 10% hydrogen peroxide (H_2O_2) (180 ml of methanol and 20 ml of H_2O_2) for 10 minutes. The slides were then washed in running tap water for 5 minutes prior to heat induced antigen retrieval in a pressure cooker. The slides were placed in boiling 1% antigen unmasking solution (Vector Laboratories) and cooked at full pressure for 2 minutes. The slides were then immediately cooled under running tap water for 5 minutes.

Table 2: Antibody specification for immunohistochemistry

Antibody	Clone	Dilution	Manufacturer	Catalogue Number	Secondary antibody detection method
ER alpha	1D5	1:100	Dako	M7047	DAKO Envision Kit (Anti-mouse)
FOXA1	ab55178	1:500	Abcam	ab55718	DAKO Envision Kit (Anti-mouse)
Total PR	PgR 636	1:200	Dako	M3569	DAKO Envision Kit (Anti-mouse)
CK5/6	D5/16 B4	1:100	Dako	M7237	DAKO Envision Kit (Anti-mouse)
CK14	LL002	1:500	BIO-RAD	MCA890F	DAKO Envision Kit (Anti-mouse)
CK18	CY-90	1:500	SIGMA	C 8541	DAKO Envision Kit (Anti-mouse)
CK19	RCK 108	1:150	Dako	GA615	DAKO Envision Kit (Anti-mouse)
E- cadherin	NCH-38	1:100	Dako	M3612	DAKO Envision Kit (Anti-mouse)
Ki67	MIB1	1:100	Abcam	ab124929	DAKO Envision Kit (Anti-mouse)
p53	DO-7	1:1000	Dako	M7001	DAKO Envision Kit (Anti-mouse)
Bcl2	124	1:200	Dako	M0887	DAKO Envision Kit (Anti-mouse)
Her-2	PN2A	1:25	Dako	K5204	DAKO Envision Kit (Anti-mouse)
Prolactin	B6.2	1drop:3000	Thermo Scientific	MA5-11955	DAKO Avidin/Biotin detection kit
Survivin	D8	1:25	Santacruz	sc-17779	DAKO Avidin/Biotin detection kit
AR	AR441	1:100	Dako	M3562	DAKO Avidin/Biotin detection kit
ER β1	PPG5/10	1:20	BIO-RAD	MCA1974S	Menarini diagnostic intellipath FLX autostainer
ER β2	57/3	1:20	BIO-RAD	MCA2279	
ER β5	5/25	1:50	BIO-RAD	MCA4676T	

The heat induced antigen retrieval was performed in pressure cooker using Vector low pH antigen unmasking solution in all the biomarkers except for ERβ1, ERβ2 and ERβ5. The Access Revelation™ buffer solution was used for antigen retrieval in ERβ1, ERβ2 and ERβ5. The primary antibody was incubated for one hour at room temperature for CK14, CK18, CK19, ERβ1, ERβ2 and ERβ5. The rest of the biomarkers were incubated with primary antibody overnight at 4^0C. The details of the scoring methods used for each biomarker is described in section 2.4.1 through to 2.4.4 (Page no. 44 to 60)

2.3.3.2 EnVision method

After completion of the above steps, Shandon cover plates were applied to the slides prior to inserting it in to a Shandon Sequenza apparatus (Thermo Pharmaceuticals). The slides were washed initially with phosphate buffered saline (PBS) prior to subjecting to casein block to eliminate non-specific background staining. We used 10% of casein (Vector Laboratories) in PBS (100

µL per slide) for 20 minutes to achieve the casein block followed with one 5 minutes wash with PBS. The slides were then incubated with the primary antibody at 100 µL dilutions overnight or for 1 hour. Overnight incubation was done in a humid chamber at 4ºC, whereas 1 hour incubation was done at room temperature. The details of the primary antibody used for each biomarker, the concentration used and the incubation times are described in Table 2. In every batch, a positive control of FBC TMA was used and a negative control was run by incubating with PBS instead of primary antibody.

After incubation with the primary antibody, the slides were washed with PBS (2x5 minutes) and 100 µL/slide of secondary antibody (Envision/HRP (DAKO)) was applied for 40 minutes. This was followed with a PBS wash (2x5minutes) and the slides were then removed from the sequencer to be placed on a humid tray. In order to visualise the reaction of the biomarker with that of the antibody, 100µL of 3, 3'- diaminobenzidine (DAB) (Vector Laboratories) was applied to each section. After 10 minutes, the excess of DAB was removed from the slides and the sections were washed in running tap water for 2 minutes. The counter staining was performed by passing the slides through copper sulphate for 1 minute, Mayer's haematoxylin for 2 minutes and Scott's substitute for 2 minutes. Each step was followed with washing the slides in running tap water for 1 minute each. The sections were dehydrated by passing through absolute ethanol (1x10 seconds; 1x30 seconds; 2x5 minutes) and cleared using serial passage through xylene (3x2 minutes) before being mounted in DPX.

2.3.3.3 Access Revelation buffer solution

Access Revelation™ buffer solution was used for both de-paraffinization and antigen retrieval in the pressure cooker for IHC staining of ERβ1, ERβ2 and ERβ5 in the intelliPATH FLX™ autostainer (Menarini diagnostic) (The immunostaining was performed by Mr Mike Shires, Laboratory Technician). It has several advantages as it is non-toxic, non-inflammable and odourless in comparison to alcohol and xylene. Moreover it reduces non-specific background staining, blocks the endogenous peroxidise and has a colour coded pH indicator.

2.3.3.4 Avidin-Biotin method

After completion of the antigen retrieval as described above, the slides were covered with Shandon cover plates prior to transferring in to a sequenza apparatus. The slides were washed with PBS for 5minutes. Subsequently, 2 drops of Avidin solution (Vector Laboratories) (300µL/slide) were applied to the slides which were left at room temperature for 10 minutes. The slides were washed with PBS (2x5 minutes) before applying 2 drops of Biotin solution (Vector Laboratories) (300µL/slide) and were left at room temperature for 10 minutes. This was followed by PBS wash (2x5 minutes) prior to the slides being incubated with the primary antibody (100µL/slide) diluted using the antibody diluent reagent (Invitrogen). The sections were left to incubate with primary antibody over night at 4ºC in a humid chamber.

The next day, slides were washed with PBS (2x5 minutes) before applying 1 drop (100µL/slide) of secondary antibody (ChemMate kit, DAKO, Bottle A). The slides were left at room temperature for 30 minutes before washing with PBS

(3x5 minutes). Subsequently, 1 drop (100µL/slide) of Avidin-peroxidase (ChemMate kit, DAKO, Bottle B) was applied for 30 minutes in the room temperature. The slides were then washed with PBS (3x5 minutes) before removing from the Sequenza apparatus. The visualisation of the antigen-antibody reaction using the DAB chromogen and counter staining methods is similar to those described under the two-step Envision system. The Avidin-Biotin method is schematically illustrated below in Figure 3.

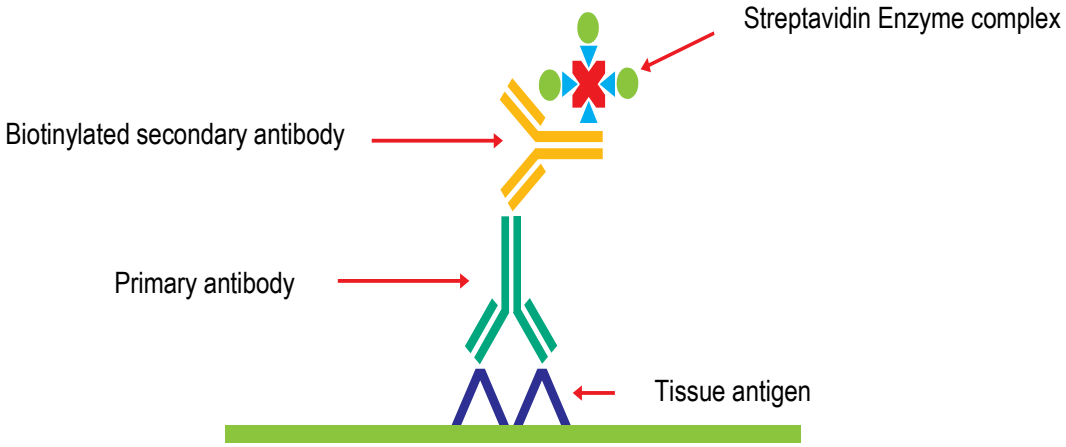


Figure 3: Labelled Streptavidin-Biotin method

Schematic represenation of streptaviding-biotin method. The multiple biniding sites between tetravalent avidin and biotinylated secondary antibody provides amplification for the reaction and hence improved sensitivity for detection of the tissue antigen.

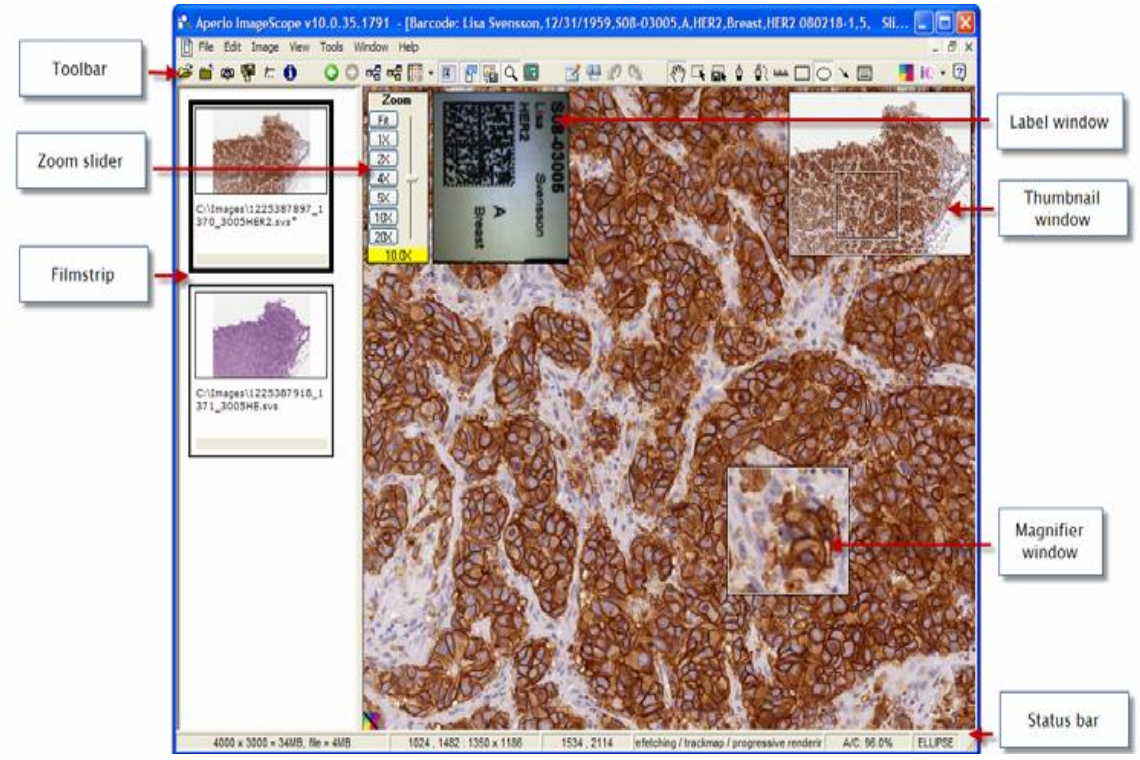


Figure 4: Aperio ImageScope™

A screen shot depicting various toolbars available to facilitate IHC scoring using the Aperio ImageScope™ software.

2.3.4 Antibody optimisation

All antibodies used were initially tested in the laboratory using positive controls as depicted in the manufacturer's specification sheet. It was a routine practice to note the concentration described within the literature or by the manufacturer as a reference point and to use a concentration above and below that for confirming the correct dilution of the antibody on whole tissue sections. The maximum dilution of the antibody which gave equivocal and best staining of the tissue antigen with least possible background and non-specific interaction was selected (Howat et al., 2014). In order to ensure optimal results in TMAs, we ran the same experiment in identical conditions using small control TMA containing 10 FBC cores. Routinely we used antigen diluent reagent (Invitrogen) for diluting all the antibodies and PBS to wash the slides except in the case of Androgen receptor (AR), where we used 2% Tween 20 (Polysorbate surfactant).

2.3.5 Immunohistochemistry scoring

The TMAs once stained were scanned (ScanScope XT, Aperio) at 20x magnification. The scanned slides were viewed and manually scored using ImageScope™ software (Aperio). The ImageScope™ as a software was easy to manoeuvre with various tool bars (Figure 4). The individual cores can be visualised using the pre-determined objective magnification at which the slides have been scanned (Figure 4). The provision of the "film strip" in a smaller window ensures that the user does not lose the orientation after magnification. The provision of axes and gridlines was especially useful when counting of the

individual nuclei was required e.g. for scoring Ki67 and Survivin. There are various other adjuncts that can be utilised during IHC scoring like the magnifying pane, thumbnail and snap shots for saving the images.

2.3.6 Statistical analysis

The clinical and pathological independent variables including, age at diagnosis, tumour size, grade, nodal status, LVI, and various biomarker expression were evaluated against overall survival (OS) as the dependent variable. The categorical and continuous variables were distinguished and appropriate statistical tests were performed. The distribution of the data was schematically checked using a histogram and normality tests were performed to determine the type of descriptive statistics required for continuous variables. A power calculation to determine the sample size was not possible due to rare nature of MBC as well as due to the retrospective and exploratory nature of the study.

Pearson's correlation coefficient was used for normally distributed continuous explanatory variables. A Spearman's correlation coefficient was used if the distribution was non-parametric or if the variable was categorical. The measure of agreement was calculated using the κ statistics for categorical variables. The type of tests used for determining the measure of association between the variables was dependent on the outcome variable. The type of test used for each outcome is described explicitly in the results section.

The strength of association between clinical and pathological variables, various bio-markers and primary outcome measure (OS) was ascertained using Cox logistic regression analysis. The prognostic variables that were significant on

univariate analysis were entered into a multivariate analysis model to identify independent predictors of OS at 5 and 10 year period. The overall survival analysis was performed using a Cox proportional hazard model. The survival curves were plotted using the Kaplan-Meier method and compared using the Log rank test. The data was analysed using SPSS (Statistical Package for Social Sciences) version 19 software and a P value of ≤ 0.05 was considered as statistically significant.

2.4 RESULTS

Immunohistochemical analysis was performed in TMAs constructed from 428 MBC tissue blocks. Cores were missing in 19 patients which were excluded from the statistical analysis. The demographic details and/or tumour characteristics (tumour size/nodal status/TNM staging) were missing in 28 patients. The follow-up data alone was not available for 37 patients. In another 143 patients, demographic data, tumour characteristics and follow-up information was not available. Therefore, the final analysis of descriptive statistics, measure of association and correlation was performed in 238 patients (Figure 5). Similarly survival analysis was performed in 229 patients, in whom the follow-up data was complete (Figure 5).

The median age of the cohort was 68 years (IQR = 17 years) (Figure 6). The tumour characteristics are summarised in Table.3. In more than three quarter of the patients the tumour was < 5 cm in size. The most common histological type was ductal cancer (88.2%), with only a small proportion being lobular (0.8%). The majority of the tumours were grade 2 or 3 and only 10% of the tumours were grade 1. Axillary nodal metastases were present in 53% of the cases. Amongst node positive cases (n=113), most were N1 (n=55, 48.6%) tumours. In comparison to FBC, early lympho-vascular invasion occurs in MBC due to the lack of distinctive tissue planes (Wagner et al., 1995, Contractor et al., 2008). However, metastatic involvement of axillary lymph nodes was only found in just over half of the cases and most had ≤ 3 involved lymph nodes.

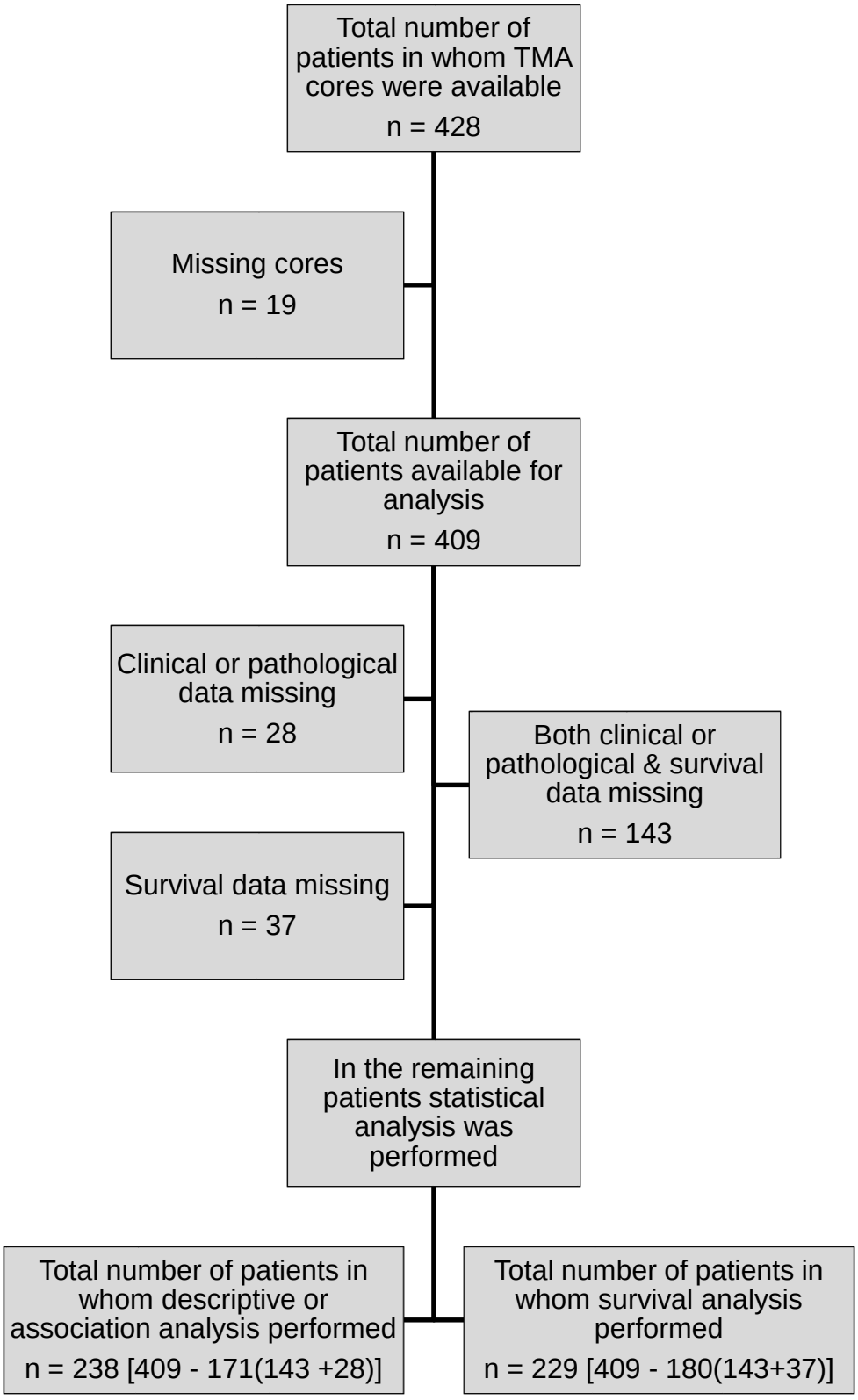


Figure 5: Schematic representation of MBC samples included and excluded from analysis

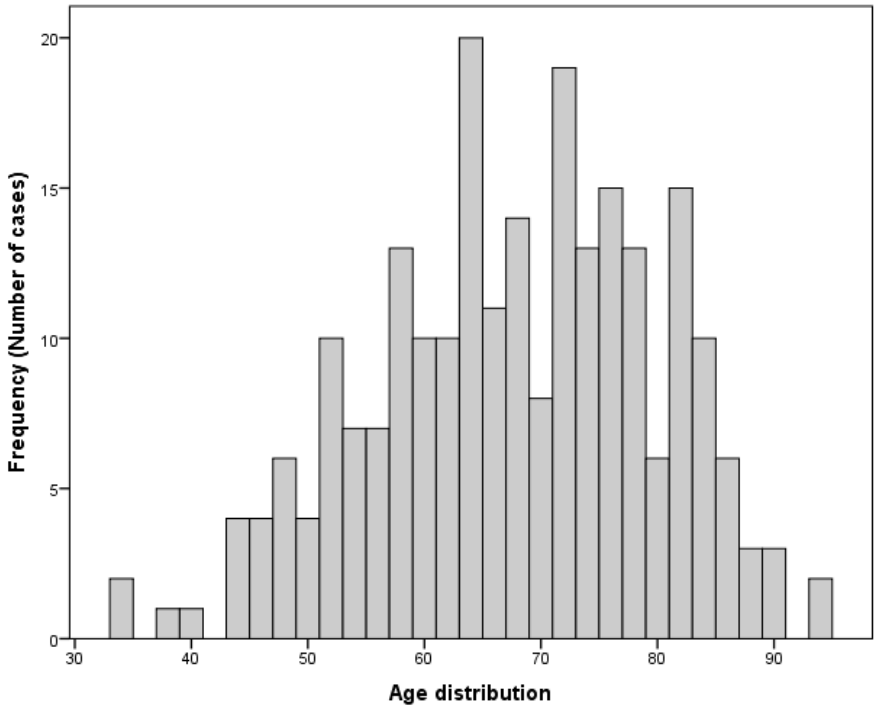


Figure 6: Histogram depicting age distributions in the male breast cancer cohort

Characteristics	Frequencies
Tumour size (n = 208)	
T1	104 (50%)
T2	91 (43.8%)
T3	11 (5.3%)
T4	2 (1%)
Histological type (n = 238)	
Ductal	210 (88.2%)
Lobular	2 (0.8%)
Mixed	7 (3%)
Special type	19 (8%)
Tumour grade (n = 235)	
G1	24 (10.2%)
G2	123 (52.3%)
G3	88 (37.4%)
Nodal status (n=223)	
Node positive	113 (51.1%)
Node negative	99 (44%)
No axillary surgery	11 (4.9%)
TNM nodal status (n = 212)	
N1	55 (25.9%)
N2	39 (18.4%)
N3	19 (9%)
Node negative	99 (46.7%)

Table 3: Histopathological characteristics of the MBC cohort

2.4.1 Hormonal Biomarkers

2.4.1.1 Oestrogen receptor α (ERα)

The percentage of positive nuclear immunoreactivity as well as the intensity of the nuclear staining was considered for determining ERα expression (Allred et al., 1998). An Allred score of > 2 was considered positive for ERα expression (Shaaban et al., 2012). The ERα immunoreactivity was nuclear specific in our cohort and there was no background cytoplasmic staining observed (Figure 7a). ERα was positive in 212/226 cases (93.8%) and negative in the rest (14/226; 6.2%).

2.4.1.2 Oestrogen receptor β1 (ERβ1)

Similar to ERα, both the percentage and intensity of the nuclear immunoreactivity was determined for ERβ1 using the Allred method (Allred et al., 1998). The ERβ1 staining was nuclear specific in most of the cases with mild to moderate background cytoplasmic staining (Figure 7b). Only nuclear staining was considered for determining the Allred score. Allred score of > 3 was considered positive for ERβ1 expression (Shaaban et al., 2012). ERβ1 was positive in 166/218 cases (76.1%) and was negative in 52/218 cases (23.9%).

2.4.1.3 Oestrogen receptor β2 (ERβ2)

ERβ2 was scored using the Allred method accounting for both the intensity and percentage of nuclear staining (Allred et al., 1998). The staining was observed in the nuclei of epithelial cells and similar to ERβ1 there was mild background cytoplasmic staining observed in most of the TMA cores (Figure 7c). Allred score of > 3 was considered positive for ERβ2 expression

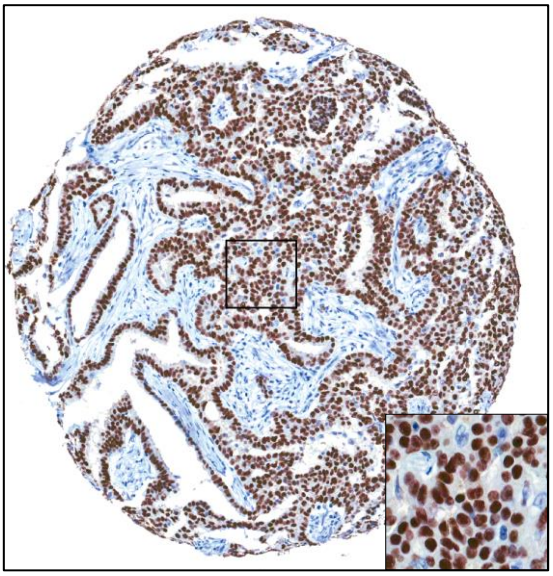


Figure. 7a

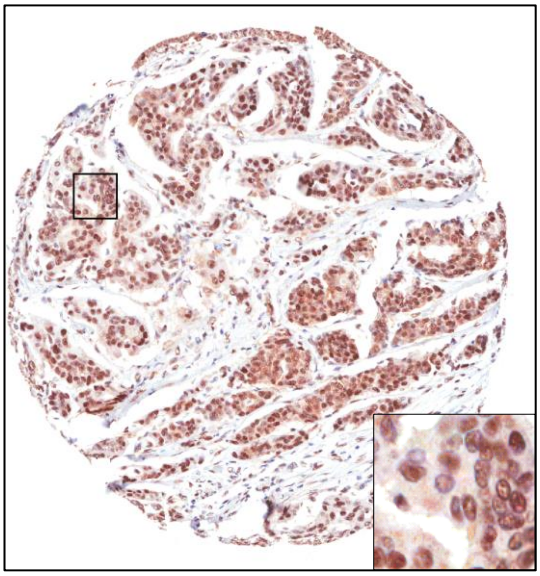


Figure. 7b

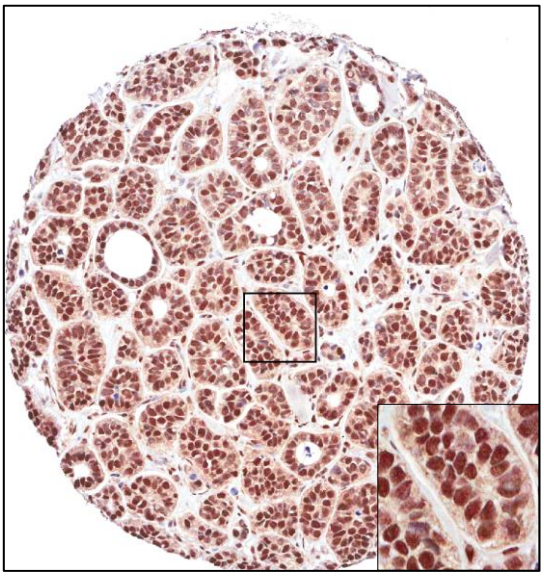


Figure. 7c

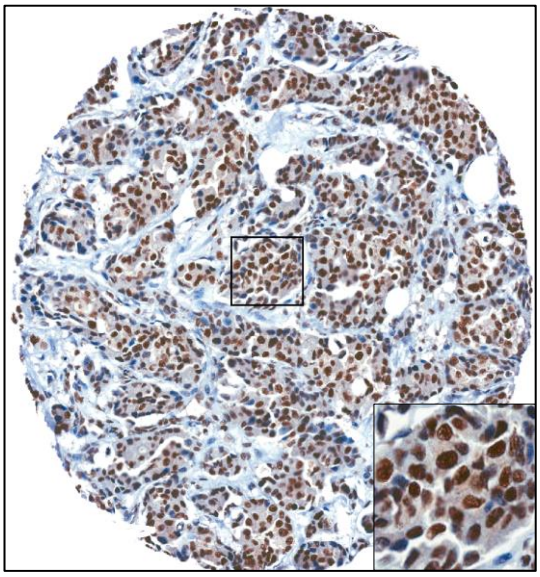


Figure. 7d

Figure 7: Expression of ERα, ERβ1, ERβ2 and ERβ5 in MBC

a) Unequivocal nuclear staining of ERα. b) Nuclear staining of ERβ1 with mild background cytoplasmic staining. c) Nuclear staining of ERβ2 with mild to moderate background cytoplasmic staining. d) Nuclear staining of ERβ5 with mild background cytoplasmic staining. Inset – Illustration of 20x magnified image of the marked square depicted in the TMA core.

(Shaaban et al., 2012). ERβ2 was positive in 156/218 cases (71.6%) and negative in 62/218 (28.4%).

2.4.1.4 Oestrogen receptor β5 (ERβ5)

The Allred method was used for determining the immunoreactivity of ERβ5 (Allred et al., 1998) and a score of > 3 was considered as positive (Shaaban et al., 2012). There was strong nuclear staining with mild to moderate background cytoplasmic staining observed in the MBC TMA cores (Figure 7d). Compared to other hormonal biomarkers, ERβ5 IHC was performed in a small cohort of our patients (n = 120) due to the exhaustion of the TMA. Similar to the other oestrogen receptors, most of the cases were positive for ERβ5 expression (97/120; 80.8%) and negative in the rest (23/120; 19.2%).

2.4.1.5 Androgen receptor (AR)

The Allred method was used for scoring AR expression (Allred et al., 1998) and a cut off of > 4 was considered as positive (Shaaban et al., 2012). The AR expression was nuclear specific with occasional background cytoplasmic immunoreactivity (Figure 8a). Compared to other hormonal biomarkers, AR was positive in 125/224 (55.8%) and was negative in the rest 99/224 (44.2%).

2.4.1.6 Progesterone receptor (PR)

The PR expression was nuclear specific in the cancer epithelial cells (Figure 8b). The Allred scoring system was used for scoring PR and a cut off of > 2 was used for determining PR expression (Allred et al., 1998, Shaaban et al., 2012). PR was positive in 186/221 cases (84.2%) and negative in 35/221 (15.8%).

2.4.1.7 Prolactin

The immunoreactivity of prolactin was seen in the cell membrane/cytoplasm of the cancer epithelial cells (Figure 8c). There was no immunoreactivity within the cancer cell nuclei or stromal elements. The intensity (1 = weak, 2 = moderate, 3 = strong) and extent (0 = no staining, 1 = 1 to 10%, 2 = 11 to 40%, 3 = 41 to 75% and 4 = >75%) of the staining was determined in each TMA core. The individual scores were then multiplied to obtain a score ranging from 0 to 12 and a score of > 3 was considered positive for prolactin expression (Bratthauer et al., 2010, Ferreira et al., 2008). In our cohort, prolactin was negative in most of the cases 148/219 (67.6%) and positive in the remaining 71/219 cases (32.4%).

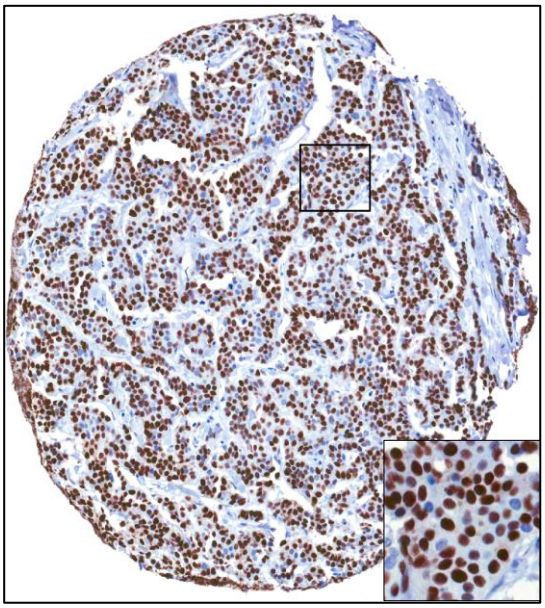


Figure. 8a

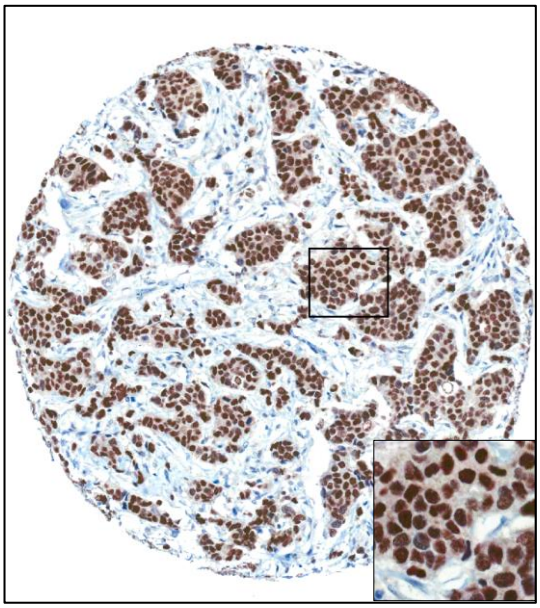


Figure. 8b

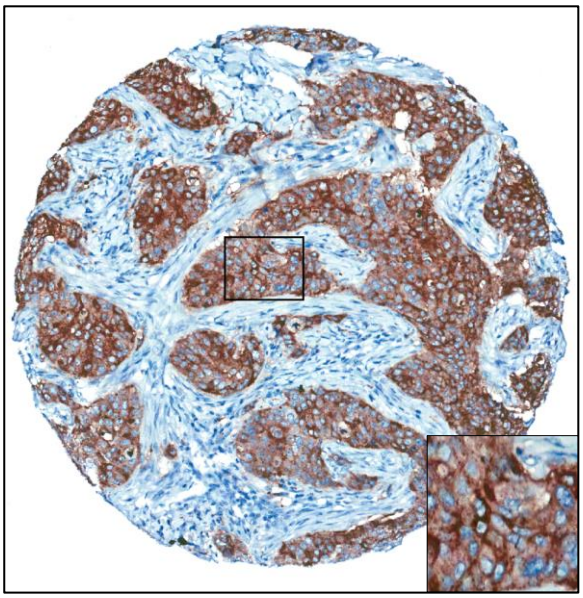


Figure. 8c

Figure 8: Expression of AR, PR and Prolactin receptor in MBC TMA cores

a) Unequivocal nuclear staining of AR in MBC. b) Nuclear staining of PR with mild background staining. c) Cytoplasmic membranous staining of prolactin. Inset – Illustration of 20x magnified image of the marked square depicted in the TMA core.

2.4.2 Cytokeratins

Analysis was undertaken for CK5/6, CK14, CK18 and CK 19 in the MBC cohort. Only unequivocal cytoplasmic and membranous staining in the cancer epithelial cells was considered for determining the cytokeratin expression (Simpson et al., 2005). The cytokeratin staining was scored semi-quantitatively into 5 categories; <5%, 5 to 25%, 26 to 50%, 51 to 75% and > 75% of the cancer epithelial cells stained. A score of 5% or above was considered as positive for cytokeratin expression in the MBC cohort.

2.4.2.1 Cytokeratin 5/6 (CK5/6)

The CK5/6 staining was detected only in the cytoplasmic and membranous compartment of the cancer epithelial cells (Figure 9a). Most of the cases were negative for CK5/6 (192/225; 85.3%) and positive in the rest (33/225; 14.7%).

2.4.2.2 Cytokeratin 14 (CK14)

Similar to other basal cytokeratin markers, only 4/221 (1.8%) cases were positive for CK14 and rest were negative (217/221; 98.2%). The CK14 staining was observed in the cancer epithelial cell cytoplasmic membrane.

2.4.2.3 Cytokeratin 18 (CK18)

The CK18 expression was positive in most of the MBC patients (217/225; 96.4%) and negative in 8/225 cases (3.6%). The CK18 expression was specific for the cytoplasmic membrane in most of the cases (Figure 9b).

2.4.2.4 Cytokeratin 19 (CK19)

The CK19 staining was observed only in the cytoplasmic membrane in the MBC cohort (Figure 9c). Similar to CK18, CK19 was positive in most of the MBC patients (211/218; 96.8%) and negative in 7/218 cases (3.2%).

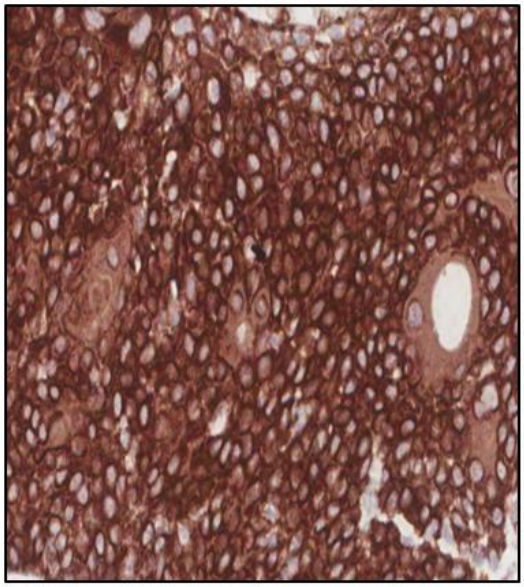


Figure. 9a

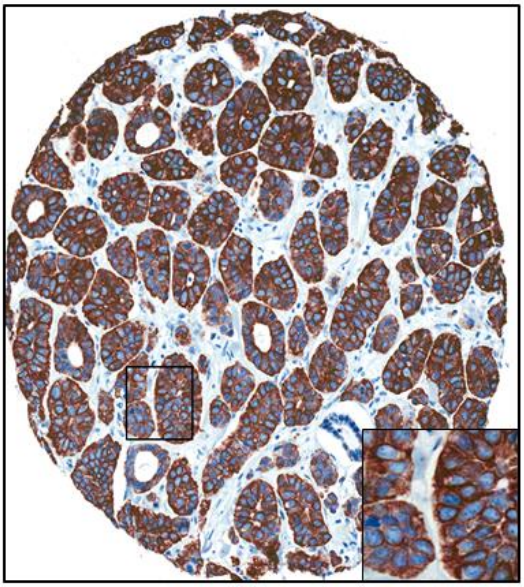


Figure. 9b

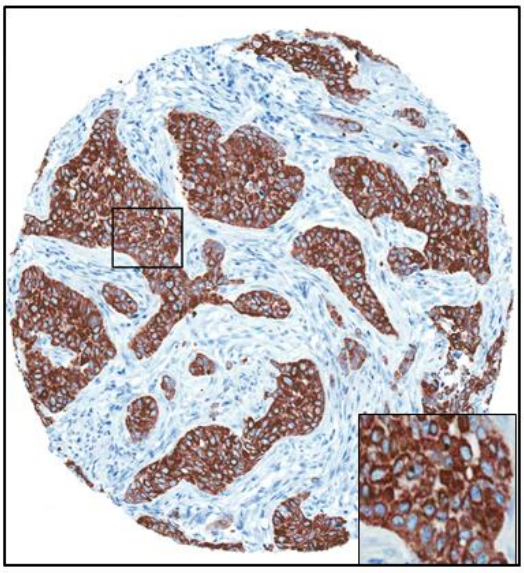


Figure. 9c

Figure 9: Expression of cytokeratins in MBC TMA cores

a) Cytoplasmic and membranous staining of CK5/6 (20x magnification). b) Cytoplasmic and membranous staining of CK18. c) Cytoplasmic and membranous staining of CK19. Inset – Illustration of 20x magnified image of the marked square depicted in the TMA core.

2.4.3 Proliferation and apoptosis markers

2.4.3.1 Ki67

The Ki67 IHC was performed adhering to the International Ki67 in Breast Cancer Working Group Consensus recommendations (Dowsett et al., 2011). As Ki67 is a proliferative marker, only unequivocal nuclear staining was considered for determining its expression in MBC tissue. Since the intensity of the nuclear staining was not shown to have any prognostic relevance, it was not determined while scoring the TMAs. In each TMA core a minimum of 500 cancer nuclei were counted and the percentage of DAB stained nuclei was determined. The Ki67 score was used as a continuous variable for descriptive analysis and for determining its association with other biomarkers and clinico-pathological prognosticators. However for survival analysis, the Ki67 score in percentage was dichotomised in to a high and low score. A score of ≥ 14% was considered as high, as it was determined against an important distinction in the underlying biology (Luminal A vs. Luminal B) compared to clinical outcomes or median Ki67 values as used in the past (Cheang et al., 2009).

The background cytoplasmic staining with SP6 and MIB1 antibody for Ki67 has been previous described in the literature and acknowledged by the International Ki67 in Breast Cancer Working Group (Dowsett et al., 2011). The IHC for Ki67 was initially performed using SP6 antibody which resulted in strong background cytoplasmic staining with mild to moderate patchy nuclear staining (Figure 10a). Hence, MIB1 antibody was used for Ki67 IHC in our MBC cohort. MIB1 antibody produced strong nuclear staining with moderate to strong

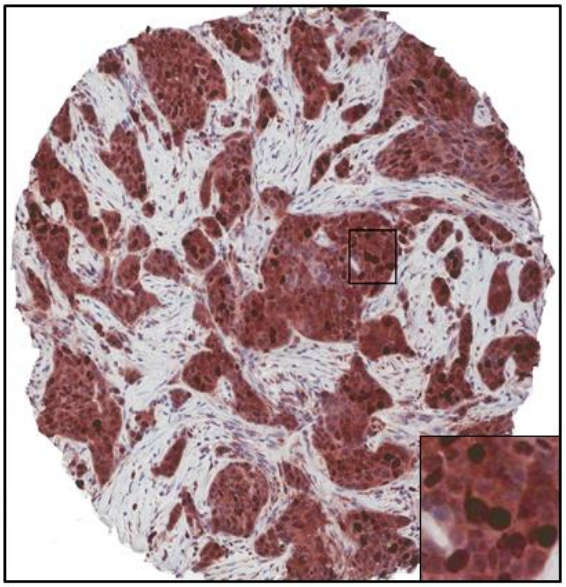


Figure. 10a

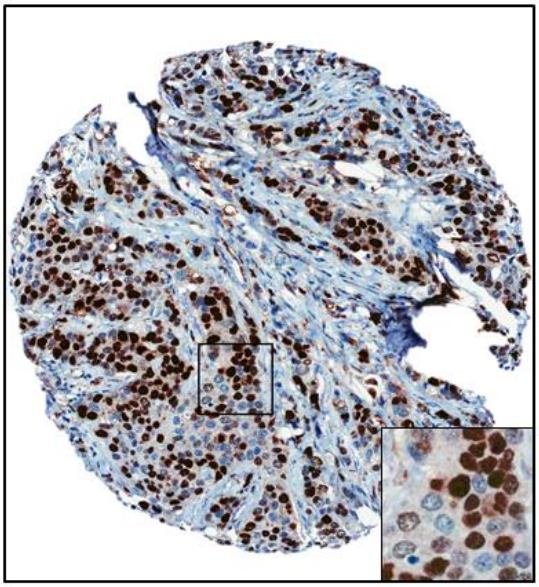


Figure. 10b

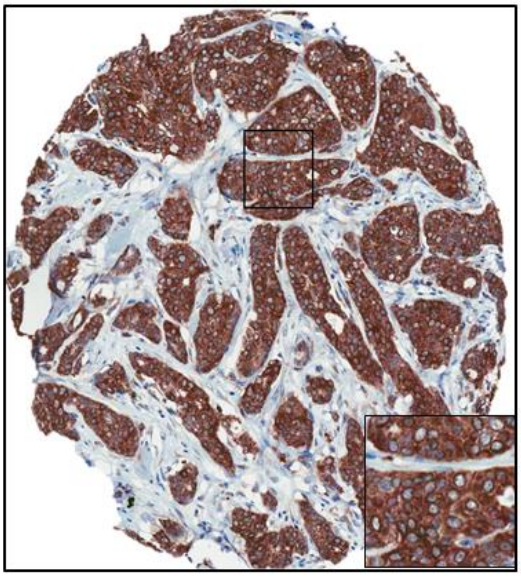


Figure. 10c

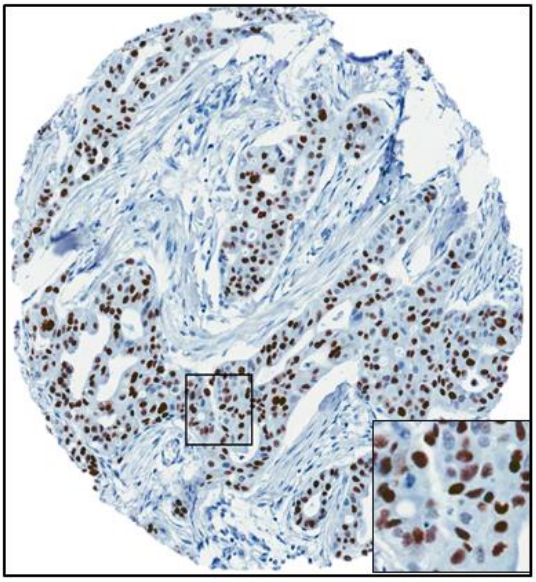


Figure. 10d

Figure 10: Expression of Ki67, Bcl2 and P53 in TMA cores

a) Ki67 IHC using SP6 antibody in FBC TMA – Nuclear staining of Ki67 observed along with strong background cytoplasmic staining b) Unequivocal nuclear staining of Ki67 using MIB1 clone of antibody in MBC TMA. c) Cytoplasmic and membranous staining of cancer cells by Bcl2 in MBC TMA d) Unequivocal nuclear staining of cancer cells by P53 in MBC TMA. Inset – Illustration of 20x magnified image of the marked square depicted in the TMA core.

intensity background cytoplasmic staining in our cohort (Figure 10b). It was noted that in most of the TMA cores only a proportion of the cancer nuclei were positive for Ki67. This could be due to the non-proliferative nature of the individual cells or the cell being in the G0 phase of the cell cycle during which Ki67 is not expressed. Optimisation of the MIB1antibody using various dilutions was performed first in full section breast cancer tissue prior to optimising it in control TMAs to ensure unequivocal nuclear staining.

The distribution of Ki67 expression in our entire cohort is depicted in Figure 11. The dichotomisation of Ki67 percentage score using a cut of ≥ 14% showed a high Ki67 expression in 63/202 cases (31.2%) and a low score in the rest (139/202; 68.8%).

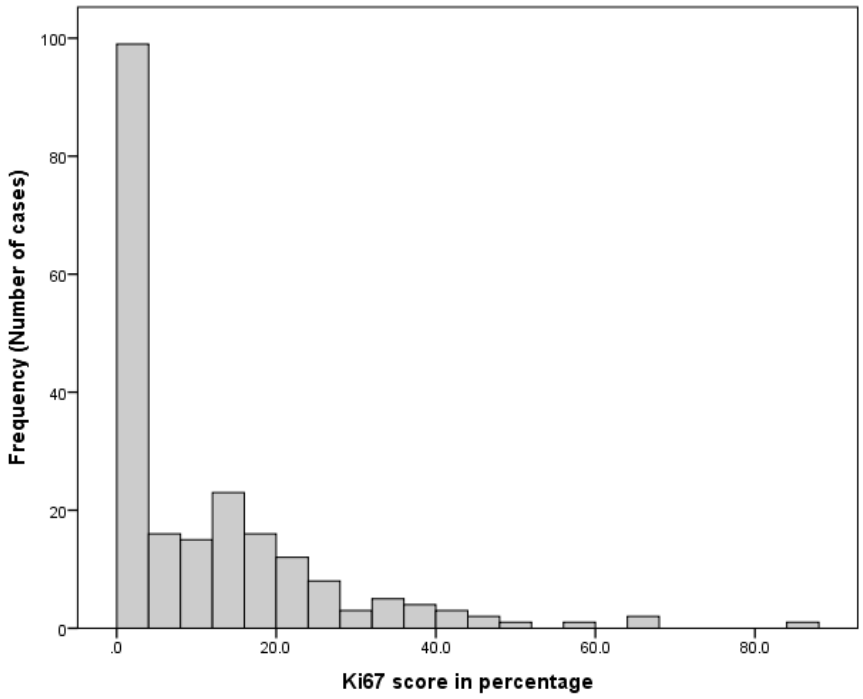


Figure 11: Histogram representing the distribution of Ki7 scores in the MBC cohort

2.4.3.2 Bcl2

Bcl-2 (B-cell lymphoma 2) is the founding member of the family of proteins that regulate cellular death (apoptosis). In humans, it is encoded by the BCL2 gene, and it induces or inhibits apoptosis. The cytoplasmic membranous staining in the cancer epithelial cells was observed for Bcl2 (Figure 10c). A clear cytoplasmic membranous staining with a cut off of 10% was used for determining Bcl2 expression (Callagy et al., 2006). The Bcl2 IHC was performed only in 161 cases and couldn't be performed in the remaining 77 cases due to exhaustion of TMA cores. Bcl2 was positive in most of the MBC patients (152/161; 94.4%) and negative only in a small minority (9/161; 5.6%).

2.4.3.3 p53

The percentage of cancer epithelial cell nuclei with unequivocal nuclear staining was determined for p53 immunoreactivity. A score of > 10% of nuclear staining was taken as the cut off for positive p53 expression (Plesan et al., 2010). The nuclear staining was specific for p53 expression in the MBC tissues (Figure 10d). Only in 42/214 cases (19.6%) p53 was positive and in the rest p53 was negative (172/214; 80.4%).

2.4.3.4 Survivin

The nuclear and cytoplasmic expression of survivin has been observed in breast cancer epithelial cells (Brennan et al., 2008, Al-Joudi et al., 2007, Younis et al., 2009). The initial attempts to optimise the survivin antibody (mouse monoclonal antibody – clone 12C4) were unsuccessful using envision, SAB and proteinase K methods. The nuclear staining was absent with envision method, whereas only patchy nuclear staining with background cytoplasmic staining was

observed with SAB and proteinase K methods. Hence a different clone (D8) of mouse monoclonal antibody was used in the MBC cohort, which has been previously validated in breast cancer tissue (Rexhepaj et al., 2010, Brennan et al., 2008). This antibody gave strong nuclear staining with background cytoplasmic staining which varied from weak to strong in the MBC cohort (Figure 12a – c). Both nuclear and cytoplasmic staining was considered for survivin. Nuclear staining was scored as the percentage of cancer epithelial cells with unequivocal nuclear staining (0 = < 5%, 1 = 5 to 20%, 2 = 21 to 50%, 3 = 51 to 75% and 4 = > 75%) and a cut off of 5% was used as previously reported (Al-Joudi et al., 2007, Hinnis et al., 2007, Tanaka et al., 2000). Whereas, the intensity (0 = no staining, 1 = weak, 2 = moderate, 3 = strong) and percentage (0 = < 5%, 1 = 5 to 20%, 2 = 21 to 50%, 3 = 51 to 75% and 4 = > 75%) of cytoplasmic staining was determined. The individual scores for intensity and percentage of positive cancer epithelial cells with cytoplasmic staining were multiplied to produce a weighted score. Cases with a weighted score of ≥ 3 were considered as positive for cytoplasmic survivin expression.

The nuclear survivin was positive in 90/211 cases (42.7%) and negative in the rest (121/211; 57.3%). The cytoplasmic survivin was positive in most of the cases (193/211; 91.5%) and was negative only in a few cases (18/211; 8.5%).

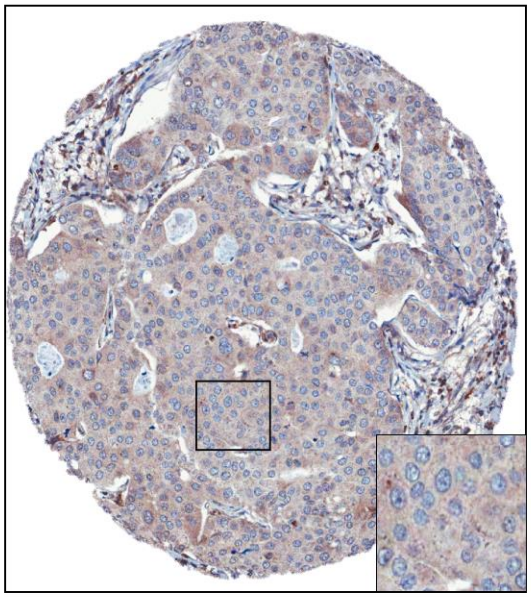


Figure. 12a

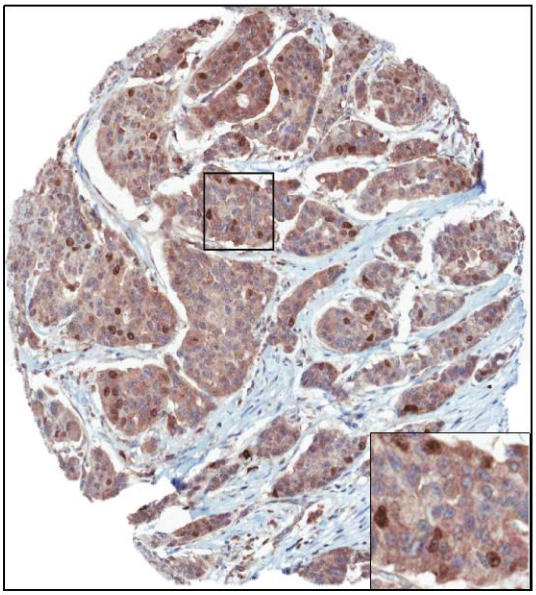


Figure. 12b

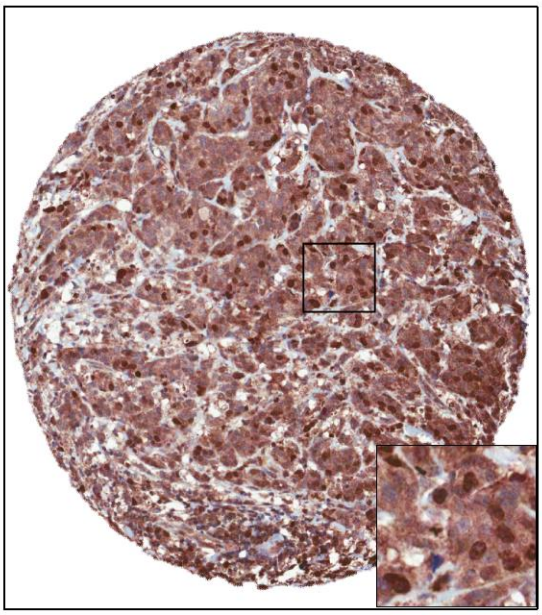


Figure. 12c

Figure 12: Expression of survivin in MBC TMA cores

a) Mild cytoplasmic expression of survivin. b) Moderate cytoplasmic and nuclear expression of survivin. c) Strong cytoplasmic immunostaining and nuclear expression of survivin. Inset - Illustration of 20x magnified image of the marked square depicted in the TMA core

2.4.4 Other biomarkers

2.4.4.1 Human epidermal growth factor receptor 2 (Her2)

The presence of Her2 was detected using the Hercep test™ (Dako). The percentage of cells with unequivocal membranous staining was determined (Figure.13a). Her2 was considered positive, when the score was 3+ on Hercep test™ test and Fluorescent in Situ Hybridisation (FISH) was performed in cases where Hercep test™ test showed a score of 2+ to confirm or refute Her2 amplification (Figure 13b). The FISH test was performed by Mrs Barbara Ozlos, Biomedical Scientist at Leeds Teaching Hospitals NHS Trust. The presence of Her2 was evaluated only 178 cases due to the exhaustion of TMA core. Her2 was negative in most of the cases 172/178 (96.6%) and positive in 6/178 cases (3.4%).

2.4.4.2 E-cadherin

The percentage of membranous staining was determined in each TMA core and a value of > 50% was considered as positive for E-cadherin expression (Callagy et al., 2006). The membranous staining was unequivocal in the MBC TMA cores with minimal background staining (Figure 13c). Most of the cases were positive for E-cadherin expression (161/212; 75.9%) and negative in only 51/212 (24.1%).

2.4.4.3 Forkhead box protein A1 (FOXA1)

The expression of FOXA1 was observed in the nucleus of the cancer cells (Figure 13 d). A semi quantitative method was used for scoring FOXA1

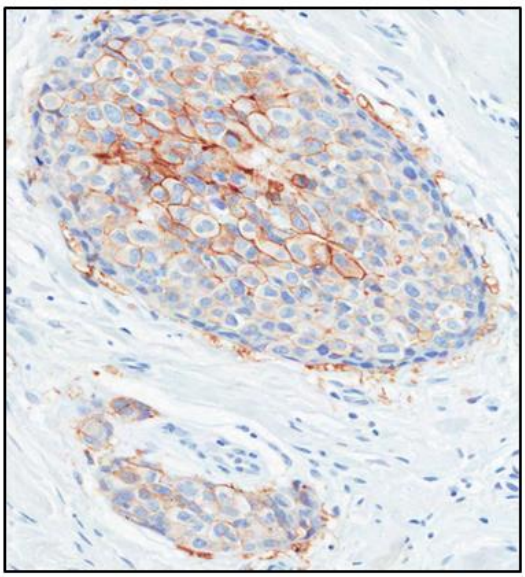


Figure. 13a

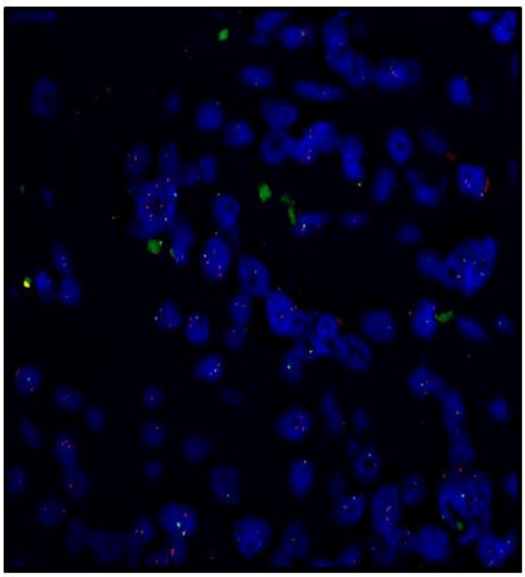


Figure. 13b

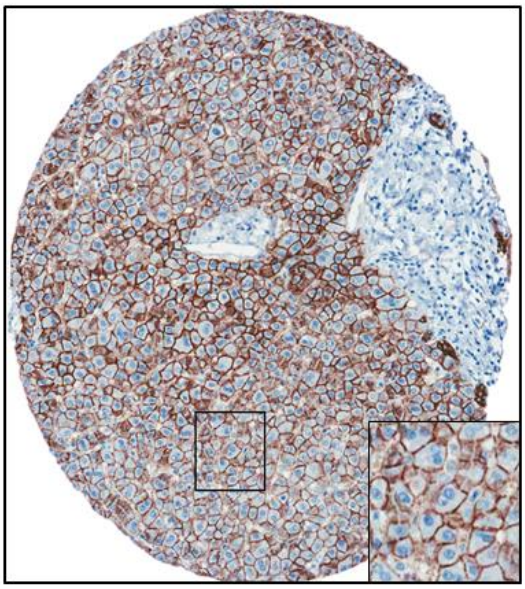


Figure. 13c

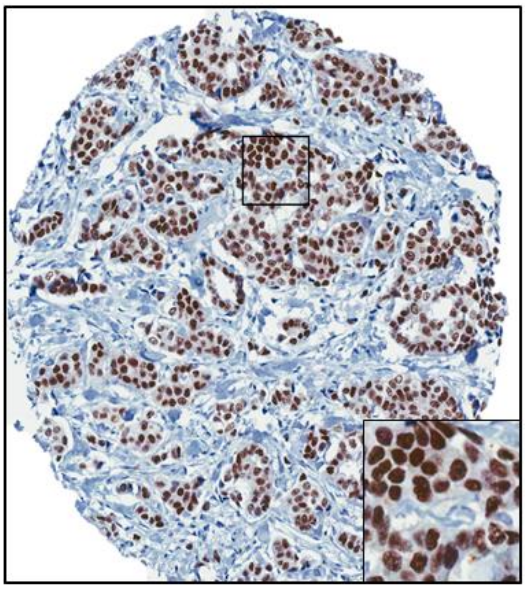


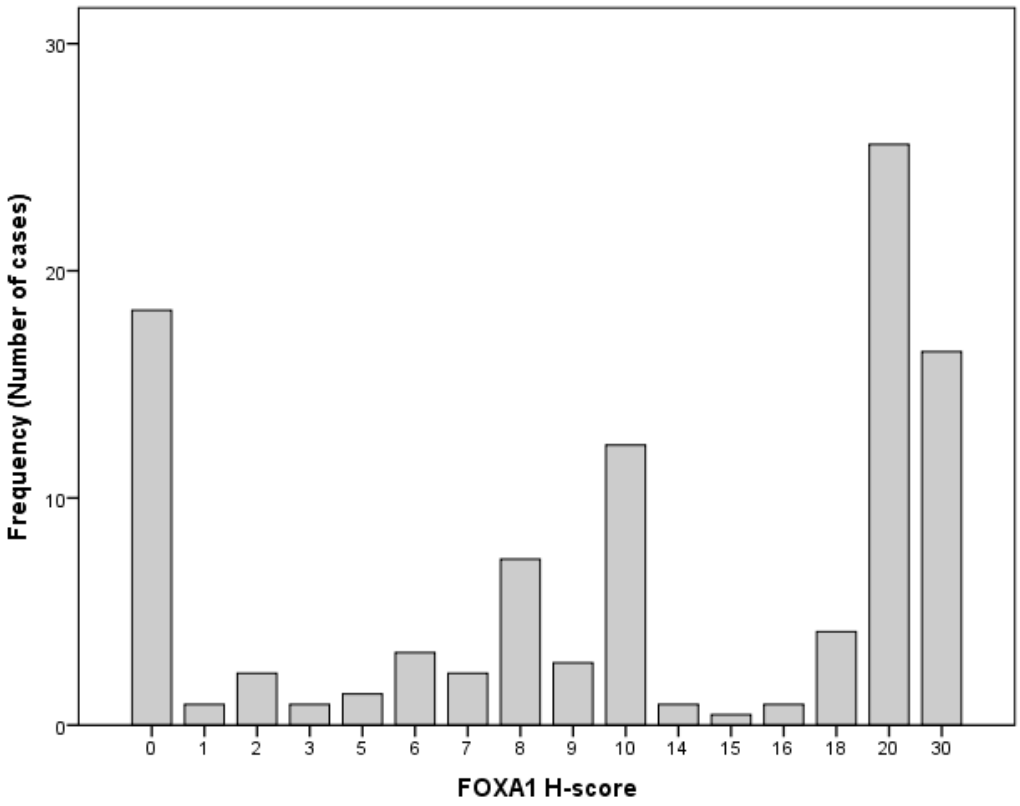
Figure. 13d

Figure 13: Expression of Her2, E-cadherin and FOXA1 in MBC TMA cores

a) Membranous staining of cancer nuclei with Hercep test^{TM}. b) Fluorescent in Situ Hybridization showing Her2 amplification. c) Membranous staining of cancer cells with E-cadherin. d) Unequivocal cancer nuclear expression of FOXA1. Inset - Illustration of 20x magnified image of the marked square depicted in the TMA core

(Badve et al., 2007). The intensity (0 = no staining, 1 = weak, 2 = moderate and 3 = strong) and percentage of nuclear staining was determined in each TMA core (0 = no expression, 1 = 1-10%, 2 = 11-20% and so on to a maximum score of 10 = 91-100%). The scores were then multiplied to give a H-score ranging from 0 to 30. The FOXA1 score was considered as a discrete continuous variable for statistical analysis due to the semi-quantitative nature of the scoring method. The distribution of FOXA1 expression in the MBC cohort is depicted in Figure 14. The H-score was determined in 219 cases and the mean score was 13.69 (SD±10.18).

Figure 14: Bar chart showing distribution of FOXA1 scores in MBC



2.4.5 Measure of association

2.4.5.1 Association between hormonal biomarkers

There was strong positive association of ERα with that of AR ($\chi^2 = 10.65$; p = 0.001) and PR (p = 0.0003, Fishers Exact test). In MBC cohort, there was no statistically significant association between ERα and ER β1 (p =0.57) or ER β2 (p = 0.53), or prolactin (p = 0.25). There was no statistically significant association between ERβ subtypes. Amongst hormonal receptors, only AR was found to have a statistically significant positive association with ERβ1 ($\chi^2 = 7.28$; p = 0.007) and ERβ2 ($\chi^2 = 13.87$; p = 0.0001). There was no association between PR and AR or ERβ sub-types. Prolactin was the only hormonal marker, which didn't have any statistically significant association with other hormonal markers.

2.4.5.2 p53

In our cohort only 20% of the cases were positive and rest were negative. The association of p53 with clinico-pathological variables and other biomarkers were explored to detect any significant result. There was no association between p53 and clinico-pathological variables (Table 4).

There was strong positive association between p53 and Ki67 (p = 0.002) in this cohort. However there was no statistically significant association with other poor prognosticators like grade or nodal status. Moreover, p53 was not shown to have any prognostic or predictive role in the overall survival of MBC patients (Table.5, Page no. 67 and Table.7, Page no. 68).

Variable	Type of test	P value
Grade	Linear-by-Linear association	0.142
Tumour size	Linear-by-Linear association	0.834
Nodal status	Linear-by-Linear association	0.969
PR	Pearson Chi-Square	0.103
ERα	Fishers' Exact	0.741
Bcl2	Fishers' Exact	1
Nuclear survivin	Pearson Chi-Square	0.944
Cytoplasmic survivin	Fishers' Exact	1
Ki67	Pearson Chi-Square	0.002

Table 4: The association of p53 with clinico-pathological variables and biomarkers.

2.4.5.3 ERα and Bcl2

Bcl2 is an inhibitor of apoptosis as well as a target for oestrogen. Hence, the association between these biomarkers were explored in MBC. There was a strong positive association between ERα and Bcl2, (p = 0.034). This was in spite of doing analysis in a smaller sample size due to the exhaustion of TMA core allowing IHC of Bcl2 only in 161 cases. However, Bcl2 was not found to have any prognostic significance in this cohort. Considering the small sample size (n = 161), the lack of power might have prevented in identifying any true significance.

2.4.5.4 Survivin

There is substantial interest in the sub-cellular location of survivin due to its prognostic significance in various cancers (Li et al., 2005). There is only a single study so far in the literature, which explored the role of survivin in MBC without identifying any statistical significance (Younis et al., 2009). In this study, the sub-cellular location of survivin was determined and its relationship with various clinical and pathological markers was explored.

The nuclear survivin expression showed no association with grade (p = 0.49) or nodal status (p = 0.191). However, there was a weak association of nuclear survivin with tumour size categories (p = 0.034). Amongst hormonal biomarkers, only PR was found to have a positive association with nuclear survivin (χ^2 = 5.98; p = 0.014). Even though there was no association between nuclear survivin and grade of the breast cancer, there was a strong positive association with Ki67 expression (χ^2 = 20.78; p = 0.001). Nuclear survivin didn't show any association with Her2 (p = 1), Bcl2 (p = 0.134) and p53 (p = 0.944)

In comparison to nuclear survivin, cytoplasmic survivin was positive in most of the cases (91%). There was no association with any of the clinical or pathological prognostic markers. However, there was a strong positive association with ERα (χ^2 = 20.57; p = 0.001) and PR (χ^2 = 8.41; p = 0.004). Similar to nuclear survivin, there was a statistically significant association between Ki67 and cytoplasmic survivin (χ^2 = 8.68; p = 0.003). In this MBC cohort, neither nuclear nor cytoplasmic survivin was found to have any prognostic or predictive role.

In FBC, the nuclear survivin was shown to be a poor prognosticator and a high cytoplasmic-to-nuclear ratio of survivin was shown to have better prognosis (Brennan et al., 2008). Brennan et al (2008) used automated in-house technology to perform consistent and reproducible survivin scoring. However with manual scoring of TMAs, It was not possible to reliably reproduce cytoplasmic-nuclear ratio in the MBC cohort. Future studies are needed to address the role of cytoplasmic-to-nuclear ratio of survivin in MBC using automated scoring.

2.4.6 Molecular subtypes in male breast cancer

Based on gene expression profile, FBC has been classified in to intrinsic sub-types with distinctive prognostic significance (Hu et al., 2006, Sorlie et al., 2003, Perou et al., 2000). The intrinsic subtypes described by Perou et al (2000) include Luminal A, Luminal B, Her 2 enriched and Basal or triple negative sub-types. These tumour sub-types correlated well with survival and metastasis free outcome with Her2 and Basal types having the worse prognosis and Luminal A having the best overall outcome (Sorlie et al., 2003). It has been since then possible to use immunohistochemical biomarkers as surrogate for classifying FBC. Carey et al (2006) successfully managed to classify and replicate the prognostic significance of intrinsic FBC sub-types using an immunohistochemical panel consisting of ER, PR, Her2, CK5/6, CK14 and epidermal growth factor receptor (EGFR). In order to facilitate more accurate clinical classification of luminal B breast cancers, it has been proposed since then to include Ki67 to the immmunohistochemical panel (Cheang et al., 2009).

In this cohort using surrogate biomarkers, MBC was classified in to the known 4 FBC intrinsic sub-types. These include, Luminal A (ERα+ and/or PR+, Her2 & Ki67-), Luminal B (ERα+ and/or PR+, Her2-, Ki67+ or ERα+ and/or PR+, Her2 +), Her 2 enriched (ERα -, PR – and Her2 +) and Basal or triple negative breast cancer (ERα -, PR - & Her2 -). As described by Cheang et al (2009), a score of ≥ 14% was considered as the ideal cut off for determining Ki67 positivity in this cohort. Amongst 238 cases available for analysis in the MBC cohort, molecular sub-types were determinable in 204 cases. Most of the cases were Luminal A (n=135; 66.2%), followed by Luminal B (n=62; 30.4%) and Basal or triple

negative sub-type (n=7; 3.4%). There was no Her2 enriched sub-type in the MBC cohort.

2.4.7 Regression analysis

A logistic regression analysis was performed to determine predictors of survival in the MBC cohort. Cox regression analysis was performed using various clinico-pathological characteristics and biomarkers as explanatory variables and overall survival (OS) as outcome variable. The analysis was performed separately for 5 and 10 year OS.

2.4.7.1 5 year OS

On univariate analysis, grade, age and CK19 expression was found to be statistically significant (Table 5). Compared to grade 1 MBC cases, those with grade 2 were shown to have poor OS (OR = 0.368; p = 0.024 (CI = 0.15 to 0.88). However, there was only a trend towards better survival amongst grade 1 MBC cases compared to grade 3 (OR =0.756; p = 0.507 (CI = 0.33 to 1.73). This might have occurred due to the relatively small number of cases with grade 1 and grade 3 cancers in this cohort compared to grade 2 MBC. There was a trend towards statistical significance with nodal status (p = 0.089), CK18 (p = 0.068) and Ki67 (p = 0.078). Variables with statistical significance or trend towards significance were entered in to the multivariate analysis model. The results showed that only age and nodal status were retained in the model (Table 6) and found to be independent predictors of survival.

2.4.7.2 10 year OS

Consistent with 5 year OS results, both age and CK19 was found to be statistically significant on univariate analysis (Table 7). The results showed loss of survival predictability for nodal status in the MBC cohort with longer follow-up.

Hence the observed 5 year OS predictive role for nodal status could have been due to chance or not having sufficient power (due to sample size) to detect a true effect. The model for multivariate analysis was created using variables that were shown to have significance or trend towards significance in the 10 year OS analysis. The results showed only age to be an independent predictor for MBC OS (OR = 1.065 (CI – 1.04 to 1.1); p = < 0.001).

Variable	HR (CI)	p value
Age	**1.073 (1.04 - 1.11)**	**0.001**
Grade 1		**0.04**
Grade 2	**0.368 (0.15 - 0.88)**	**0.02**
Grade 3	0.756 (0.33 - 1.73)	0.5*
Nodal status	1.885 (0.91 – 3.91)	0.09*
ERα	3.002 (0.41 – 21.89)	0.278
ER β1	0.600 (0.30 – 1.19)	0.145
Erβ2	0.817 (0.38 – 1.75)	0.602
ER β5	1.225 (0.42 – 3.59)	0.711
PR	1.210 (0.47 – 3.11)	0.692
AR	0.648 (0.34 - 1.25)	0.194
Prolactin	0.711 (0.33 – 1.52)	0.378
CK5/6	1.119 (0.40 – 3.16)	0.832
CK14	2.148 (0.29 – 15.68)	0.451
CK19	**0.263 (0.10 – 0.68)**	**0.006**
CK18	0.334 (0.10 - 1.09)	0.068*
*Ki67	1.825 (0.93 - 3.57)	0.078*
Bcl2	1.819 (0.25 – 13.39)	0.557
p53	1.037 (0.40 – 2.69)	0.941
Nuclear survivin	1.634 (0.85 – 3.15)	0.143
Cytoplasmic survivin	0.743 (0.26 – 2.10)	0.575
Her2	3.316 (0.45 – 24.49)	0.240
E-cadherin	1.326 (0.55 – 3.20)	0.531
FOXA1	0.995 (0.96 – 1.03)	0.743

Table 5: Cox univariate regression analysis result for 5 year overall survival in MBC cohort

The variables in bold were found to be significant. * represents variables that were found to have a trend towards significance and entered in to the multivariate modelling.

Variable	Multivariate analysis - 5 year OS	
	OR (CI)	p value
Age	1.072 (1.03 - 1.12)	0.001
Nodal status	2.369 (1.05 - 5.36)	0.04

Table 6: Cox multivariate regression analysis result for 5 year overall survival in MBC cohort

Variable	OR (CI)	p value
Age	**1.061 (1.035 – 1.089)**	**0.0001**
Grade 1		0.113
Grade 2	0.537 (0.27 – 1.08)	0.080*
Grade 3	0.630 (0.42 – 1.69)	0.630
Nodal status	1.357 (0.78 – 2.35)	0.277
ERα	1.861 (0.58 – 5.93)	0.294
ER β1	0.740 (0.43 – 1.27)	0.276
ERβ2	0.794 (0.44 – 1.42)	0.438
ER β5	0.816 90.39 – 1.72)	0.593
PR	1.044 (0.54 – 2.00)	0.897
AR	0.891 (0.54 – 1.47)	0.650
Prolactin	0.764 (0.42 – 1.37)	0.370
CK5/6	0.582 (0.21 – 1.60)	0.295
CK14	1.026 (0.14 – 7.41)	0.980
CK19	**0.284 (0.12 – 0.67)**	**0.004**
CK18	0.675 (0.21 – 2.15)	0.506
Ki67	1.312 (0.76 – 2.26)	0.330
Bcl2	1.794 (0.44 – 7.37)	0.418
p53	0.839 (0.40 – 1.77)	0.646
Nuclear survivin	1.179 (0.71 – 1.96)	0.524
Cytoplasmic survivin	1.137 (0.45 – 2.84)	0.783
Her2	1.667 (0.23 – 12.13)	0.614
E-cadherin	1.639 (0.81 – 3.33)	0.171
FOXA1	1.005 (0.98 – 1.03)	0.711

Table 7: Cox univariate regression analysis result for 10 year overall survival in MBC cohort

The variables in bold were found to be significant. * represents variables that were found to have a trend towards significance and entered in to the multivariate modelling.

2.4.8 Survival analysis

The data for OS was available for this MBC cohort. The Cox proportional hazard model was used for performing survival analysis and Kaplan-Meier survival plots were produced separately for 5 and 10 year OS. The 5 year OS analysis showed CK18 ($p = 0.05$) and CK19 ($p = 0.003$) to be statistically significant (Figure 15a & b). There was a trend towards statistical significance with Ki67 ($p = 0.07$) and nodal status ($p = 0.082$) (Figure 16a & b). The results also showed statistically significant survival advantage for grade 2 MBC over other grades ($p = 0.034$). This could have occurred as a result of small number of grade 1 cases in the MBC cohort ($n = 29$), leading to lack of power to determine any statistically significant outcome. This was further affirmed, when 10 year OS analysis was performed and only CK19 was shown to be statistically significant ($p = 0.002$) (Figure 17).

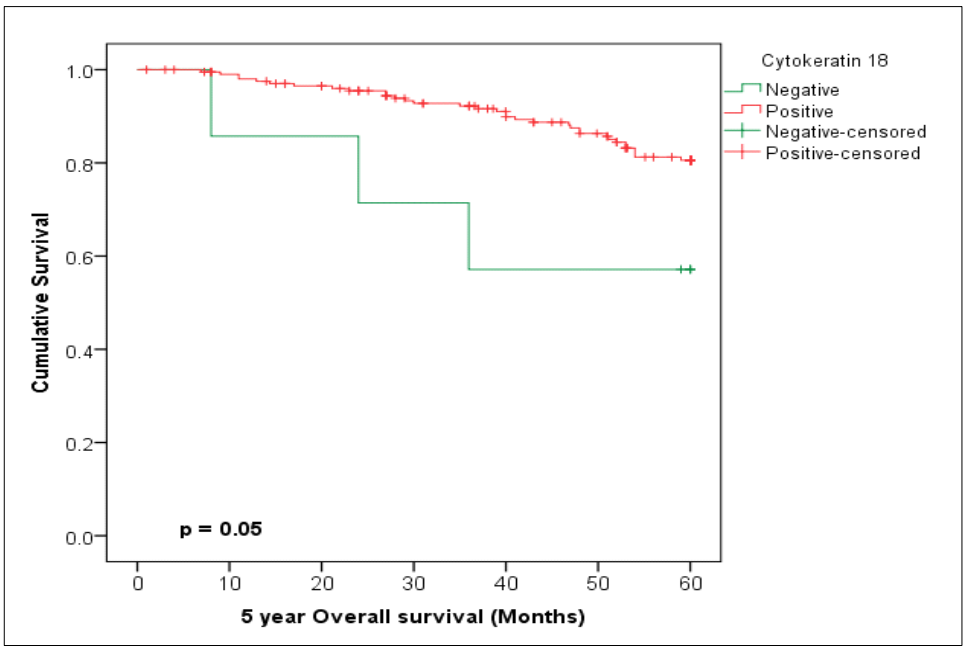


Figure. 15a

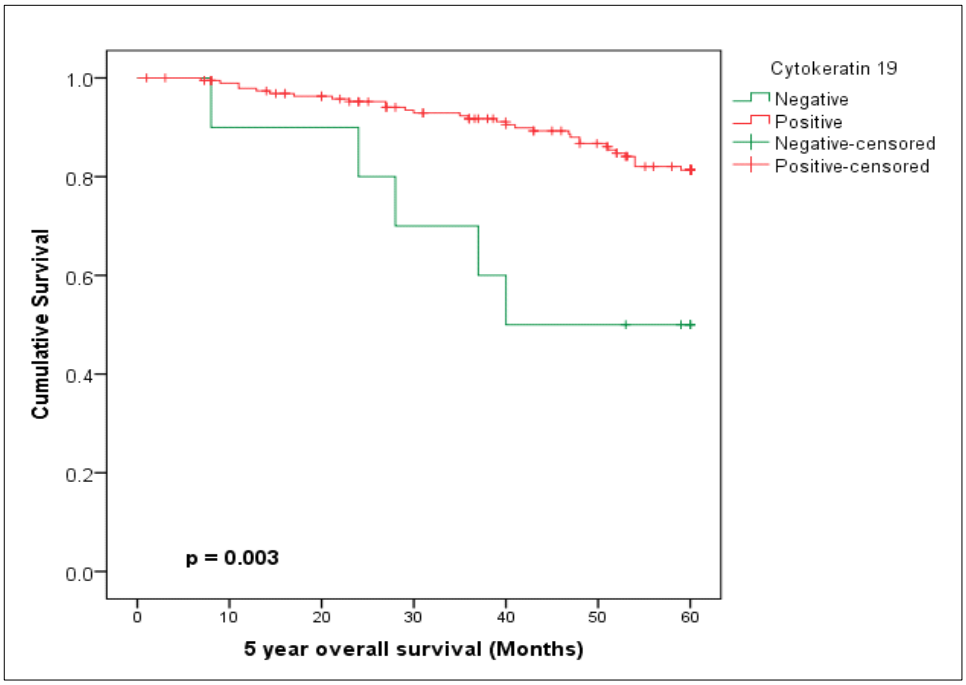


Figure. 15b

Figure 15: Kaplan-Meier survival curve at 5 years for CK18 and CK19

Statistically significant poor OS at 5 years for (a) CK18 and (b) CK19 negative male breast cancers. Patients were censored when they ceased to be followed up for any reason but had not died due to breast cancer.

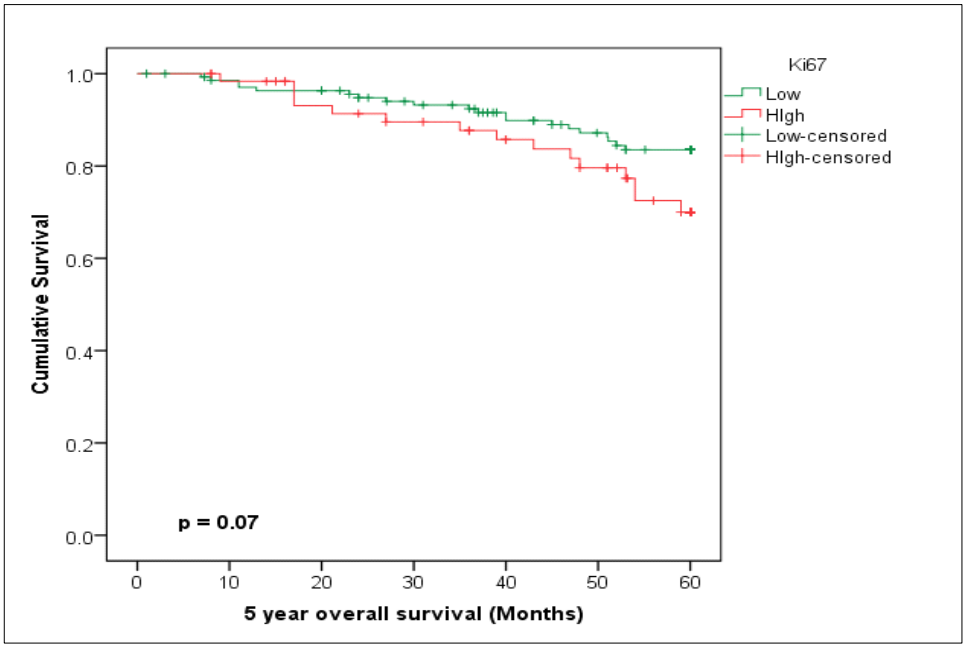


Figure. 16a

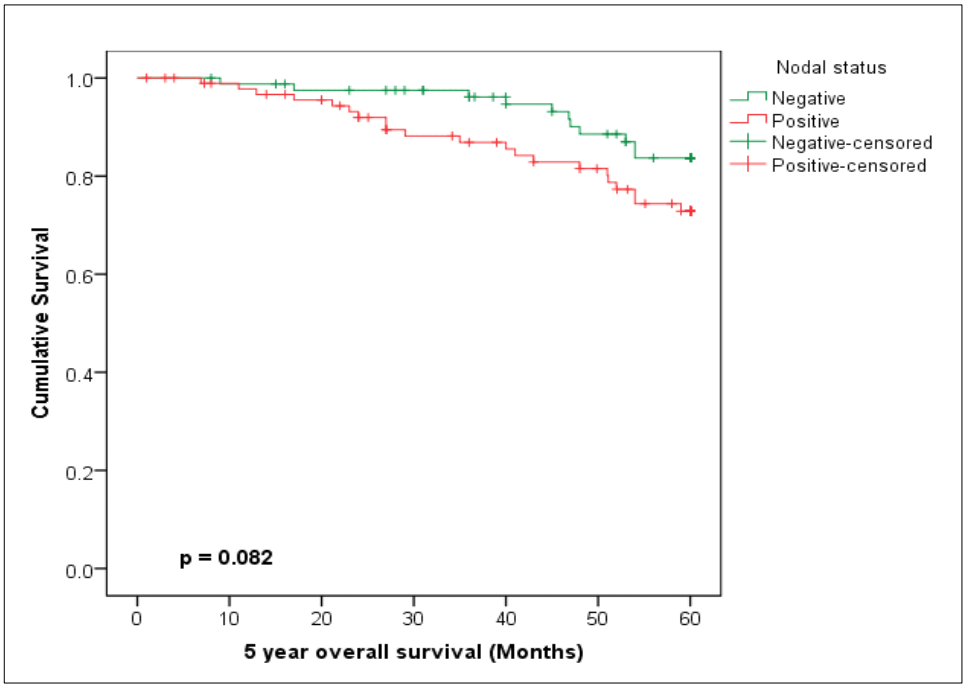


Figure. 16b

Figure 16: Kaplan-Meier survival curve at 5 years for Ki67 and nodal status

A trend towards statistical significance was observed for (a) Ki67 and (b) nodal status at 5 years in the male breast cancer cohort. Patients were censored when they ceased to be followed up for any reason but had not died due to breast cancer.

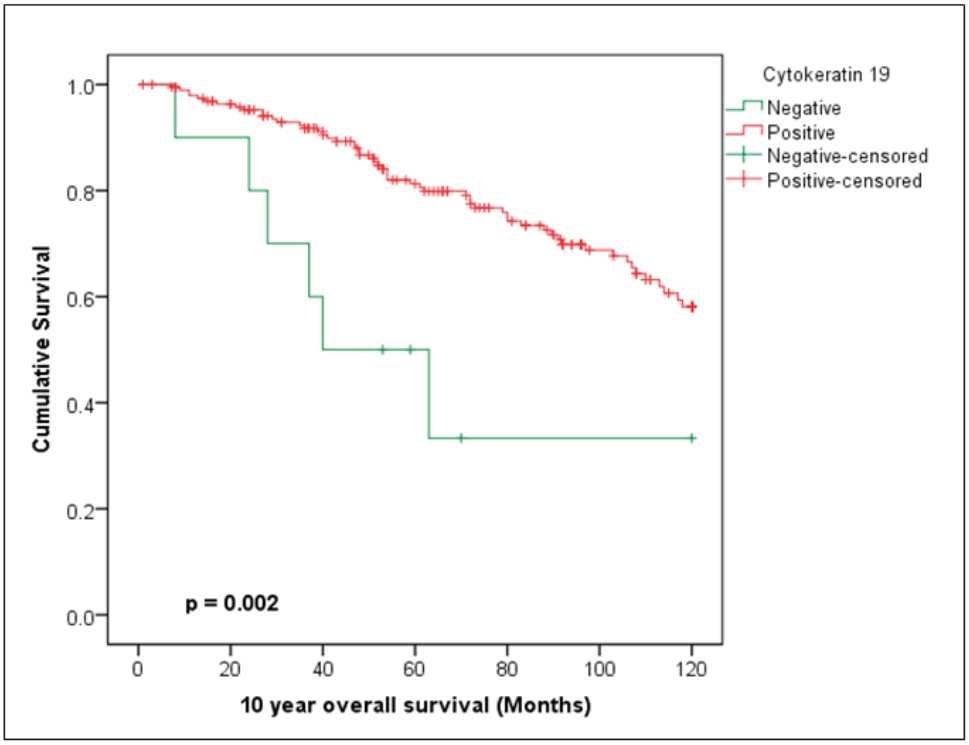


Figure. 17

Figure 17: Kaplan-Meier survival curve at 10 years for CK19

Survival curve showing statistically significant poor 10 year overall survival for CK19 negative male breast cancer patients. Patients were censored when they ceased to be followed up for any reason but had not died due to breast cancer.

2.5 DISCUSSION

Even though there has been an increase in the incidence of MBC, it remains a rare cancer compared to FBC (CRUK, 2012). The low incidence of the disease results in a paucity of research compared to FBC. The current management of MBC is mostly transcribed and extrapolated from the information that is available from research undertaken in FBC. An attempt was made here to understand the biology of MBC using immunohistochemical analysis of various biomarkers.

A high expression of ERα and PR was seen in this MBC cohort (93.8% and 84.2% respectively). Similar higher expression of steroid receptors in MBC compared to FBC has been observed by other researchers (Contractor et al., 2008, Shaaban et al., 2012). It has been hypothesised that, an up-regulation of the steroid receptors in an oestrogen depleted environment may be responsible for this (Muir et al., 2003). However, in comparison to FBC, ER α was not found to have any prognostic significance in MBC. This was further substantiated by a recent study combining data from copy number variation and gene expression demonstrating that even though most MBC are ER positive, they share features of ER negative FBC (Johansson et al., 2013). The different hormonal milieu in which male and female breast cancer develop could be another factor influencing gender specific prognostic significance of ERα. Finally, the high positive expression of ERα in MBC necessitates the need for large studies with sufficient power to detect any true prognostic effect.

The functional studies in MBC have shown a greater role for AR compared to ER (Weber-Chappuis et al., 1996b). However, AR expression in MBC has been variably reported within the literature ranging from 0 to 95% (Contractor et al., 2008). Around 55% of the cases were AR positive in this MBC cohort, but did not add to prognostic significance. This finding further confirms the lack of a definitive prognostic role for hormonal receptors in MBC compared to FBC based on the current published literature.

It has been shown that, in an oestrogen depleted environment, such as in post-menopausal women or MBC, up-regulation of ER can lead to an increased response to oestrogen targets such as Bcl-2 (Nahleh and Girnius, 2006). Once up-regulated, Bcl-2 can lead to malignant transformation through genomic modification and downstream target activation (Muir et al., 2003, Rayson et al., 1998). Even though Bcl-2 was shown to be over expressed in MBC compared to FBC (Muir et al., 2003, Rayson et al., 1998, Weber-Chappuis et al., 1996b), its prognostic significance remains uncertain. In this study, 94% of the cases were Bcl2 positive and there was a strong positive association between ER α and Bcl2 ($p = 0.03$). However, regression analysis did not reveal any statistically significant prognostic role for Bcl2 in this cohort of MBC (Tables 5 and 7, Page 67 and 68 respectively).

The high throughput gene expression studies have classified FBC into distinctive intrinsic sub-types with prognostic significance (Hu et al., 2006, Sorlie et al., 2003, Perou et al., 2000). It was possible to represent these intrinsic sub-types with distinct prognostic significance using surrogate biomarkers immunohistochemically in FBC (Carey et al., 2006, Nielsen et al., 2004).

Therefore an attempt was made to evaluate the established FBC sub-types in this MBC cohort. The results showed a distinctive pattern with most MBC being either luminal A or luminal B subtype with no Her2 enriched subtype. Similar findings were observed by other researchers evaluating FBC subtypes in MBC (Abreu et al., 2016, Shaaban et al., 2012, Kornegoor et al., 2012). Recent evidence through high throughput gene expression profiling has confirmed the existence of unique MBC intrinsic subtypes, luminal M1 and M2 (Johansson et al., 2012, Johansson et al., 2011). The male subtypes did not cluster along with the known FBC subtypes. However unlike FBC, it has not been yet possible to represent these intrinsic subtypes immunohistochemically in MBC. The above evidence further substantiates that, MBC should be considered as a different molecular entity compared to FBC.

The expression of various cytokeratins depends on their differentiation with CK5/6 & CK14 representing the basal/myoepithelial cells and CK18/CK19 that of luminal cells (Ciocca et al., 2006). The basal cytokeratin expression in MBC was quite low, with CK5/6 positive in 14% of cases and CK14 in only 1.8% cases. Conversely, most of the MBC cases (around 96% for CK18 and CK19) were positive for luminal cytokeratins. Similar high expression of luminal cytokeratins was shown in a smaller MBC cohort (n = 32) (Ciocca et al., 2006). Even in FBC, the expression of CK19 was shown to be in the range of 89 - 94% (Shao et al., 2012, Delgallo et al., 2010). CK19 negativity has been correlated with poor prognostic factors like ER/PR negativity and positive Ki67 expression in FBC (Fujisue et al., 2012). CK19 negativity was also found to be an independent predictor of poor overall survival and local recurrence in young women (Parikh et al., 2008). A similar finding was observed in this cohort of

MBC patients with CK19 being a predictor of OS on univariate analysis (p = 0.006). Similarly, both CK18 and CK19 positive patients were found to have a statistically significant better OS (Results section: Figure 15a & 15b, Page no. 70 and Figure 17, Page no. 72).

In this cohort, age seems to be the single most statistically significant predictor of OS. However, its clinical significance could be of uncertain nature for the following reasons. The OS could have been influenced by various factors in a cohort with median age of 68 years. The varying length of patient follow-up due to the data originating from multiple centres around the world could have influenced the results. Finally, the retrospective nature of the study could have introduced selection and analytical biases influencing the results. However, it should be emphasized that, within these limitations, this series has got the second largest sample of MBC patients with adequate follow-up data within the published literature.

The preliminary results of the International Male Breast Cancer Program which has got the largest number of MBC patients to date was recently presented at the San Antonio Breast Cancer Symposium (Cardoso, 2014). Amongst 1483 eligible MBC patients with a median age of 68.4 years, ERα was positive in 99%, PR positive in 81% and 97% were AR positive. Her2 was positive in 9% of the MBC patients, compared to 3.4% of the patients in this cohort. Ki67 was positive in 39% of the patients and was comparable to the expression levels noted in this cohort (31.2%). The researchers didn't identify any prognostic role for grade, Her2, AR, Ki67 and FBC surrogate subtypes in MBC patients. They showed better OS for MBC patients with high expression of ERα and PR (Allred

score of 7 or 8) compared to low expression (Allred score 3 to 6). However due to low percentage of ERα (1%) and PR (19%) negativity they could not test the OS in this group compared to Allred positive (score 3 to 8) group of MBC.

The nodal status was found to be an independent predictor of OS at 5 years. However, the predictability was lost at 10 year OS analysis, indicating that it is not a reliable variable to predict long term survival in MBC. Amongst all biomarkers only CK19 was found to be an independent predictor for 5 year OS. However, its independent predictability was lost when the analysis was performed at 10 years. The effect of attrition due to longer follow-up, small number of CK19 negative cases, the sample size and power of the study were some of the factors that might have influenced the results.

Survival analysis performed using Cox proportional hazard model showed statistically better 5 year OS for CK18 and CK19 positive MBC cases (Result section: Figure 15a & 15b, Page no. 70 and Figure 17, Page no. 72). However, only CK19 was shown to be statistically significant at 10 year follow-up. There was also a trend towards statistical significance at 5 year OS analysis with Ki67 ($p = 0.07$) and nodal status ($p = 0.082$). Even though there was no OS benefit identified for Ki67, there was poor relapse free survival amongst node positive patients in the International Male Breast Cancer Program (Cardoso, 2014).

2.5.1 Limitations of the study

Before summarizing and interpreting the findings, it is important to consider some of the limitations of this study. The retrospective nature of the study may have introduced biases. However considering the low incidence and prevalence of MBC, other methodologies would have been time consuming and impractical. A criticism is the missing clinical, pathological and follow-up data in approximately 40% of the patients reducing the power of the study. The lack of information about the surgical and adjuvant treatment given along with variation in the treatment modalities by virtue of the diverse sample should have confounded the outcome variable. However, the diverse nature of origin of the MBC tissue samples from various geographical areas around the world does indeed increases the external validity of the results.

The primary outcome variable was OS, due to the lack of data on breast cancer specific survival (BCSS). The latter should have improved the internal validity and reproducibility of the findings. The findings of this study should be considered along with the advantages and disadvantages of TMA as a tool for histopathological evaluation in large number cancer cases. The benefits and disadvantages of tissue microarray as a tool for histopathological evaluation have been described extensively in the literature (Ilyas et al., 2013). In this cohort, TMA cores were missing only in 19 cases, which was acceptable considering the diverse nature of this cohort. The multitude of biomarkers evaluated in this study resulted in TMA exhaustion, which affected the evaluation of some of the biomarkers more than others.

2.6 CONCLUSION

The study showed potential prognostic role for epithelial cytokeratins, especially for CK19 and to some extent for CK18 in MBC patients. However age was shown to be the single most independent predictor for survival in the MBC cohort. The study confirmed that the FBC molecular sub-types represented using surrogate biomarkers are expressed differently in MBC. Moreover the expression profile and role of biomarkers with known prognostic significance in FBC was shown to be different in this MBC cohort. Overall the findings of the study suggest inherent differences amongst male and female breast cancer at protein level. However, it also emphasizes the need for future studies with larger sample size and adequate follow-up. This will help in the identification of clinically relevant protein biomarkers that can be used in the management of MBC.

CHAPTER 3

External Validation of ImmunoRatio™ Image Analysis Application for ERα and Ki67 Determination in Breast Cancer

3.1 INTRODUCTION

Automated IHC image analysis systems were introduced in the early 1990's with the intention to produce consistent, accurate and reproducible results without potential human errors associated with manual microscopic analysis (Kirkegaard et al., 2006). However it is only in the last few years that software has been developed for the assessment of high-throughput automated assessment of the histological slides. The majority of automated IHC analysis software's are only available commercially and therefore its academic or clinical availability is limited to a few centres. Automated IHC image analysis systems are either fully or semi-automated but usually require input from the pathologist to delineate the areas to be evaluated.

A fully or semi-automated IHC image analysis systems could be effective when scoring large cohort of histological samples. An automated image analysis system can reduce the time required to score more than 1000 TMA cores (428 cases in triplicate) as in this MBC cohort. Another advantage is the consistency (by reducing inter and intra-observer variability) with which IHC scoring can be performed in a large cohort cases. However there is no single platform that can be utilised for scoring a range of nuclear, cytoplasmic and membrane biomarkers as evaluated in this MBC cohort. The availability of a common platform should have also helped in calculating the cytoplasmic-to-nuclear ratio of survivin (Brennan et al., 2008) in the MBC cohort.

ImmunoRatio™ is a web based freely available fully automated IHC image analysis software developed for the evaluation of nuclear proteins, ERα, PR and

Ki67 (Tuominen et al., 2010). Even though the free web based platform improves the accessibility, similar to other fully automated platforms, the role of ImmunoRatio™ is limited to the analysis of IHC in tissue microarrays. In this study it was decided to evaluate the automated scoring of ERα and Ki67 using ImmunoRatio™ software in breast cancer TMAs. ERα was selected as a paradigm as it provides unequivocal nuclear staining with antibodies in clinical use and has got a well-established scoring method that has been in use over many years. Moreover, ERα is predictive for response to adjuvant hormonal treatment (Osborne, 1998) and around 70% of breast tumours are ERα positive (Murphy and Watson, 2002).

The ability of cancer to proliferate uncontrollably (Hanahan and Weinberg, 2011) can be evaluated immunohistochemically by measuring the expression of proliferative biomarkers. Antigen KI-67 (Ki67) is a nuclear protein that in humans is encoded by the MKI67gene and is the most well established proliferative marker in breast cancer (Dowsett et al., 2011). Ki67 is currently considered as the assay of choice for measuring tumour proliferation in clinical trials (Dowsett et al., 2011). However, Ki67 has proven to be a more challenging nuclear biomarker to analyse using automated systems (Mohammed et al 2012; Fasanella et al 2011). Hence, Ki67 was chosen as the second biomarker to be evaluated using the ImmunoRatio™ software in this study.

3.2 AIM

The aim was to externally validate ImmunoRatio™ image analysis software in the determination of ERα and Ki67 expression in a pre-stained series of breast cancer TMAs. The ERα expression was evaluated in consecutive 100 cases of male and female breast cancer TMAs and Ki67 in 64 cases of male breast cancer TMAs.

The objectives were to:

1) Determine the agreement between manual and ImmunoRatio™ scoring
2) Establish the practical applicability of ImmunoRatio™ software for scoring nuclear biomarkers

(Work presented in this chapter has been published and details can be found in appendix 7, Page no. 188)

3.3 METHODOLOGY

The TMAs that were pre-stained with ERα and Ki67 antibodies were selected from the Leeds Breast Research Group tissue archived at the Leeds Institute of Cancer and Pathology. Consecutive male (n=100) and female (n=100) breast cancer TMAs without any artefacts or missing cores were used for the evaluation of ERα. For the evaluation of Ki67, only MBC TMAs was used as per the recommendation of the International Ki67 in Breast Cancer Working Group (Dowsett et al., 2011). This was due to the following reasons: It is well established that non-nuclear and multi-compartmental IHC staining can reduce the sensitivity of automated image analysis systems (Bolton et al., 2010). The SP6 antibody used for the evaluation of Ki67 in the FBC produced moderate to strong background cytoplasmic staining compared to the MIB1 antibody used in the MBC cohort. Hence from a cohort of 428 MBC cases, only 64 cases with unequivocal nuclear staining with weak or no background cytoplasmic staining were selected (Figure 18). Ethical approval for the use of male and female breast cancer tissues were obtained from the Leeds West (06/Q1205/156) and East (06/Q1206/180) Research Ethics Committees respectively. These TMA cores were made in triplicate from formalin fixed paraffin embedded cancer blocks as described previously (Hamilton-Burke et al., 2010, Shaaban et al., 2012) in the methodology section (Chapter 2, Methodology section 2.3, Page no. 27-31).

3.3.1 Immunohistochemistry

Immunohistochemistry for ERα and Ki67 was performed using standard published protocols as previously described (Chapter 2, Methodology section 2.3, Page no.32-36). The stained slides were scanned at x 20 objective magnification using Aperio

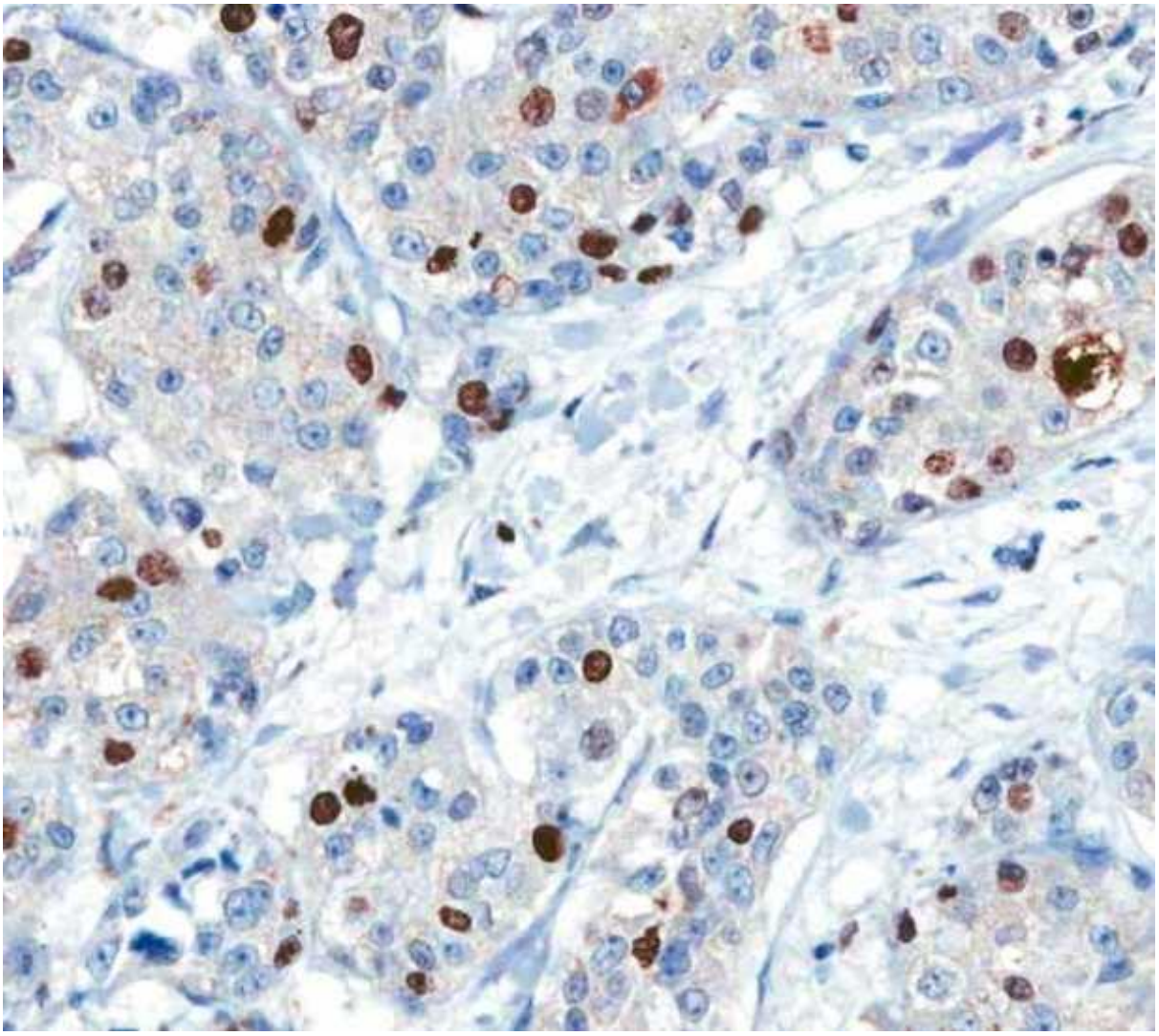


Figure 18: Ki67 immunohistochemistry staining with MIB1 antibody

Ki67 IHC using MIB1 antibody showing unequivocal nuclear staining with weak background cytoplasmic staining (20x magnified image)

ScanScope™ and then visualised for manual scoring using the ImageScope™ viewing software. The manual scoring was performed under guidance of a consultant breast histopathologist (AMH). The Allred method (Allred et al., 1998) was used for ERα manual scoring. In each TMA core, all representative cancer nuclei were counted and the proportion of unequivocally DAB stained nuclei were estimated for Ki67 manual scoring (Dowsett et al., 2011).

3.3.2 ImmunoRatio™ software

The ImmunoRatio™ software calculated the percentage of DAB stained nuclei area over the total nuclear area for each biomarker and represented it as Labelling Index (LI) (Tuominen et al., 2010). The software uses the uploaded microscope image, optional blankfield correction and threshold adjustment parameters entered into the website for determining the LI. The blank field image or an in-built algorithm (when it was not available as in the basic mode) was used for correcting illumination and colour balance. The corrected image was then separated into DAB and haematoxylin stained components using the colour deconvolution method. Subsequently each of the component area was subjected to threshold adjustments using the parameters entered in to the website (advanced mode) or using adaptive IsoData thresholding method (basic mode). The nucleus segmentation was then performed in the components to delineate the nucleus from other cell types. Once the nuclear area was defined within the DAB and haematoxylin stained components, the images were overlaid on each other to calculate the percentage of DAB stained area. The final output displayed both the input image and the colour de-convoluted image produced by the software to calculate the labelling index (Figure 19).

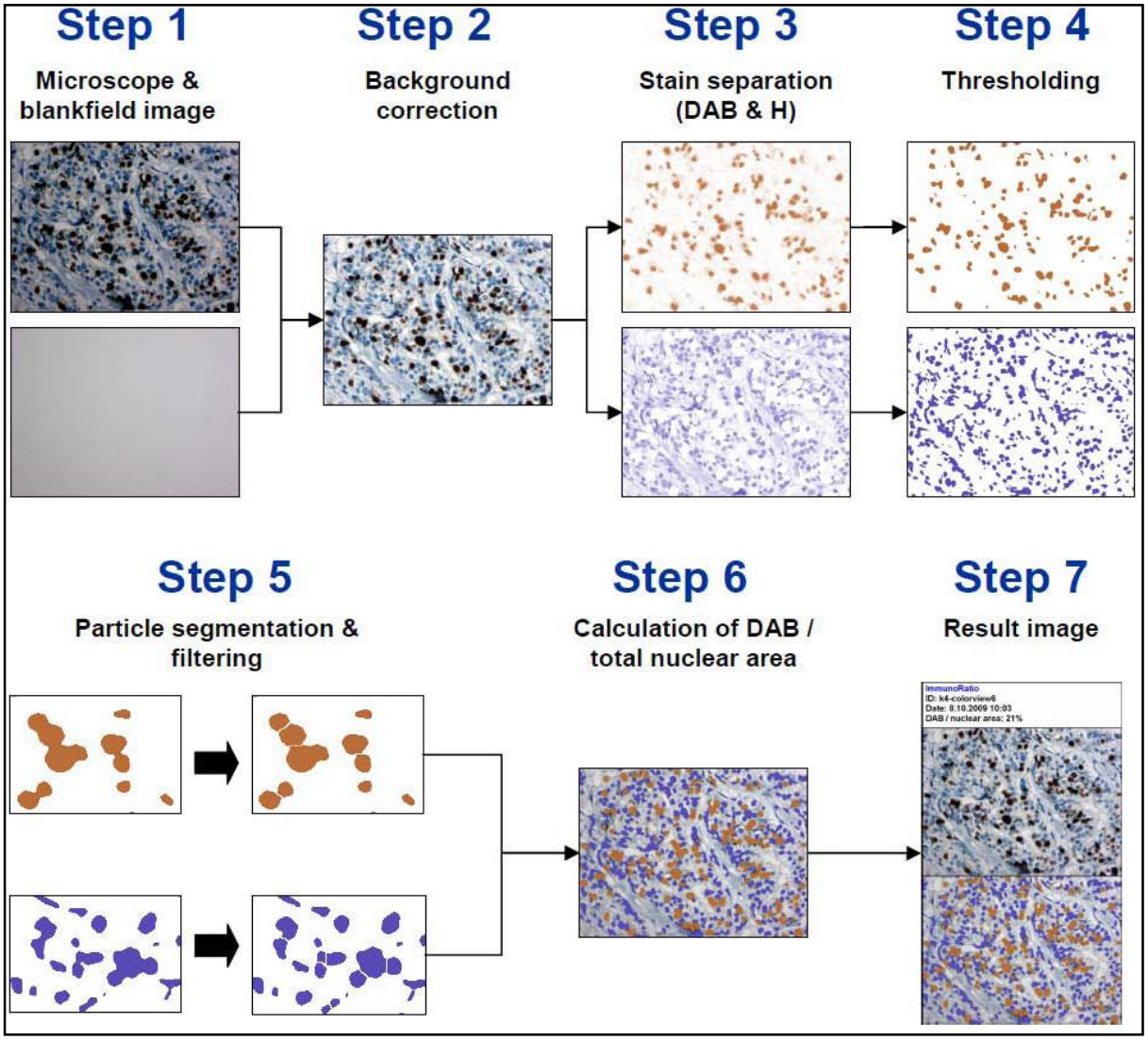


Figure 19: Schematic representation of various steps involved in the processing of uploaded image in the ImmunoRatio™ software

The figure shows the use of colour deconvolution algorithm which separates DAB stained nuclei from haematoxylin counter stain followed by adaptive thresholding to allow nuclear segmentation (Figure obtained with permission from Tuominen et al. (Tuominen et al., 2010))

3.3.3 Image acquisition and evaluation using ImmunoRatio™

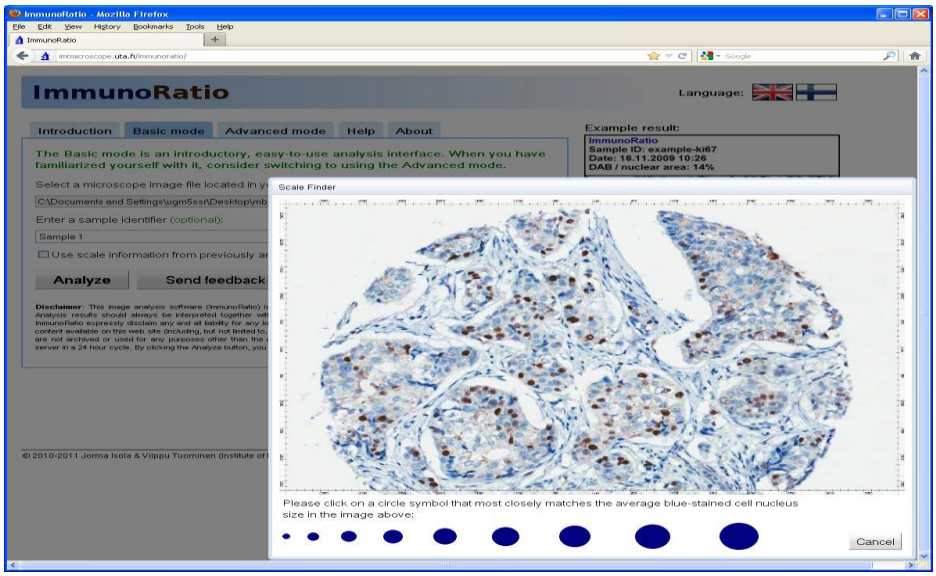
The pictures of the representative TMA core in its entirety were taken using the camera function available in the ImageScope™ software. The captured image was then saved in a JPEG file that was combatable for uploading into the ImmunoRatio™ web based analysis software. The individual JPEG files were uploaded in to the ImmunoRatio™ website for analysis (Figure 20). The software allows the option of analysing the uploaded image in either the basic or advanced mode. In the advanced mode, it offers the option for blank field correction, alteration of image scale and one can also adjust for the threshold of the brown and blue stained nuclei. Once the representative mode was selected, the software calculates the percentage of DAB stained nuclear area over the total nuclear area in the representative TMA core as the LI (Figure 21). The software gives the option of saving the images into the computer from their website for future reference.

3.3.4 Statistical analysis

The variation between the manual and ImmunoRatio™ scoring was analysed using Spearman's and Pearson's correlation for categorical and continuous variables respectively. The measure of agreement was calculated using Kappa statistics for categorical variables with a κ of 0.40 -0.59 considered as moderate agreement, 0.60-0.79 as good agreement and ≥ 0.80 as very good agreement (Landis and Koch, 1977). A weighted Kappa was calculated using the R package "psy" (Performed by Dr Helene Thygesen). The CI was then computed using Efron's bias-corrected and accelerated (BCa) method based on 10,000 bootstrap samples (Efron, 1987). The measure of agreement for a continuous

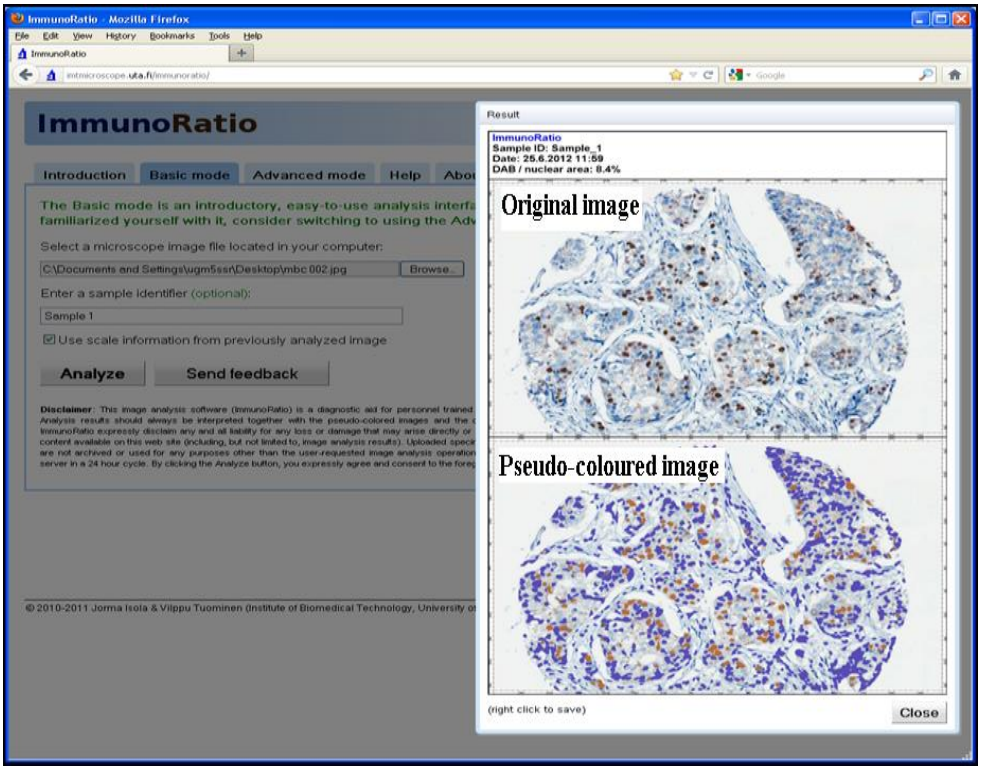
variable was determined using the Bland-Altman plot. The remaining statistical analysis was performed using SPSS version 19 (SPSS, Chicago, IL, USA).

Figure 20: JPEG image of ERα stained TMA core after uploading in to the ImmunoRatio™ website



In this screen shot, the website is asking to select the size (blue circles seen at the bottom of the image) of the ERα stained nuclei prior to calculating the Labelling Index (LI).

Figure 21: Illustration of Labelling Index calculated by ImmunoRatio™ software



Example of Labelling Index (LI) calculated by ImmunoRatio™ for ERα in one the TMA cores (LI =8.4%). The “original image” refers to the image uploaded to the website & “Pseudo-coloured image” represents the colour deconvoluted image produced by the ImmunoRatio™ software. In the latter image, DAB stained areas are represented with brown staining and the blue areas correspond to haematoxylin stained cancer nuclei.

3.4 RESULTS

3.4.1 ERα

The ImmunoRatio™ software calculates the LI, which only accounts for the proportion of the stained nuclei. However the current gold standard for scoring ERα, the Allred score is a sum of the proportion and intensity of staining (Allred et al., 1998). The pre-stained TMA series in our cohort were scored using the Allred method. To allow a comparison with ImmunoRatio™, these scores was converted such that only the proportion of DAB stained nuclei were considered for comparison (Table 8).

Allred score for proportion of stained cancer nuclei	Equivalent ERα manual score	Equivalent ImmunoRatio™ ERα scoring
Negative	0	0
1	1/100	0 to 1%
2	1/10	>1% but ≤ 10%
3	1/3	> 10 but ≤ 33%
4	2/3	> 33 but ≤ 66%
5	1	> 66% up to 100%

Table 8: **Manual and ImmunoRatio™ ERα scores converted into categories to assist analysis**

The median age for breast cancer diagnosis was 58.5 years (IQR = 27) amongst women compared to 70.5 years (IQR = 18) for men in our ERα cohort. Manual scoring showed ERα negative (Allred score = 0) in 43 cases, whereas 26 cases were negative with ImmunoRatio™ (Table 9). There was excellent correlation between the manual and ImmunoRatio™ ERα IHC scoring (Spearman's correlation = 0.872; p = 0.000). We observed only a moderate

Figure 22: Schematic representation of the inability of the ImmunoRatio™ software to differentiate cancer nuclei from stromal elements

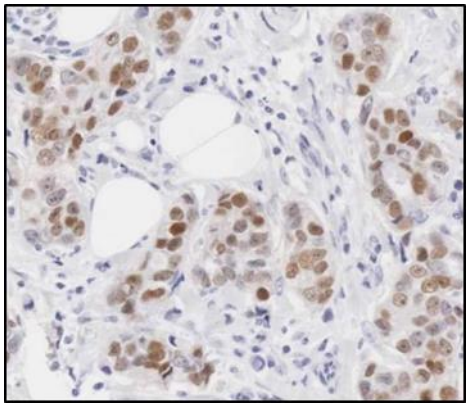


Figure. 22a

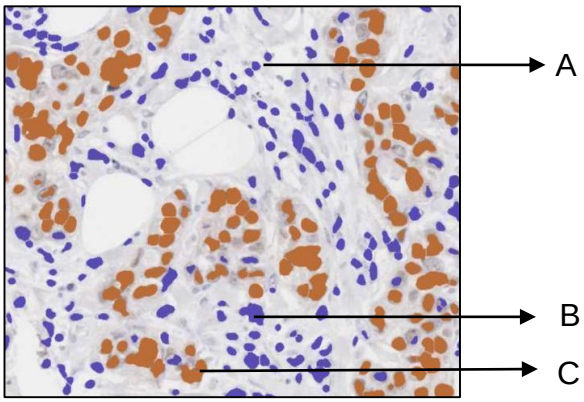
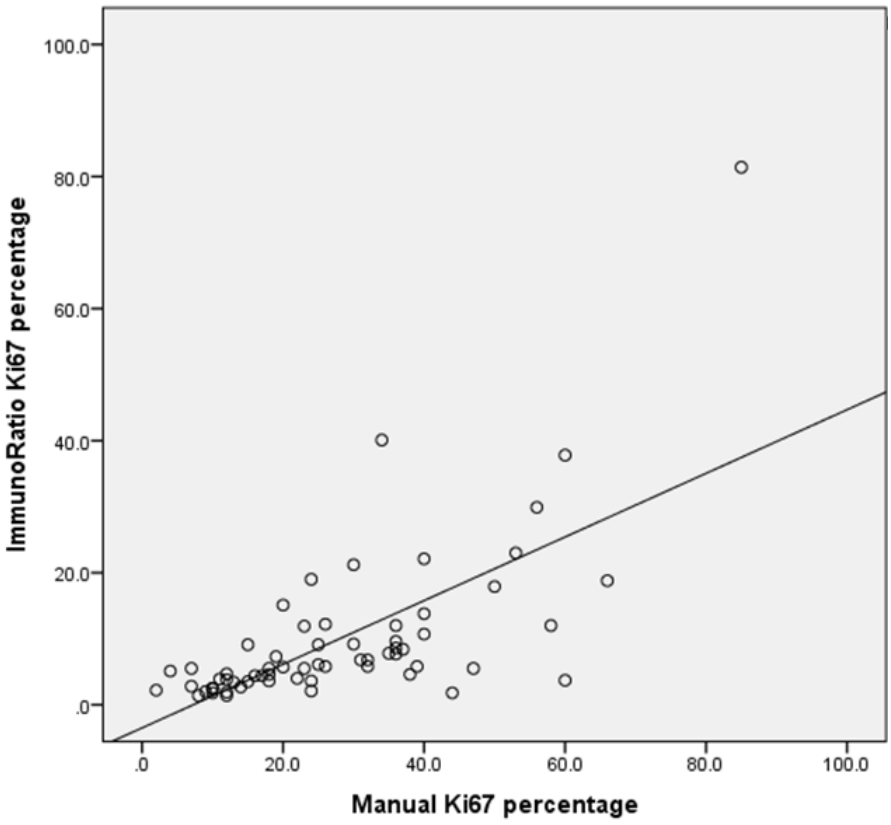


Figure. 22b

a) The original ERα immunohistochemistry image saved in the JPEG format; b) The same image after colour deconvolution by ImmunoRatio™ software. The inability of the software to differentiate true cancer nuclei (C; coloured brown) from the stromal elements (A and B) is depicted.

Figure 23: Correlation between manual and ImmunoRatio™ Ki67 proliferation index



Scatter plot showing the correlation between manual and ImmunoRatio™ Ki67 proliferation index. (Pearson correlation = 0.675, p = <0.001)

Table 9: **Comparison of manual and ImmunoRatio™ ERα scores**

		ImmunoRatio™ ERα						Total
		0	1	2	3	4	5	
Manual ERα	0	24	5	14	0	0	0	43
	1	0	0	4	0	0	0	4
	2	2	1	10	0	0	0	13
	3	0	0	5	13	4	1	23
	4	0	0	1	27	21	10	59
	5	0	0	0	3	16	39	58
Total		26	6	34	43	41	50	200

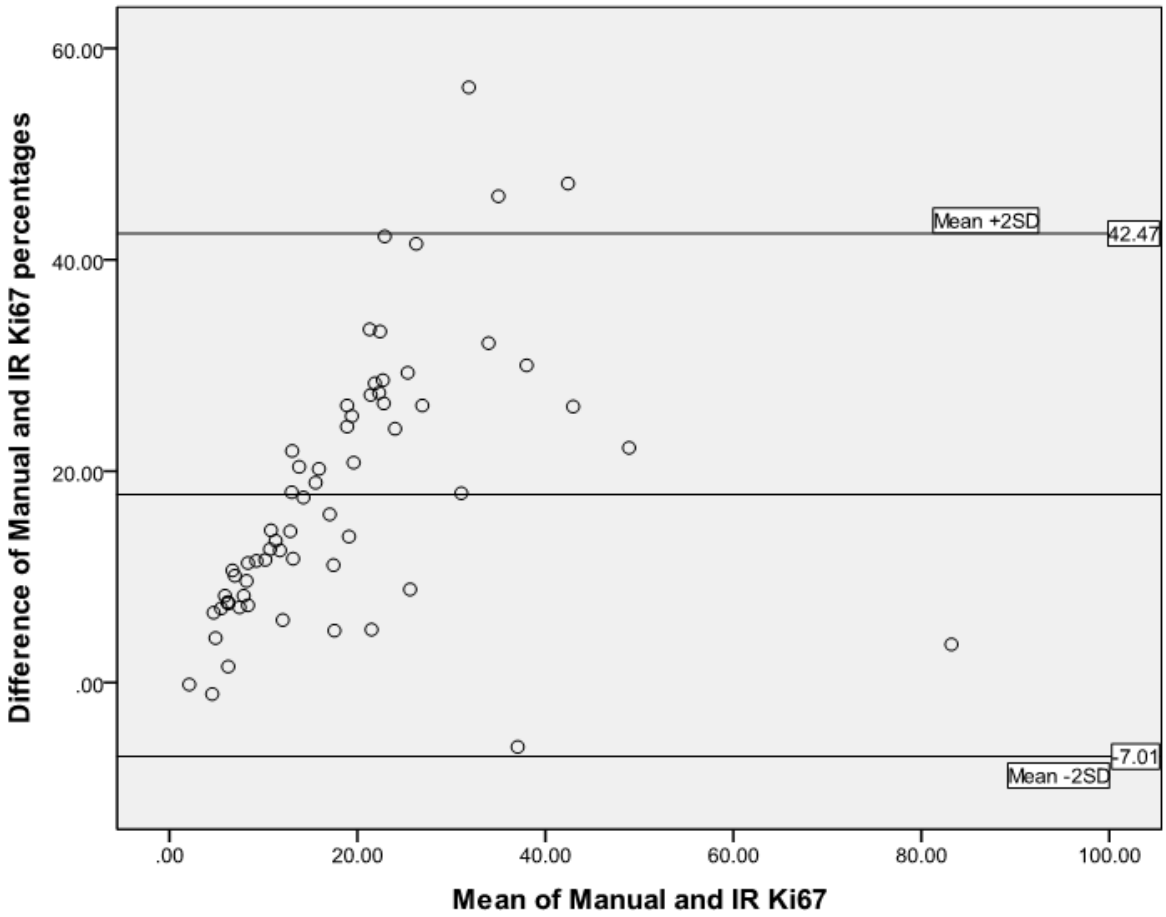
agreement (κ = 0.421) between the manual and ImmunoRatio™ scores. This may have been due to the underestimation of the ImmunoRatio™ LI score due to the inability of the software to differentiate cancer nuclei from the non-cancer stromal elements (Figure 22).

The weighted Kappa was 0.874 (CI - 0.839 to 0.902) and was calculated using absolute weights (i.e. disagreements are weighted by the absolute difference between the automated and manual scoring). The discrepant cores were re-evaluated to ensure that there were no human errors in the manual scores and it was reproducible. The discrepancy with ImmunoRatio™ was most evident with a low (Allred score = 0) or high manual Allred score (Allred score = 4 and 5). The latter could have been due to the underestimation of the LI due to the inability of the ImmunoRatio™ software to differentiate cancer nuclei from the stromal elements (Figure 22b). Whilst amongst ERα negative cores (Allred score = 0) using manual method (n=43), ImmunoRatio™ identified background staining, blotching and inflammatory infiltrate staining as false positive in 19 cores, resulting in an Allred score of 1 (n=5) or 2 (n=14).

3.4.2 Ki67

The median age at breast cancer diagnosis was 65 years (IQR = 21) amongst the Ki67 cohort (n = 64). A direct comparison between Ki67 manual and ImmunoRatio™ scores were possible as both calculated the proportion of DAB stained nuclei in the TMA core. There was a statistically significant correlation (Figure 23) between the manual and ImmunoRatio™ Ki67 score (Pearson correlation = 0.675; p = 0.000). However, when the Bland-Altman plot (Figure 24) was used for determining the measure of agreement between the scores, it became evident that majority of the values were spread away from zero. This indicated that both the manual and ImmunoRatio™ scores produced different results for the expression of Ki67. Similarly it can be seen that the manual scoring was higher than that of the observed Ki67 expression using ImmunoRatio™. The mean difference between the manual and ImmunoRatio™ Ki67 score was 17.76 (SD ± 12.64; CI - 7.023 to 42.54). This discrepancy may have been due to the inability of the software to differentiate between the cancer nuclei and the stromal elements resulting in an underestimation of ImmunoRatio™ LI.

Figure 24: Measure of agreement between manual and ImmunoRatio™ Ki-67 scores



Bland-Altman plot showing the manual Ki67 score being significantly higher compared to the ImmunoRatio™ Ki67 score.

Figure 25: Schematic representation of strong counter staining obscuring the colour deconvoluted ImmunoRatio™ image

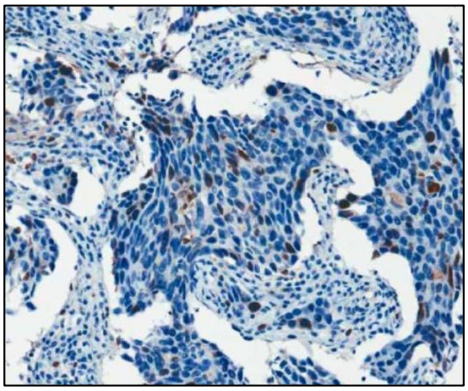


Figure. 25a

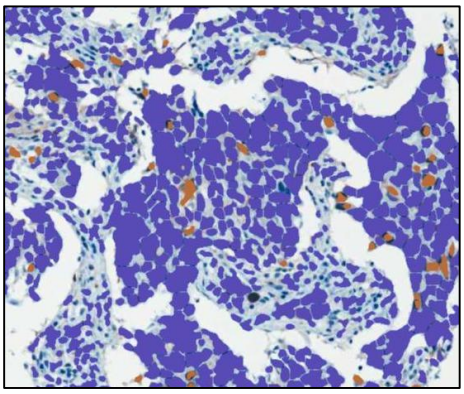


Figure. 25b

a) The original ERα immunohistochemistry image saved in the JPEG format; b) The colour deconvoluted ImmunoRatio™ image. The strong counter staining in this TMA core resulted in obscuring the DAB stained nuclear area in the processed image leading to lower labelling index.

3.5 DISCUSSION

Hormone receptor status is one of the most important factors which determine the choice of adjuvant therapy for patients with breast cancer. Amongst hormonal receptors, ERα remains the most reliable predictor for response to endocrine therapy (Speirs and Walker, 2007). Evaluation of ERα by IHC therefore is standard practice in the management of breast cancer patients. There has been a growing interest in the use of a panel of biomarkers for prognostic and predictive purposes in breast cancer due to the heterogeneous nature of the malignancy. As a proliferative marker, Ki67 gene expression is utilised in Oncotype Dx™ and Mammaprint™ assays for determining suitability for the choice of adjuvant treatment in breast cancer patients. In clinical trials, IHC evaluation of Ki67 expression is used as the assay of choice for determining the proliferative nature of the cancer (Dowsett et al., 2011).

Manual scoring is the gold standard for IHC analysis, but automated image analysis systems have shown promising improvement over the last decade (Mohammed et al., 2012, Rexhepaj et al., 2008). Centralisation of the services and the resulting increased laboratory workload and volume is one of the driving forces behind the development of automated image analysis systems. Various automated image analysis systems are available commercially, which include those that are fully automated (Slidepath Tissue IA System (Mohammed et al., 2012); Beecher Instruments, TMAx™; (Akbar et al., 2015)) and others which are semi-automated (Applied Imaging, Ariol™; Aperio, TMA-Lab II™; Aperio IHC Nuclear Version 10 algorithm, Aperio Technologies).

The aim was to externally validate ImmunoRatio™ a freely available web based fully automated image analysis system (Tuominen et al., 2010) for scoring nuclear biomarkers in a cohort of breast cancer patients. The ease of navigation and rapid processing of the uploaded images within the ImmunoRatio™ website is advantageous. To perform the algorithm, user input is required to select nuclei which best match the size of the stained nuclei to be evaluated. However the software cannot differentiate cancerous nuclei from those of other cell types. There was poor discrimination of nuclei of normal breast epithelial cells and inflammatory cell infiltrates surrounding cancer cells. This is arguably one of the biggest challenges faced by automated systems which still cannot outperform a trained histopathologist (Gurcan et al., 2009). However researchers have recently identified fully automated computer generated algorithm to differentiate between cancer and non-cancerous areas in TMA cores (Akbar et al., 2015). The authors of the study have also internally validated the algorithm using ERα scoring in breast cancer TMAs and have shown promising results (Akbar S et al 2015).

The smallest available nucleus size was selected in all the TMA cores analysed in this study. Altering the size of the selected nuclei or other features in the advanced mode did not impact on the ImmunoRatio™ ERα or Ki67 score. Hence majority of the images (> 90%) were analysed using the basic mode platform available in the ImmunoRatio™ website.

The standardisation of counter staining is very important during IHC image evaluation of automated image analysis systems. Though the counter staining was relatively homogenous in the ERα cohort, occasionally strong counter

staining obscured the area under evaluation by the ImmunoRatio™ software (Figure 25a and 25b). ImmunoRatio™ is also susceptible to the known inability of the fully automated image analysis systems to differentiate between areas of section damage, dye precipitates and out of focus scanning.

It is well documented within the literature that automated image analysis systems perform less reliably when there is a multi-compartmental staining pattern (Bolton et al., 2010). The International Ki67 in Breast Cancer Working group acknowledge the cytoplasmic staining of Ki67 using various antibodies and recommended that only the nuclear staining should be taken in to consideration while analysing Ki67 (Dowsett et al., 2011). The background cytoplasmic staining obscured and diminished the ability of the ImmunoRatio™ software to distinguish true nuclear Ki67 staining. Hence only TMA cores with unequivocal nuclear staining and only mild background cytoplasmic staining was selected for analysis (Methodology section 3.3.1, Figure 18, Page no. 86).

In comparison to some of the commercially available fully automated systems (Mohammed et al., 2012, Rexhepaj et al., 2008), ImmunoRatio™ does not account for the intensity of the nuclear staining. Nuclear intensity is measured as a component of analysing ERα using the Allred method, which is considered as the current gold standard. However there is some evidence to suggest that the presence of 10% ERα positive tumour cells should be the ideal cut-off for predicting the response to adjuvant hormonal therapy (Pertschuk et al., 1996). This view was further supported by the recommendation of the American Society of Clinical Oncology and College of American Pathologists, who advised that ERα expression should be reported as the percentage of positive

nuclei due to the clinically observed response of some breast cancer with low levels of ERα expression (Harvey et al., 1999). Therefore ERα LI as determined by ImmunoRatio™ can be applied in a research setting but cannot yet be recommended for analysis of individual patients in the context of treatment tailoring. Even though the current standard for determining the Ki67 proliferative index is similar to that determined using ImmunoRatio™ LI, there were significant discrepancies observed between manual and ImmunoRatio™ scoring. Hence a fully automated image analysis system should be able to differentiate cancer nuclei from stromal elements, account for the background cytoplasmic staining and also have options available to select representative areas to be scored within the TMA core for it to be reliably used for determining Ki67 proliferative index.

3.6 CONCLUSION

The external validation of ImmunoRatio™ showed excellent correlation between manual and ImmunoRatio ERα scoring in breast cancer TMAs. However it was not possible to validate ImmunoRatio™ for analysing Ki67 in breast cancer due to antibody specificity issues and the inability of the software to process multi-compartmental staining. It is recommended that there should be further development of the ImmunoRatio™ software and external validation of it before it can be used for the evaluation of Ki67 in breast cancer.

CHAPTER 4

Gene Expression Profiling in Matched Male and Female Breast Cancer Patients

4.1 INTRODUCTION

It is well established that as with most of the solid tumours, the development and progression of breast cancer is a multistep process with genomic alteration playing a vital role by affecting gene expression and hence the various downstream cellular processes. The heterogeneous nature of FBC is well established at genomic level and has been correlated to its prognosis (Sorlie et al., 2003, Perou et al., 2000, Hu et al., 2006). Female breast cancer has been classified using gene expression profiling in to intrinsic subtypes (Luminal A, Luminal B, Her2 positive and Basal-like/triple negative), with unique clinical and biological behaviours (Sorlie et al., 2003, Perou et al., 2000, Hu et al., 2006). Based on genomic aberrations, FBC has been further classified in to three unique genomic subgroups (Fridlyand et al., 2006), which has been externally validated (Chin et al., 2006). More recently, using a high resolution array comparative genomic hybridisation (aCGH) method, the FBC genomic subgroups has been further differentiated into six subgroups (Jonsson et al., 2010).

However there is paucity in the knowledge of MBC at transcriptomic level, with very few studies exploring its potential. Earlier studies conducted identified comparable genetic changes in male and female breast cancer using a comparative genomic hybridisation (CGH) method (Tirkkonen et al., 1999, Ojopi et al., 2002). A more recent study confirmed these findings and showed similarities between male and female breast cancer in the chromosomal gains observed at 1q, 8q, 16p and 17q as well as for losses at 8p, 16q, 13q and 11q (Rudlowski et al., 2006). Similar findings were observed by Piscuoglio et al with

gains of 1q, 8q and 16p and losses of 1p, 16q and 17p in MBC (Piscuoglio et al., 2016).

More recent studies have however made use of the aCGH technology (Johansson et al., 2012, Johansson et al., 2011, Tommasi et al., 2010), which provides high resolution and better genetic yield compared to the metaphase based CGH studies. These high throughput gene expression studies revealed significant differences amongst similarities between male and female breast cancers with chromosomal gains being more common in MBC compared to deletions in FBC (Johansson et al., 2011). Based on these results, they classified MBC into two distinctive groups, male simple and complex (Johansson et al., 2011). They identified significant similarities amongst the male complex and female complex sub-groups. Conversely, the male simple group was unique and was different from all the other 6 established FBC subgroups. They also showed that, the male simple group was less aggressive and hence associated with a better outcome. The similarities observed between male and female breast cancer through CGH studies have suggested that, the fundamental genes involved in breast cancer development and evolution may be fairly similar (Rudlowski et al., 2006, Tirkkonen et al., 1999).

However, Callari et al., (2011) showed that there was a substantial difference between male and female breast cancer with differentially regulated biological process observed on gene ontology (GO) analysis. Particularly interesting was the difference noted in the up-regulation of the proteins associated with eukaryotic translation (EIF4 and ribosomal proteins) as well as the lack of preference for anaerobic metabolism noted in MBC (Callari et al., 2011).

Furthermore, there was down regulation of ERBB2 and growth factor associated genes with over expression of BCL2 in MBC cases (Callari et al., 2011). It has been also shown that up-regulation of EIF4E selectively enhances the translation of genes involved in tumourigenesis such as MYC and BCL2 (Graff et al., 2008). The enhancement of protein synthesis observed on gene expression analysis has been linked to the development of MBC through increased cell growth, proliferation and activation of the tumour oncogenes.

Similarly, cluster analysis based on the expression of receptor correlated genes (ERBB2, AR and PGR) showed significant differences between male and female breast cancers (Callari et al., 2011). The AR correlated genes were expressed more in MBC, whereas both ERBB2 and PGR related genes in FBC. In comparison to FBC, they noted that there was a 10 fold decreased expression of genes correlated with ERBB2 in MBC (Callari et al., 2011). Similarly, hierarchical clustering of MBC based on differentially expressed genes associated with PGR positive and negative tumours identified from internal and publically available FBC data sets showed significant differences between male and female breast cancer (Callari et al., 2011). Overall these results indicate a diminished role for ERBB2 with AR being the driving gene in MBC biology and a possible different role for PGR in MBC compared to FBC.

Johansson et al., (2012) successfully classified MBC in to two subgroups, luminal M1 and M2, based on gene expression profiling. This correlated well with that of the male complex and simple subgroups identified respectively using genomic aberration study (Johansson et al., 2011). The luminal M1 and M2 subgroups were unique and different from that of the already established

FBC molecular subgroups (Giordano et al., 2004). On GO analysis it was observed that, the luminal M1 subgroup was associated with up-regulation of genes involved in cell cycle, cell migration, cell adhesion and cell division as well as HOX genes (Johansson et al., 2012). Conversely, higher expression of immune response and ER related genes were observed among the luminal M2 group (Johansson et al., 2012). The luminal M1 subgroup was the more frequent and aggressive phenotype in comparison to the luminal M2, which was associated with significant survival advantage.

More recently candidate driver genes have been identified in MBC (Johansson et al., 2013). The study conducted in 53 MBC and 359 FBC tumours identified 30 candidate driver genes in MBC and 67 in FBC tumours. However there were only 2 candidate driver genes that were in common (TAF4 and CD164). Whereas the analysis performed only in the MBC data set identified 45 candidate driver genes. In MBC, amongst the candidate driver genes only 3 were established cancer genes, MAP2K4, LHP and ZNF217. These findings suggests MBC harbour distinct candidate driver genes compared to FBC.

The current level of evidences published within the literature shows significant similarities amongst differences in somatic genetic changes observed in male and female breast cancer. Hence the need for further studies to confirm or refute the somatic genetic changes observed in MBC.

4.2 HYPOTHESIS

Current level of evidence suggests difference amongst similarities between male and female breast cancer. Hence it was hypothesised that, there will be significant differences between male and female breast cancer at transcriptomic level, leading to changes downstream ultimately resulting in the biological differences between male and female breast cancer.

4.2.1 Aim

1. Identify differentially expressed genes of statistical significance in matched male and female breast cancer through gene expression profiling.
2. Amongst differentially expressed genes, identify the protein expression of a gene or a group of genes with clinical significance and evaluate its prognostic significance in MBC cohort using IHC.

4.3 METHODOLOGY

4.3.1 Gene Expression Profiling

Matched male (n=15) and female (n=10) breast cancer patients were identified for gene expression profiling. The patients were identified from an existing database maintained at Pathology Department, Leeds Institute of Molecular Medicine, University of Leeds. The groups were matched for known prognostic pathological variables (type of cancer, hormonal receptors and lymph nodal status) as well as for the treatments received (surgical and adjuvant). Only invasive ductal cancers which were G1/2 with no lymph nodal involvement or N1 disease (1 to 3 involved nodes) were selected.

Even though macro dissection was not performed, H&E staining of the sections were performed to ensure that more than 80% of the FFPE sections contain breast cancer. The selected FFPE block was sectioned at 10 µm thickness (performed by Ms. Jennifer Pollard) and a minimum of four untreated sections were sent to ALMAC diagnostics™ in RNase-free-tube. Samples were randomised and re-named to prevent order bias and batching effect by ALMAC diagnostics™ (ALMAC diagnostics, Craigavon, United Kingdom). The gene expression profile experiment was undertaken by ALMAC diagnostics™. A summary of the methodology for gene expression profiling as disclosed by ALMAC diagnostics™ is detailed below.

4.3.1.1 RNA isolation

“All FFPE tissues (pre-cut section “curls” in RNase-free tubes) were first de-paraffinized with a xylene-based extraction followed by ethanol dehydration.

Tissue samples were then disrupted during an overnight incubation with a Proteinase K lysis buffer. Nucleic acids were bound to a glass fiber filter in the presence of a chaotropic salt under conditions that were optimized specifically for RNA recovery. Bound RNA was subjected to a series of wash steps to remove contaminating cellular components, and any residual DNA was digested by incubation with DNase. A second round of Proteinase K digestion followed by further wash steps was done to improve the final purity of the RNA, which was then eluted from the glass fibers in a small volume of low-salt elution buffer. The RNA samples were examined for concentration and purity using a spectrophotometer and the standards used are detailed in appendices 3 (Page no. 180)".

4.3.1.2 Amplification of total RNA

"Total RNA was amplified using the NuGEN™ WT-Ovation™ FFPE RNA Amplification System. First-strand synthesis of cDNA was performed using a unique first-strand DNA/RNA chimeric primer mix, resulting in cDNA/mRNA hybrid molecules. Following fragmentation of the mRNA component of the cDNA/mRNA molecules, second-strand synthesis was performed and double-stranded cDNA was formed with a unique DNA/RNA hetero-duplex at one end. In the final amplification step, RNA within the hetero-duplex was degraded using RNaseH, and replication of the resultant single-stranded cDNA was achieved through DNA/RNA chimeric primer binding and DNA polymerase enzymatic activity. The amplified single-stranded cDNA was purified for accurate quantitation of the cDNA and to ensure optimal performance during the fragmentation and labelling process. The single stranded cDNA was assessed using spectrophotometric methods in combination with the Agilent Bioanalyzer.

The details of the quality control performed in the single stranded cDNA are described in appendices 4 (Page no. 181)".

4.3.1.3 cDNA processing

"The appropriate amount of amplified single-stranded cDNA was fragmented and labelled using the FL-Ovation™ cDNA Biotin Module V2. The enzymatically and chemically fragmented product (50-100 nt) was labelled via the attachment of biotinylated nucleotides onto the 3'-end of the fragmented cDNA. The resultant fragmented and labelled cDNA was added to the hybridization cocktail in accordance with the NuGEN™ guidelines for hybridization onto Affymetrix GeneChip® arrays. Following hybridization for 16-18 hours at 45°C in an Affymetrix GeneChip® Hybridization Oven 640, the array was washed and stained on the GeneChip® Fluidics Station 450 using the appropriate fluidics script, before being inserted into the Affymetrix autoloader carousel and scanned using the GeneChip® Scanner 3000". The array data generated was subjected to assessment of GeneChip quality control and expression data integrity. The details of the quality control and data integrity analysis performed are described in appendices 5 (Page no. 182-186). Three MBC samples which failed quality control and data integrity analysis were removed before continuing the analysis".

4.3.1.4 Data summarisation and normalisation

"The scanned Almac Breast DSA™ microarray image data was translated into signal intensities using Robust Multichip Algorithm (RMA) to determine the transcript abundance. The RMA using Affymetrix' Expression Console version 1.1, was applied to the raw data in order to generate the processed data. The

first step was to correct the probe signals for background. The background correction component of RMA assumes the observed signal is a combination of an exponentially distributed true signal and a normally distributed noise component. Once the background subtraction has been performed, all arrays are normalised using a quantile normalisation procedure. It was assumed that the distribution of probe abundances is nearly the same in all samples for quantile normalisation procedure. Finally a median polish was used to calculate a summary value for each probe set".

4.3.1.5 Differential gene selection

The ALMAC diagnostics™ used an in-house method to select differentially expressed probe sets using background, variance and fold change filters using FeatureSelection_Workflow version 102. All probe sets (60,856) on the Almac Breast DSA™ array was subjected to general filtering with background and variance filtering. Any probe set with an expression on the background level (Background p-value = 0.3) was removed. Similarly, any probe set with a variance below the mean global variance was removed in an intensity dependent manner (α-value = 0.9). The group and statistical filtering was performed in 14,959 probe sets that passed the general filtering criteria. The data was then filtered using an advanced fold change filter to identify those probe sets that are differentially expressed. The student's t-test was performed on the probe sets passing general and fold change filters to determine the statistical significance after multiple test correction (pFDR).

The preliminary findings of the GEP experiment including bioinformatics data and functional enrichment analysis was provided by ALMAC diagnostics™.

Subsequent bioinformatics data mining, analysis, interpretation of the results and drawing inference for further validation was undertaken through guidance from Dr. Alastair Droop, Bioinformatics research fellow at University of Leeds.

4.3.2 Western Blot

The experiment was performed under the guidance of Dr E T Verghese (PhD Research Fellow, University of Leeds) to determine the specificity of the Fibronectin antibody used for immunohistochemical analysis.

4.3.2.1 Reagents used for Western Blot

1. 4 x NuPage® Lithium Docedyl Sulphate (LDS) Sample Buffer (pH 8.5)

The stock was obtained from Life Technologies, Carlsbad, USA. The components of NuPage® LDS sample buffer (500 ml) is as follows,

Glycerol	10% (v/v)
Tris base (pH 8)	424 mM
Tris-hydrochloric acid	564 mM
LDS	2% (w/v)
Ethylene di-amine tetra-acetic acid (EDTA)	0.5 mM
Serva blue G250	1 mM
Phenol red	0.7 mM

2. 20 x NuPAGE® MOPS Sodium Dodecyl Sulphate (SDS) Running Buffer (pH 7.2)

The stock was obtained from Life Technologies Carlsbad, USA. The components of the 20 x NuPAGE® MOPS SDS Running Buffer (500 ml) is as follows,

3-(N-morpholino) propane sulphuric acid (MOPS)	1.0 M
Tris Base (pH 8)	1.0 M
Sodium dodecyl sulphate (SDS)	70 mM
EDTA (pH 8)	20 mM

3. 20 x NuPAGE® MOPS SDS Transfer Buffer (pH 7.2)

The stock was obtained from Life Technologies Carlsbad, USA. The components of the 20 x NuPAGE® MOPS SDS Transfer Buffer (500 ml) is as follows,

Bicine	500 mM
Bis-Tris	500 mM
EDTA (pH 8)	20 mM

4. β – Mercaptoethanol

Obtained from Sigma Aldrich, St. Louis, USA (Catalogue no: M3148).

5. Precast NuPAGE® Novex® 4-12% Bis-Tris Gel (0.1mm, 10 wells)

Obtained from Invitrogen, Carlsbad, USA (Catalogue no. NP0321)

6. Amersham Hybond - P Polyvinylidene Difluoride (PVDF) membrane

Obtained from GE Healthcare, Little Chalfont, UK (Catalogue no. RPN2020F)

7. Tween – 20

Obtained from Fisher Scientific, Waltham, USA (Catalogue no: PBP337-100).

8. Tris-Buffered Saline Tween (TBST)

The components of TBST are Tris-Buffered Saline (TBS) (pH 7.4) and Tween 20. 1 x TBS (1 litre) was made by mixing 60 ml of stock solution of 2.5 M NaCl and 20 ml of 1 M Tris-Hydrochloric acid (Tris-HCl) (pH 7.4) in deionised water.

Sodium chloride	0.15 M
Tris-HCl	0.02 M
Tween -20	0.1% (v/v)

9. Polyclonal Swine Anti Rabbit horse radish peroxidase (HRP) conjugated secondary antibody

Obtained from Dako, Santa Clara, USA (Catalogue no: P0217).

10. Protein Molecular Weight Markers

SeeBlue® Plus2 Pre-stained Protein Standard was used. 500 µl stock was obtained from ThermoFisher Scientific, Paisley, UK (Catalogue no. LC5925).

11. SuperSignal® West Femto Chemiluminescent Substrate

Obtained from Thermo Scientific, Waltham, USA (Catalogue no: 34095)

4.3.2.2 Gel Electrophoresis of Proteins

1. Sample preparation

Standardised protein was added to 1 x LDS sample buffer (pH 8.5) containing 5% (v/v) of β-mercaptoethanol in a 0.5 ml Eppendorf tubes to make a total volume of 20µl. The Eppendorf tubes were then closed, the cap pierced using a 21 gauge needle, spun and then placed in a hot plate at 105^0 for 5 minutes. The tubes were then transferred to ice for 5 minutes. The tubes were centrifuged for 30 seconds at 8,000g and then gently re-suspended and placed on the ice.

2. Gel preparation

The Precast NuPAGE® Novex® 4-12% Bis-Tris Gel was carefully removed from package and gently rinsed with tap water. The 1 x NuPAGE® MOPS SDS

running buffer was made by mixing 475 ml of deionised water and 25 ml of 20 x NuPAGE® MOPS SDS running buffer. The running buffer is added to the top of the gel until it fills quarter of the tank. Twenty microgram of sample and 5 µl SeeBlue® plus2 marker was added to the lanes. Electrophoresis was then performed in Xcell SureLock™ Mini-Cell system (life Technologies, Model no: EI0001) at a constant current of 180V for 1 to 1.5 hours.

3. Transfer of proteins

Amersham Hybond-P Polyvinylidene difluoride (PVDF) membrane was used for transferring the proteins separated through gel electrophoresis. The 1 x NuPAGE® MOPS SDS transfer buffer was used to prime the filter paper and sponges used for the protein transfer. The membrane itself is activated by soaking in neat methanol for 30 second, washed in deionised water for 10 minutes with constant agitation on a shaker and finally soaked in the transfer buffer. The pre-cast gel was removed before loading into the transfer module with the membrane in place over the gel ensuring that there are no air bubbles. The transfer module is then filled with the transfer buffer and run at a constant current of 30 V for 1.5 hours.

4. Immunoblotting of proteins

The PVDF membrane with the transferred proteins was blocked in 20 ml of blocking buffer for 1.5 hours at room temperature. The blocking buffer consists of Tris buffered saline with tween (TBST) and 5% (w/v) of Marvel skimmed milk. The membrane was then washed 3 x 10 min with TBST with gentle agitation on spiramix. The membrane was then transferred to 5% (w/v) skimmed dried milk

(Marvel) in TBST containing primary antibody, anti-fibronectin 1 (anti-FN1) in 1:250 concentration. The membrane was then allowed to incubate with the primary antibody overnight at 4^0C on a roller. The membrane was then washed 3 times in TBST for 10 minutes each. The membrane was then incubated in 5% (w/v) skimmed dried milk (Marvel) in TBST with secondary antibody (Polyclonal swine anti-rabbit HRP conjugated) at 1: 5000 dilutions for one hour at room temperature. Subsequently the membrane was again washed 3 times in TBST for 10 minutes each. During optimisation of secondary antibody, samples were tested by incubating without primary antibody to ensure there aren't any false positive results. The membrane was then incubated with 200µl of SuperSignal® West Femto Chemiluminescent substrate for 5 minutes to facilitate visualisation of the antibody. The membrane was covered in a cling film ensuring that there are no air bubbles and visualised using ChemiDoc® MP imaging system (Bio-Rad, model no.170-8280), with exposure time varying depending on the signal intensity. The control ladder used was Precision Plus Protein™ Dual Colour Standards (BioRad, Cat no. 161-0374).

4.3.3 Immunohistochemistry

IHC analysis for Fibronectin was performed in MBC TMAs 1 to 6, TMA 8 and was not performed in MBC TMA 7 due to core exhaustion. A rabbit polyclonal anti-fibronectin 1 (anti-FN1) antibody (Prestige Antibodies® Powered by Atlas Antibodies, Product no: HPA027066), which was validated in Human Protein Atlas project was used in this cohort.

The IHC was performed with EnVision method as described earlier (Chapter 2, Methodology Section 2.3.3, Page no. 33-34). The main differences was that instead of using PBS, PBS with 0.2% Tween (PBST) was used for washing the slides and the slide racks were not passed through copper sulphate prior to counter staining with Mayer's haematoxylin. The anti-FN1 antibody was used at 1:50 concentration in the MBC TMAs and overnight incubation was done in a humidified chamber at 4^0C. The anti-rabbit secondary antibody conjugated with HRP was used (DAKO® –Envision Kit) for visualisation of the antigen-antibody reaction. In each batch, a breast cancer control TMA was used for positive control and negative control was run by incubating with PBST instead of primary antibody.

The anti-FN1 antibody was optimised in male and female breast cancer full slide sections to confirm the location and intensity of the staining. Subsequently, a control FBC TMA was used to confirm the ideal concentration of the antibody in TMAs.

The manual IHC scoring was performed after scanning the slides at 20x magnification (ScanScope XT, Aperio) using the ImageScope™ software (Aperio). A random selection of TMA cores (n = 100) were also scored by an independent researcher (Dr E T Verghese, PhD Research Fellow, University of Leeds) to ensure minimal inter-observer variability.

4.3.4 Statistical analysis

The clinical and pathological variables, like age at diagnosis, tumour size, grade and nodal status were available for analysis in the MBC TMA cohort. The main

outcome variable was overall survival (OS). The correlation between fibronectin expression and various clinical and pathological variables were calculated using Spearman's correlation coefficient. The measure of agreement was calculated using the K statistics. The strength of association between clinico-pathological variables and fibronectin with that OS was ascertained using Cox logistic regression analysis. The prognostic variables that were significant on univariate analysis were then entered in to a multivariate analysis model to identify independent predictors for OS at 5 year. The survival analysis was performed using the Cox proportional hazard model. The survival curves were plotted using Kaplan-Meier method and compared using the Log rank test. The data was analysed using SPSS version 19 software and a P value of ≤ 0.05 was considered as statistically significant.

4.3.4.1 Bioinformatics

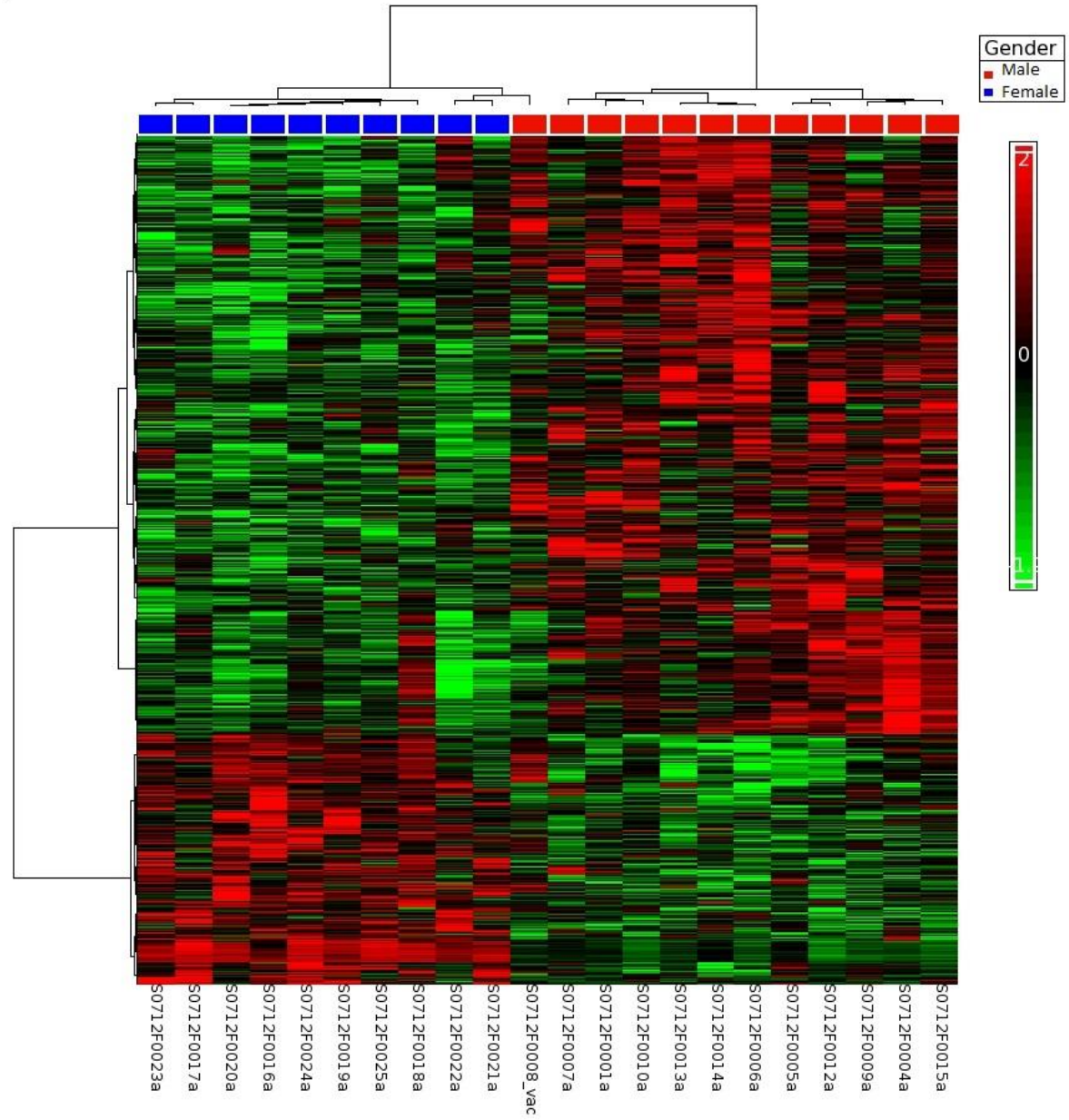
The ALMAC Diagnostics Breast DSA™ research tool consisting of 60, 856 probe sets were used in the gene expression profiling analysis. The Robust Multichip Algorithm (RMA) using Affymetrix™ Expression console version 1.1 was applied to the raw data in order to generate the processed data. Differentially expressed probe sets were identified using background, variance and fold change filters with FeatureSelection_Workshlow version 102. Further details of the initial bioinformatics analysis performed by ALMAC Diagnostics™ are detailed in Appendix 6 (Page no. 187).

4.4 RESULTS

Bioinformatics analysis (performed by ALMAC Diagnostics™) was undertaken to identify probe sets that were differentially expressed. The general filtering was carried out using background expression and variance filters to identify a list of 14,949 probe sets. The criteria used was to remove any probe set with an expression on the background level (p = 0.3). Similarly, any probe set with a variance below the mean global variance was removed in an intensity dependent manner (α-value = 0.9). The data was then subjected to advanced fold change filtering based on the intensity and variation of probe sets between the groups. Student's t-test was performed to determine differentially expressed probe sets and false discovery rate (pFDR) was calculated to account for multiple test correction. There were 735 differentially expressed genes using the less stringent criteria (Significance level for log fold change = 0.05) and 117 genes were differentially expressed using the stringent criteria (Significance level for log fold change = 0.01). Hierarchical agglomerative clustering was undertaken amongst differentially expressed genes, using both less stringent (Figure 26) and stringent criteria (Figure 27). The agglomerative clustering starts with each sample/gene as a separate cluster and merge them into successive larger clusters. The analysis was performed in Partek GS v 6.5.

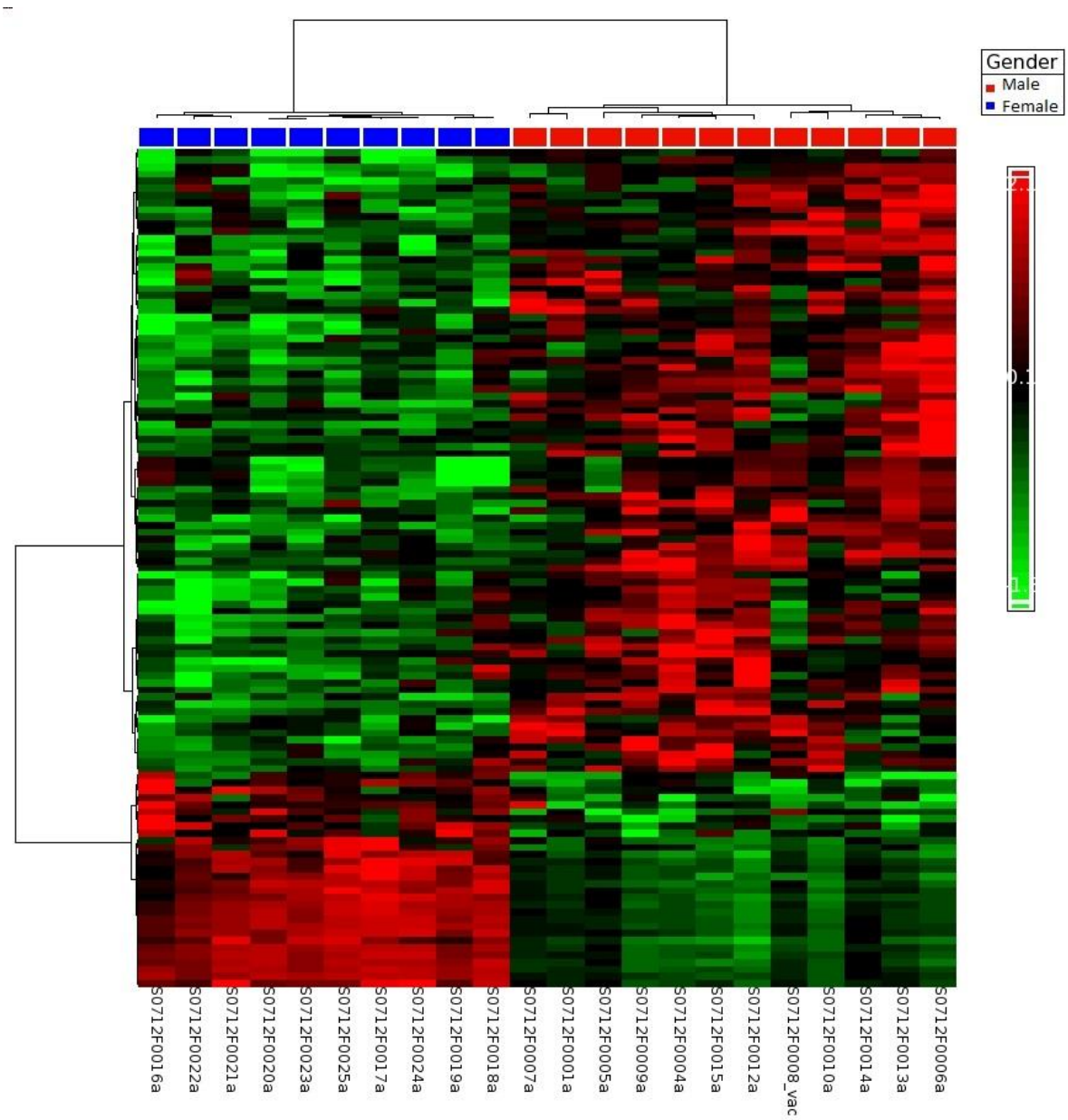
Functional enrichment analysis of differentially expressed genes identified with the less stringent criteria (n=735) was undertaken using Ingenuity Pathway Analysis (IPA v9.0 build 116623, Ingenuity® Systems, content version 3211) and Gene Ontology (GO) analysis using Functional Enrichment Tool (FET version 1.0).

Figure 26: Heat map showing hierarchical clustering using less stringent criteria ($p \leq 0.05$)



Hierarchical agglomerative clustering showing differentially expressed genes in the male and female breast cancer samples using less stringent criteria ($p \leq 0.05$). Rows: individual genes; Columns: individual tissue samples. Pseudo-colours indicate transcript level below, equal to or above the mean (green, black and red respectively).

Figure 27: Heat map showing hierarchical clustering using stringent criteria ($p \leq 0.01$)



Hierarchical agglomerative clustering showing differentially expressed genes in the male and female breast cancer samples using stringent criteria ($p \leq 0.01$). Rows: individual genes; Columns: individual tissue samples. Pseudo-colours indicate transcript level below, equal to or above the mean (green, black and red respectively).

The significance of the association between differentially expressed genes to the total number of genes in a canonical pathway/GO term was expressed as a ratio and Fisher's exact test was performed to establish statistical significance. Benjamini & Hochberg adjusted p-value was calculated to account for multiple testing and represented as pFDR. The top 10 enriched canonical pathways and biological processes are represented respectively in Table 10 and 11.

Table 10: Top 10 enriched canonical pathways

No	Ingenuity Canonical Pathways	Ratio	p-value	p-value (FDR)
1	Hepatic Fibrosis / Hepatic Stellate Cell Activation	15/147	0.0000	0.0115
2	EIF2 Signalling	11/101	0.0001	0.0162
3	Actin Cytoskeleton Signalling	16/238	0.0007	0.0589
4	Glioma Invasiveness Signalling	8/60	0.0008	0.0589
5	Leukocyte Extravasation Signalling	15/199	0.0013	0.0676
6	Intrinsic Prothrombin Activation Pathway	5/34	0.0014	0.0676
7	ILK Signalling	14/193	0.0020	0.0851
8	Glucocorticoid Receptor Signalling	18/295	0.0026	0.0912
9	VEGF Signalling	9/99	0.0028	0.0912
10	Rac Signalling	10/123	0.0035	0.1033

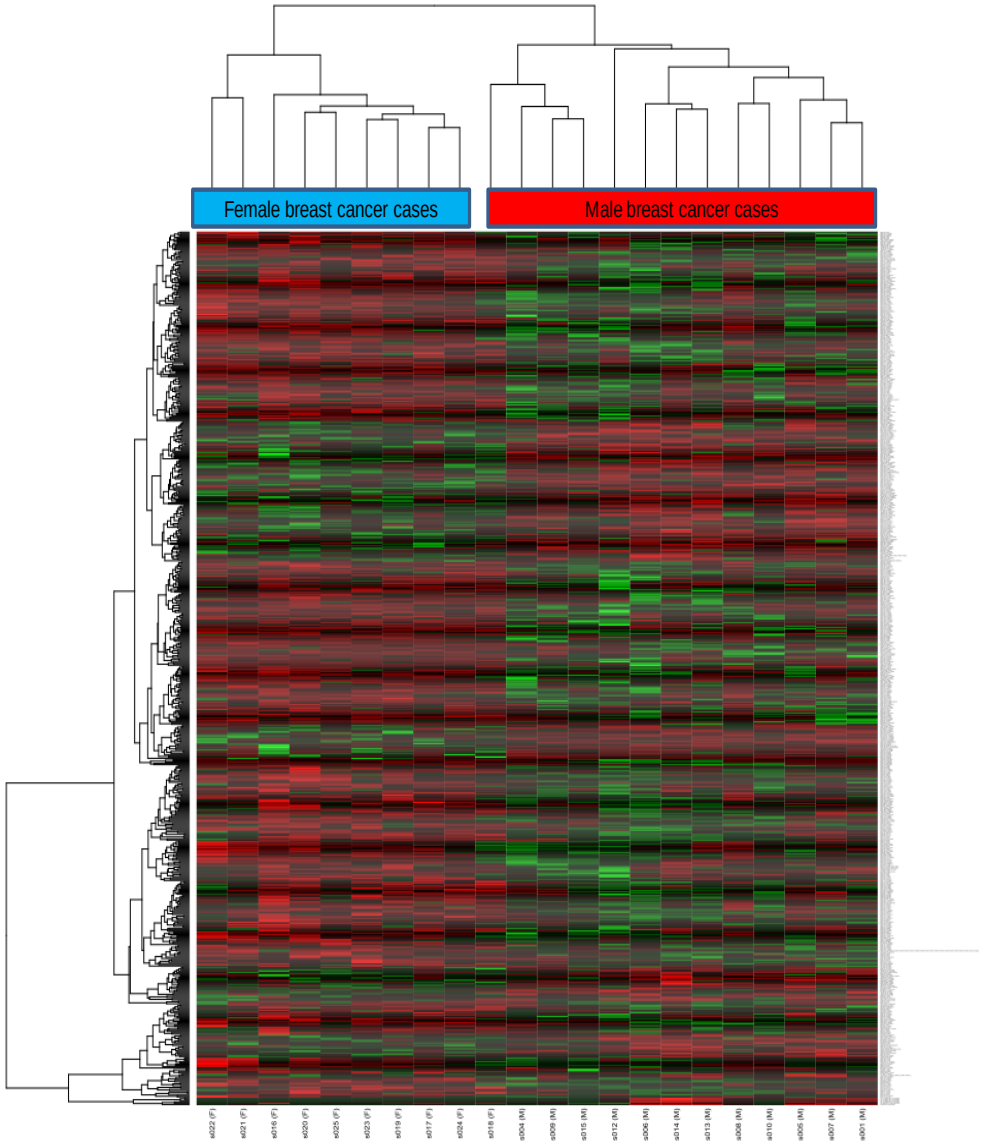
Table 11: Top 10 enriched biological processes

No	GO Biological Process	Ratio	p-value	p-value (FDR)
1	GO:0007155 cell adhesion	68/708	0.0000	0.0000
2	GO:0022610 biological adhesion	68/709	0.0000	0.0000
3	GO:0031589 cell-substrate adhesion	19/137	0.0000	0.0002
4	GO:0006928 cellular component movement	38/458	0.0000	0.0003
5	GO:0030199 collagen fibril organization	8/26	0.0000	0.0009
6	GO:0009653 anatomical structure morphogenesis	63/1023	0.0000	0.0016
7	GO:0048519 negative regulation of biological process	93/1754	0.0000	0.0029
8	GO:0048856 anatomical structure development	113/2253	0.0000	0.0029
9	GO:0007160 cell-matrix adhesion	14/105	0.0000	0.0040
10	GO:0048523 negative regulation of cellular process	86/1610	0.0000	0.0040

The data and results as detailed above provided by Almac Diagnostics™ were scrutinised and was subjected to further data mining. Since the cohort contained matched male and female breast cancer patients, there was selection and analytical bias originating from differentially expressed sex linked genes. Hence sex linked genes / probe sets were excluded (n = 79) and analysis was repeated in the remaining 656 probe sets. The hierarchical clustering performed using the new gene set generated a more representative cluster of differentially expressed genes (Figure 28). The majority of genes were up-regulated (n = 487) and rest were down-regulated (n = 169).

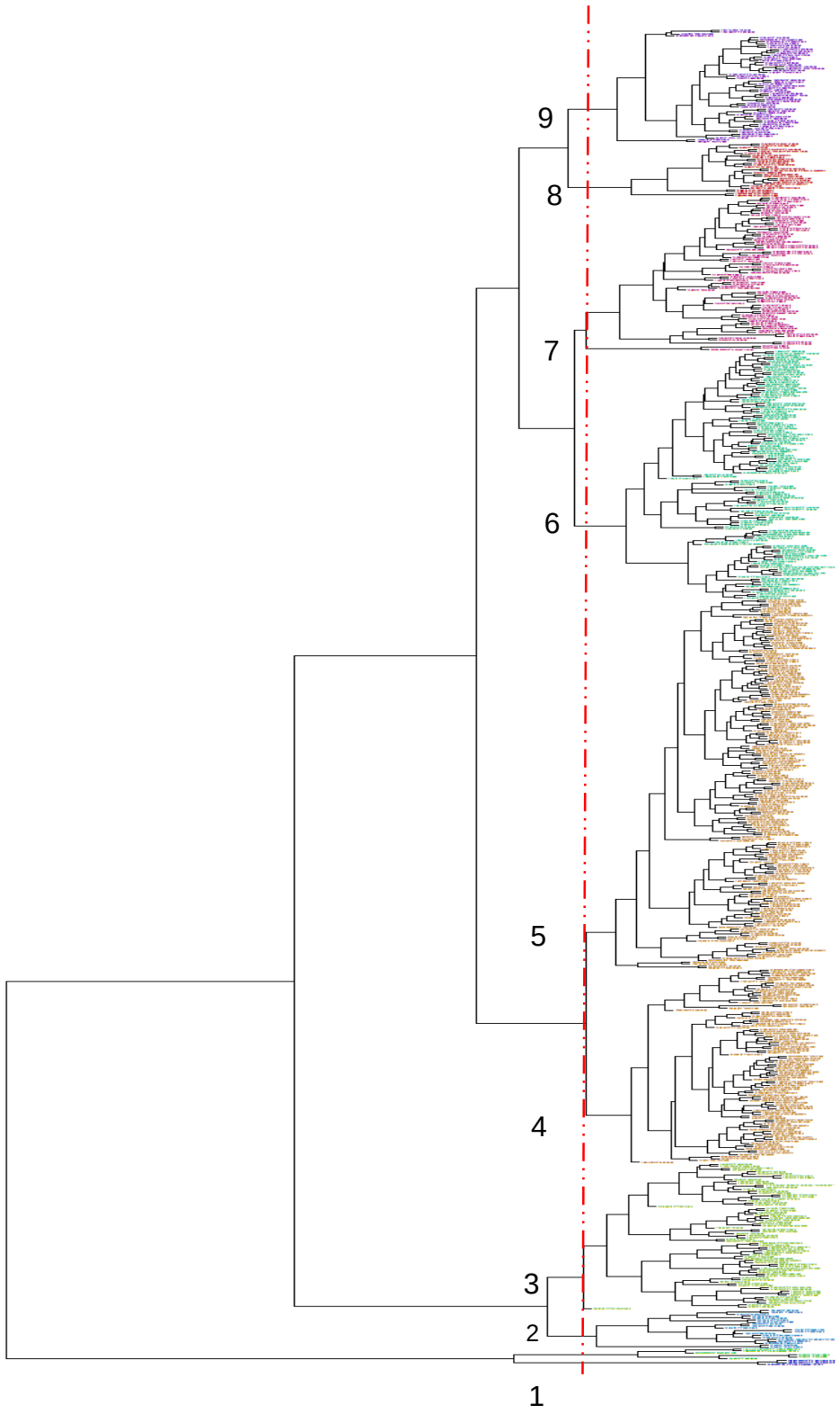
The subsequent data mining concentrated on identifying either a cluster or a group of genes with a role in the natural history of breast cancer. The data was initially analysed using the group of genes represented in 9 major clusters or within its sub-clusters (Figure 29). The genes in each cluster were entered into Oncomine™ and/or ToppGene™ database to extract biological insights into the data. In addition to this, the top 10 up and down regulated genes encoding for proteins were individually analysed (Table 12 and 13). The role of each individual gene in cancer and in particular breast cancer was extracted from publically available data sets and literature. The information was mostly gathered from various resources available within the National Center for Biotechnology Information (NCBI) website. In the NCBI site the “Gene Expression Omnibus” (GEO) and “Gene” depositories were mainly used for data mining. The bibliography there in was reviewed to identify published evidence supporting the role of individual genes in breast cancer.

Figure 28: Hierarchical agglomerative clustering of less stringent genes after excluding sex linked genes



Hierarchical agglomerative clustering showing non-sex linked differentially expressed genes with less stringent criteria ($p \leq 0.05$) in the matched male and female breast cancer groups. Rows: individual genes; Columns: individual tissue samples. Pseudo-colours indicate transcript level below, equal to or above the mean (green, black and red respectively).

Figure 29: Representation of the 9 clusters selected from the heat map and representative genes used for data mining.



The red dotted line represent the level at which 9 clusters was selected. Different colour was assigned to each cluster of genes.

Table 12: The top 10 up-regulated non-sex linked genes in matched male and female breast cancer samples

Gene Symbol	Gene Description	Chromosomal Location	Fold-Change (A vs. B)*	p value	pFDR**
GJA1	Gap junction protein, alpha 1	6p11.1	8.26212	0.000436	0.000045743
STC2	Stanniocalcin 2	5p11	7.61776	0.006737	0.00010712
SERPINA6	Serpin peptidase inhibitor	14p11.1	6.67724	0.001699	0.000081345
PI15	Peptidase inhibitor 15	8p11.1	5.96069	0.012226	0.00012912
SULF1	Sulfatase 1	8p11.1	5.30301	0.000853	0.00006264
FHL2	Four and a half LIM domains 2	2p11.1	4.89209	0.000652	0.000056279
NPNT	Nephronectin	4p11	4.84637	0.015899	0.00014557
COL1A1	Collagen, type I, alpha 1	17p11.1	4.70691	0.008838	0.00011447
MMP11	Matrix metallopeptidase 11 (stromelysin 3)	22p11.1	4.62931	0.006139	0.00010772
FN1	Fibronectin 1	2p11.1	4.33815	0.001337	0.00007548

*"A" represent male samples and "B" represent female breast cancer samples. **pFDR (Positive false discovery rate) represent the significance after correction for multiple testing.

Table 13: The top 10 down-regulated genes amongst non-sex linked genes in matched male and female breast cancer samples

Gene Symbol	Gene Description	Chromosomal Location	Fold-Change (A vs. B)*	p value	pFDR**
ERBB4	v-erb-a erythroblastic leukaemia viral oncogene homolog 4	2p11.1	-4.47078	0.001486	0.000076079
HMGCS2	3-hydroxy-3-methylglutaryl-CoA synthase 2 (mitochondrial)	1p11.1	-3.09543	0.054133	0.00032524
ZC3H7B	Zinc finger CCCH-type containing 7B	22p11.1	-2.81424	0.002774	0.000086649
IGJ	Immunoglobulin J polypeptide	4p11	-2.7442	0.033072	0.00021324
TFAP2B	Transcription factor AP-2 beta	6p11.1	-2.74218	0.033909	0.00021769
PBX1	Pre-B-cell leukaemia homeobox 1	1p11.1	-2.68953	0.024997	0.00017727
KCNC3	Potassium voltage-gated channel	19p11	-2.58495	0.021384	0.00016349
SLC20A2	Solute carrier family 20 (phosphate transporter)	8p11.1	-2.55447	0.049414	0.0002981
ACACB	Acetyl-CoA carboxylase beta	12p11.1	-2.54331	0.011321	0.00012559
PAK1	p21 protein (Cdc42/Rac)-activated kinase 1	11p11.11	-2.43954	0.037239	0.00023361

***"A" represent male samples and "B" represent female breast cancer samples. **pFDR (Positive false discovery rate) represent the significance after correction for multiple testing.

The data mining as detailed above showed an emerging common theme of extracellular matrix (ECM) remodelling amongst up-regulated genes. This

indeed was reflected in the Gene Ontology results showing up-regulation of biological processes involved in ECM remodelling (Table11, Page no. 125). The literature was reviewed to identify genes that were shown to be differentially expressed in breast cancer stroma compared to normal breast. The literature review was limited to FBC due to the lack of published comparative studies in MBC. The following studies with publically available data set of genes were identified (Ma et al., 2009, Bergamaschi et al., 2008, Karnoub et al., 2007, Finak et al., 2008). The differentially expressed gene list (non-sex linked genes) from this study was compared with that of the exhaustive gene list extracted from public repositories. This enabled to generate a list of common differentially expressed genes (Table 14)

Table 14: Differentially expressed stromal associated genes in this study and in other publically available data sets

Study details	Differentially expressed genes
Ma et al (2009) [>3 fold differential expression in IDC stoma compared to normal breast stroma]	ANTXR1, ASPN, CACNG4, CD276, CFL1, COL10A1, COL11A1, COL12A1, CTHRC1, EIF5, ENTPD7, EPS8, FN1, FNDC1, GLUL, GREM1, INHBA, ITGA6, ITGBL1, MFAP2, MMP11, MVP, PXDN, RAB31, SPARC, SULF1, SYNPO2L, TUBA1C
Ma et al (2009) [> 3 fold differential expression in DCIS stoma compared to normal breast stroma]	ASPN, COL10A1, COL11A1, COL12A1, COL8A1,CTHRC1,FN1,FNDC1,GJB2,GREM1,IGJ,INHBA, SULF1
Finak et al (2008)	C1orf31, CAPS, CXCL14, GK, ITGB8, ITGBL1, MYB, SPP1
Bermaschi et al (2008)	CD44, CERCAM, COL12A1, COL1A1, COL11, COL3A1, COL4A5, COL5A3, COL6A1, COL6A3, FBN1, HTRA3, ITGA6, ITGB3BP, ITGB5, ITGB8, ITGBL1, LAMC1, MMP11, MMP13, MMP2, PI15, SERPINA1, SERPINA3, SERPINA6, SMOC1, SPARC, TNC
Karnoub et al (2007)	CERCAM, COL10A1, COL11A1, COL8A1, CRYAB, CTHRC1, EMP1, FN1, GREM1, ITGBL1, MMP11, MSR1, RUNX2, SPP1, SULF1

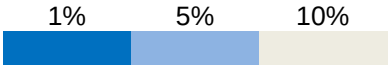
The differential expression of genes identified in this cohort was from matched male and female breast cancer samples. Hence it is possible that, the identified genes may not be truly differentially expressed in MBC. The differential expression observed may be due to either a slight increase (amongst up-regulated genes) or decrease (amongst down-regulated genes) in the mean expression level of that particular gene in all the MBC samples compared to that in all of the FBC samples.

Hence in order to identify genes that are truly differentially expressed between MBC and FBC, comparison with publically available data sets were undertaken. The differentially expressed genes were individually evaluated in Oncomine™ website to determine their expression in FBC compared to that in the normal breast tissue. The results showed that 134 out of 487 genes were up-regulated in FBC compared to normal breast tissue (Table 15). Similarly 39 out of 169 genes were down-regulated in FBC compared to normal breast tissue (Table 15). The genes thus identified (134 up-regulated and 39 down-regulated) were compared with genes that were known to have functional role (publically available repositories) in ECM remodelling or associated with breast cancer stroma. Amongst the up-regulated genes, 15 genes were involved in ECM remodelling (Table 16). However none of the down-regulated genes were related to ECM remodelling.

Each of these genes, its role (prognostic or predictive) in breast cancer and its relevance in ECM remodelling was explored in the literature. The protein by product of these genes and its expression in the breast cancer was analysed. Amongst the genes identified, it was decided to validate the protein expression

Table 15: Common differentially expressed genes identified in this cohort and in the FBC vs. Normal breast analysis of the publically available data sets (Collated from the Oncomine™ website)

Down-regulated genes	Up-regulated genes			
HMGCS2	PTPRD	MVP	FN1	CTHRC1
IGJ	FAM36A	MICAL2	SLC39A11	SIAH2
TFAP2B	ANTXR1	SORD	RASAL2	VCAN
ACACB	C1orf31	UBE2N	ARFIP2	COL11A1
TFAP2B	PTK7	VDR	COL10A1	EDIL3
EDNRB	CTNNB1	TGFB3	COL8A1	AGR3
FBXO2	SYNCRIP	MICB	INHBA	ASPN
FRMD4A	NAV1	MICB	--- /// MFAP2	MUC1
HNRNPA0	EIF2AK1	MSR1	GJB2	CNTNAP2
N4BP2L2	PIP5K1A	GPX8	PXDN	CDH11
ATP5J2	TUBA1C	NUDT3	MUC1	COL12A1
PDHA1	SNX14	PALLD	C1QTNF3	MMP13
BTNL9	HTATSF1	SKIL	COL12A1	SERPINA1
GMPR2	CELF1	LASS2	PLAU	MMP11
HSPA1A /// HSPA1B	KIAA2026	SMAD2	SPOCK1	COL1A1
DDX17	IL1B	RUFY1	COL3A1	NPNT
MAML3	VEZF1	TANC1	INHBA	SULF1
LIMCH1	MIA3	GIPC1	POSTN	
HNRNPA0	TGFB3	TNFAIP6	RAB31	
PIK3R1	S100A13	PDXDC1	CD276	
HSPD1	GMFB	FN1	SLITRK6	
SAP18	RNF144B	HNRNPUL2	RAB31	
MAT2A	IPPK	GPSM2	ITGB8	
UBA7	ACTR2	ETNK1	FBN1	
DTWD1	GK	PIGX	COL1A2	
PILRB	PPPDE1	DYNLT3	COL12A1	
STAT5B	FN1	FAM36A	SULF1	
PPP3R1	RUNX2	TCF12	GAS7	
EIF3CL /// EIF3C	TBCK	MLLT4	VCAN	
RPL28	ITPK1	PXDN	MFHAS1	
NSMCE4A	FARP1	SPP1	ENTPD7	
C15orf52	PREPL	CERCAM	MFAP5	
HNRNPF	SH3RF1	HTRA3	COL3A1	
EIF3F	TAX1BP3	FNDC1	LUM	
PPP3R1	C18orf1	MYO6	VAV3	
HNRNPAB	STK3	GREM1	COL6A3	
N4BP2L2	POSTN	MS4A7	CD109	
TRIM28	ZNF207	PTPLB	CACNG4	
IRX2	PXDN	CDH11	RAB31	



The cell colour is determined by the best gene rank percentile as given in the Oncomine™ website.

of Fibronectin gene due to its role in ECM. The role of Fibronectin in MBC was then evaluated by determining its expression in TMA's constructed from FFPE tissue blocks.

Table 16: Genes involved in ECM remodelling

Gene symbol	Gene name
FN1	Fibronectin 1
COL12A1	Collagen, Type XII, alpha 1
COL1A1	Collagen, Type I, alpha 1
COL1A2	Collagen, Type I, alpha 2
COL3A1	Collagen, Type III, alpha 1
COL6A3	Collagen, Type VI, alpha 3
COL10A1	Collagen, Type X, alpha 1
COL11A1	Collagen, Type XI, alpha 1
COL8A1	Collagen, Type VIII, alpha 1
SPARC	Secreted Protein, Acidic, Cysteine-Rich (Osteonectin)
FBN1	Fibrillin 1
ITGB8	Integrin beta8
MMP11	Matrix Metallopeptidase 11(Stromelysin 3)
MMP13	Matrix Metallopeptidase 13
SERPINA1	Serpin Peptidase Inhibitor (alpha-1 antitrypsin)

4.4.1 Immunohistochemistry Evaluation of Fibronectin Expression

The IHC analysis was performed in MBC TMA's and details of the TMA's used are described in Chapter 2 (Methodology section 2.3.2, Page no. 29-31). The only difference being the exclusion of TMA 7 (Canadian group) due to core exhaustion. Therefore, the descriptive statistics, measure of association and correlation analysis with known prognostic markers were performed in 208 patients. Whereas, the 5 year survival analysis was performed in 199 patients in whom the follow-up data was complete. The clinical and salient pathological characteristics of the cohort are as described earlier (Chapter 2, Result section 2.4, Page no. 41-43).

The literature was reviewed to establish various antibodies and scoring methods used for IHC analysis of FN in FFPE tissue (Tavakoli et al., 2011, Sudo et al., 2013, Swiatoniowski et al., 2005, Ioachim et al., 2002). In addition to this The Human Protein Atlas™ website was searched to identify other validated antibodies. After careful consideration, a rabbit polyclonal antibody against FN was selected for IHC evaluation in the MBC cohort (Product No: HPA 027066, Atlas Antibodies, SIGMA Life science, St. Louis, USA).

Even though the FN antibody has been independently validated, its specificity for FN was determined using Western Blot method. The specificity of the antibody to bind with FN was determined in transformed immortalised fibroblasts (Verghese et al., 2011), MCF-7 and MDA-231 cell lines using Western Blot (Figure 30a). The results showed a band corresponding to the molecular weight of Fibronectin (440kDa) in the fibroblast cell line. Moreover

there was no reactivity seen with MCF-7 or MDA-231 cell lines. This finding supported the fact that, FN is not routinely expressed by the breast cancer epithelial cells (Taylor-Papadimitriou et al., 1981, Gould et al., 1990, Friedman et al., 1984).

The antibody optimisation was performed initially in full section of female normal and breast cancer tissue prior to testing in the control TMA. The antibody produced optimal staining at 1:50 concentration using the envision method. The FN expression in the breast cancer stroma was only determined. The intensity of the FN immunoreactivity was determined and graded as negative, mild, moderate and strong. The FN immunoreactivity was specific to the stroma but occasional cancer epithelial cell staining was observed (Figure30 b-d).

The FN scoring was performed by two investigators (in 1/3rd of the cases) to determine the reproducibility of the scoring method (Dr S Sundara Rajan and Dr E T Verghese). The statistical analysis showed good agreement between the investigators (Table 17) ($\kappa = 0.68$). It was evident that most of the discrepancies were between mild to moderate and moderate to strong scoring grades.

Table 17: Representation of fibronectin immunohistochemistry scoring by two independent investigators

		Investigator 1				
	Fibronectin Scoring	**Negative**	**Mild**	**Moderate**	**Strong**	**Total**
Investigator 2	**Negative**	3	0	0	0	**3**
	Mild	1	23	1	0	**25**
	Moderate	0	5	15	0	**20**
	Strong	0	0	6	18	**24**
	Total	**4**	**28**	**22**	**18**	**72**

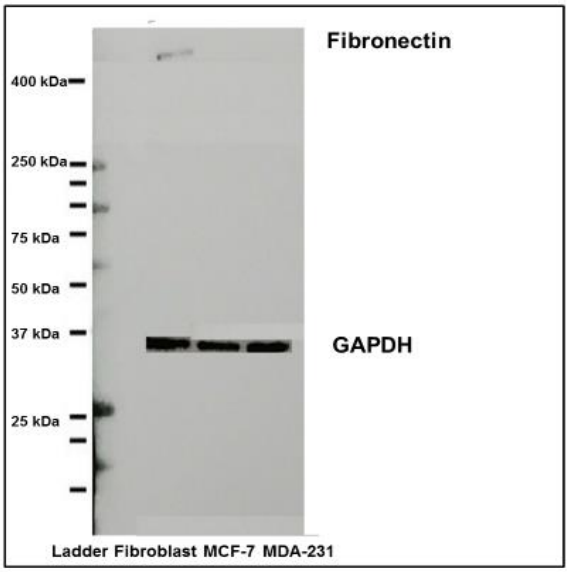


Figure. 30a

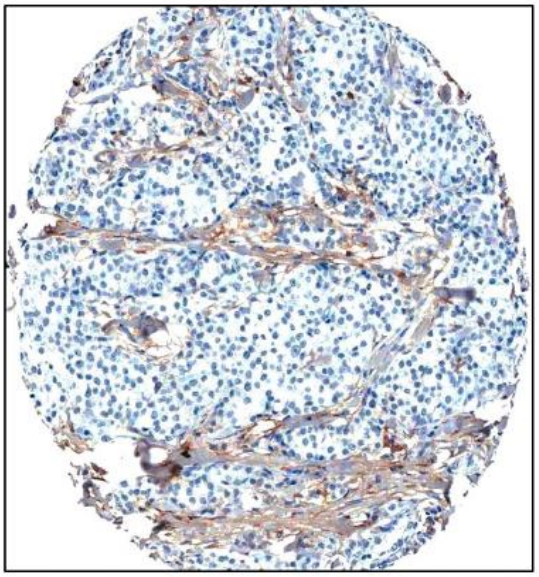


Figure. 30b

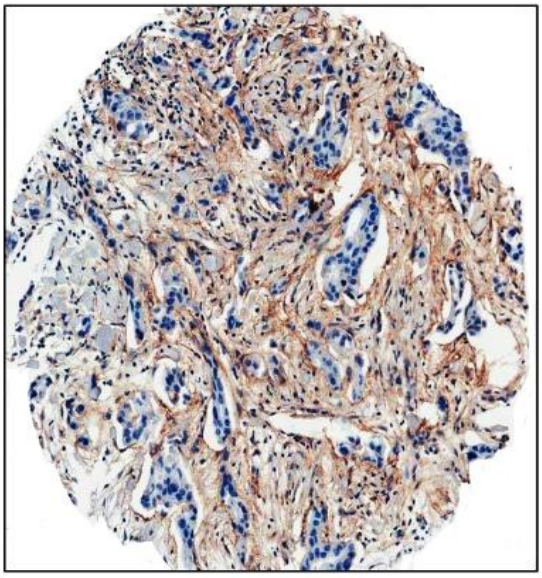


Figure. 30c

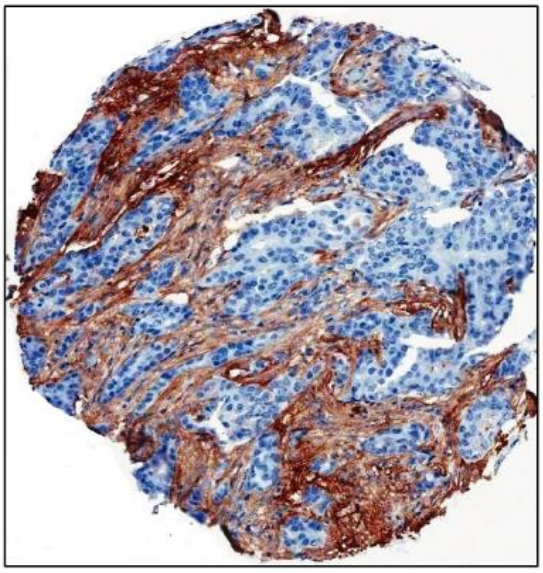


Figure. 30d

Figure 30: Specificity of anti-FN antibody & fibronectin expression in MBC TMA cores

(a) Western blot showing specificity of anti-FN antibody (HPA 027066). A single band within 10% of predicted molecular weight of FN (440 kDa) was observed in transformed fibroblasts and no bands were identified in epithelial cell lines (MCF-7 and MDA-231). GAPDH was used as the loading control. Mild (b) Moderate (c) and Strong (d) stromal staining observed with anti-Fibronectin antibody in the male breast cancer TMA's.

There was a clear distinction between those scored as negative or mild to those scored as moderate or high. Hence statistical analysis was performed by dichotomising the FN IHC score into high expression / positive (moderate and strong scores) and low expression / negative (negative and low scores). The result of the FN scoring was available in 190 MBC cases. In the MBC cohort, 114 cases (60%) were FN positive and the rest were negative (n =76 and 40%). Fibronectin showed a strong positive association with nodal status ($\chi^2 = 7.78$; $p = 0.02$). Fibronectin didn't show any statistically significant association with other known tumour specific prognostic indicators. There was a weak association between FN and AR ($\chi^2 = 3.732$; $p = 0.054$). However fibronectin didn't show any statistically significant association with other known prognostic biomarkers (hormonal biomarkers, cytokeratins, Ki67 and HER2).

The 5 year overall survival analysis was performed using Kaplan-Meier survival curve. There was a definite trend towards poor 5 year OS with positive FN expression ($p = 0.06$) (Figure 31). On Cox univariate regression analysis, FN showed only a trend towards predicting poor OS at 5 years (OR = 5.476 (CI = 0.75 to 40.2); $p = 0.09$). In keeping with the results of the univariate analysis, FN was not found to be an independent predictor of OS on multivariate analysis.

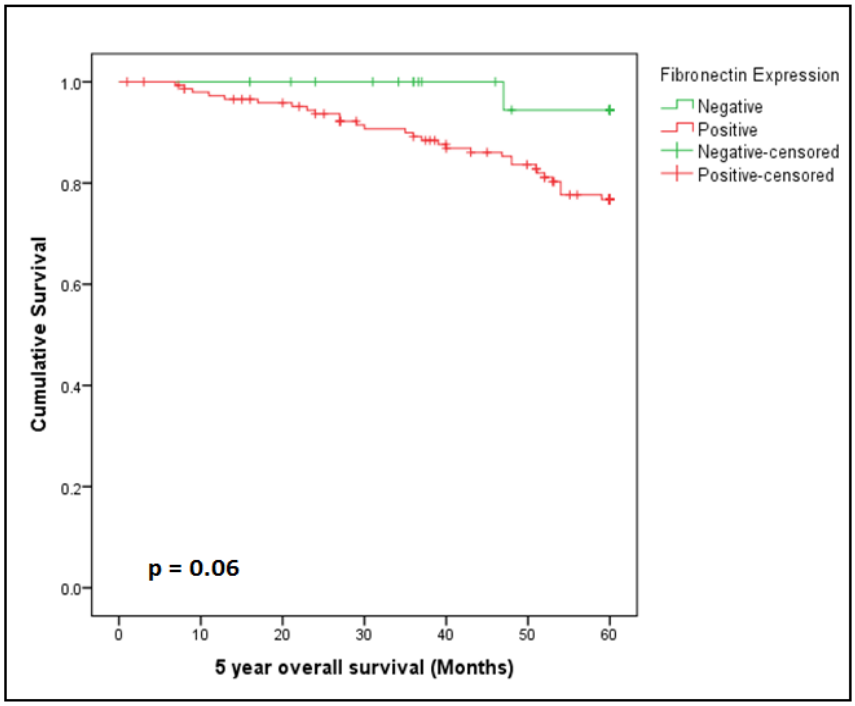


Figure. 31

Figure 31: Kaplan-Meier survival curve at 5 years for Fibronectin in Male Breast Cancer

The survival curve shows a trend towards poor 5 year overall survival with positive Fibronectin expression. Patients were censored when they ceased to be followed up for any reason but had not died due to breast cancer.

4.5 DISCUSSION

The heterogeneous nature of FBC is well established and has got prognostic significance (Hu et al., 2006, Sorlie et al., 2003, Perou et al., 2000). Compared to FBC, there is a paucity of studies in MBC to understand the molecular biology. Recently, Johansson et al (2012) has classified MBC into two distinct sub-groups luminal M1 and M2 (Johansson et al., 2012). They correlated well with the previously described male complex and simple sub-groups which were identified using aCGH (Johansson et al 2011). The luminal M1 and M2 sub-groups were unique and different from that of the established FBC sub-types (Giordano et al., 2004). However there were also similarities noted between male and female breast cancer suggesting that the fundamental genes involved in the disease process might be similar (Johansson et al., 2011, Rudlowski et al., 2006, Tirkkonen et al., 1999).

The bidirectional cross talk between the breast epithelial and stromal cells has been implicated in normal breast growth and differentiation as well as in the progression of malignancy (Schedin and Hovey, 2010). Stroma is the major component of the adult human breast (connective tissue 60% and fat 20%) with only 20% constituting for the epithelial elements (Howell et al., 2005, Wiseman and Werb, 2002). The stroma can be either inter-lobular or intra-lobular and supports both the vascular and lymphatic system as well provides scaffolding for the growth of the epithelial cells. The major component of the Inter-lobular stroma is collagen type 1, followed by collagen type 2 both making the inter-lobular stroma denser (Polyak and Kalluri, 2010). Fibronectin was found to be

more abundant in the normal inter-lobular stroma compared to intra-lobular stroma using immunofluorescence method (Ferguson et al., 1992).

There are pertinent differences in the development of male and female breast. The female breast ductal system undergoes rapid allometric growth and differentiation from puberty onwards under the stimulation of oestradiol resulting in the formation of the terminal ductal lobular units (TDLU) (Pelekanou and Leclercq, 2011). Conversely in the male breast there is minimal glandular formation with abrupt termination of the terminal ducts. Hence there is very little intra-lobular stroma compared to female breast. The inter-lobular stroma may be variable and may be comparable to female breast but less in volume. Hence it would be reasonable to hypothesise that there will be inherent differences in the development and progression of MBC compared to FBC.

It has been shown in FBC that, substantial changes occur in the ECM during normal growth and development. However during progression of normal breast to malignancy, only minimal gene expressional changes occur in the inter-lobular stroma compared to the epithelial compartment, where most of the changes seem to be taking place (Vargas et al., 2012). Most of the studies in this field were conducted using intra-lobular stroma and identified the changes were occurring mainly in the epithelial compartment (Knudsen et al., 2012, Ma et al., 2009). There is no published evidence within the literature on the changes which occur in the stromal compartment during progression from normal male breast to malignancy.

The results of the GEP of matched male and female breast cancer patients showed 656 differentially expressed genes (after exclusion of the sex-linked genes). Most of the genes (n = 487) were up-regulated. The results of Gene Ontology, individual functions of the top 10 up-regulated genes and hierarchical cluster group analysis showed the role of differentially expressed genes in ECM. The main themes emerging from the GO analysis were cell adhesion, cell-substrate adhesion, cell-matrix adhesion and collagen fibril organization (Chapter 4, Result section 4.4, Table 11, Page no. 125). The analysis undertaken by Johansson et al (2012) in MBC also showed similar GO biological process in the luminal M1 subgroup.

Subsequent analysis and comparison with publically available data sets of normal vs. FBC stroma identified 64 common differentially expressed genes (Chapter 4, Result section 4.4, Table.14, Page no. 130). Similarly there were 134 up-regulated and 39 down-regulated common genes, when this cohort was compared with normal vs. FBC GEP results available in the public repositories (Chapter 4, Result section 4.4, Table.15, Page no. 132). Together there were a total of 15 common differentially expressed genes identified from this analysis (all of them up-regulated in MBC) that was shown to have a role in the ECM during cancer progression. Most of these genes were involved in secreting proteins responsible for the production (COL12A1, COL1A1, COL1A2, COL3A1, COL6A3, COL10A1, COL11A1, COL8A1 and FBN1) or degradation (SPARC, MMP11, MMP13 and SERPINA) of ECM. Both FN1 and ITGB8 were unique in that the former facilitates cell-cell / cell-matrix adhesion and the latter is a transmembrane receptor that bridges cell-cell and cell-matrix interaction.

The FN1 gene encodes for fibronectin, which has unique functions in the ECM (Hynes and Yamada, 1982) and epithelial mesenchymal transition (EMT) (Park and Schwarzbauer, 2014, Foroni et al., 2012, Sung et al., 2011). Fibronectin is a large molecular weight glycoprotein (440kDa) present in the plasma (plasma fibronectin) and connective tissue (cellular fibronectin) (Hynes and Yamada, 1982). There are structural and functional similarities and difference amongst cellular and plasma FN (Hynes and Yamada, 1982). The plasma fibronectin is a dimer compared to the cellular fibronectin, which can be either a dimer or a polymer. The main functional difference is the more active role played by the cellular fibronectin in transformed cells and in haem-agglutination (Hynes et al., 1978, Yamada and Kennedy, 1979).

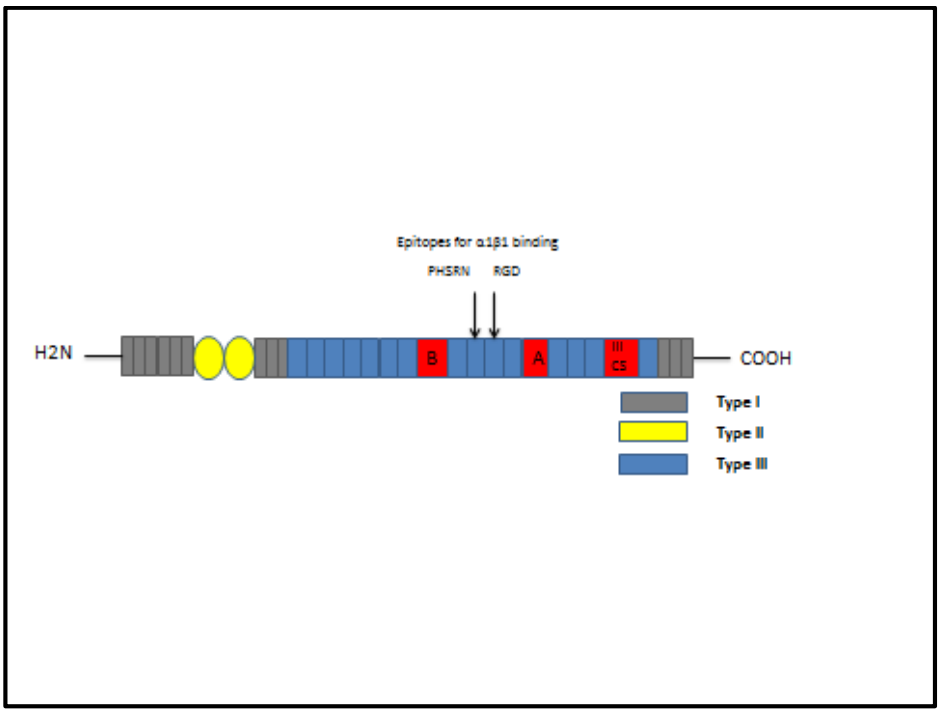


Figure 32: Model of fibronectin monomer

The three different repeating units are represented in different colours. The disulphide bond at the carboxyl terminal joins the monomers. The alternative splicing of the RNA at three repeats (represented in red box) results in EDA, EDB and III CS splice variants.

FN molecule is a heterodimer of similar but not identical polypeptides connected together by disulphide bonds (Kornblihtt et al., 1985) (Figure 32). The diversity of FN protein has been attributed to the alternating splicing occurring in at least 3 regions, extra domain A (EDA), extra domain B (EDB) and type III homology connecting segment (IIICS) in the pre-mRNA (Gutman and Kornblihtt, 1987, Zardi et al., 1987, Schwarzbauer et al., 1987, Kornblihtt et al., 1984). The inclusion or exclusion of these regions is tissue and species specific (Schwarzbauer et al., 1987, Schwarzbauer et al., 1983). The alternative splicing in these 3 regions can result in the formation of various known splice variants of FN.

The EDB sequence of the FN is almost undetectable in normal adult tissue and is expressed more during tissue remodeling, i.e. embryogenesis in foetal tissues, wound healing and during angiogenesis in tumours (Carnemolla et al., 1989, Ffrench-Constant et al., 1989). There are two types of oncofoetal FN reported within the literature, the one generated by alternative splicing of EDB of the FN pre-mRNA and the other generated by O-glycosylation in the splicing region of type III CS (Carnemolla et al., 1989). The oncofoetal FN (i.e. EDB-FN and IIICS sequence FN) is not expressed in normal breast tissue and found to be expressed in the tumour vessels and stroma (Kaczmarek et al., 1994). The oncofoetal FN isoforms also seems to be differentially distributed in the neoplastic tissues i.e. the EDB-FN present more in the tumour associated vessels and the IIICS sequence FN in the extracellular matrix tumour stroma (Kaczmarek et al., 1994).

There are two mechanisms described within the literature explaining the synthesis of fibronectin. A positive feedback resulting from an increase in stiffness of the ECM by FN will lead to increase in production of the FN by the epithelial cells (Williams et al., 2008). There is also a paracrine mechanism of regulation of FN synthesis by breast stromal cells through hormonal (oestrogen and/or progesterone) regulation of mammary epithelial cells (Woodward et al., 2001, Ferguson et al., 1992).

There is evidence within the literature supporting the role of FN in promoting cell matrix adhesion and/or spreading in a variety of mediums including collagen, gelatine and fibrin (Grinnell et al., 1980, Grinnell, 1978). FN acts as a mediator for the adhesion of fibroblasts to collagen (Pearlstein, 1976). In cell culture studies, it was observed that adult human fibroblasts attaches to FN in the presence of either type 1 or 4 collagen (Terranova et al., 1986) and laminin can competitively block FN binding to type 4 collagen (Terranova et al., 1986). It has been proposed that such competition between laminin and FN may be contributory in preventing the fibroblasts from invading the epithelial side of the basement membrane (Yamada, 1991). This could also explain the lack of FN in the basement membrane of adult tissue as the normal tissue matures from embryonic state with the presence of FN in the basement membrane.

Fibronectin can influence cell differentiation and morphogenesis (Yamada, 1991). FN is essential for the developmental branching (morphogenesis) of salivary gland, lung and kidney through facilitating cleft formation by conversion of cell-cell adhesion to cell-matrix adhesion (Sakai et al., 2003). Similarly FN was shown to be an essential effector for the breast acinar morphogenesis

(Williams et al., 2008). Fibronectin was down regulated during the normal acinar morphogenesis and over expression in MCF-10A cells results in increase in the acinar size with loss of the acinar lumen (Williams et al., 2008).

The loss of cell surface FN has been implicated in the tumourigenic potential of various transformed and untransformed cell lines (Chen et al., 1979, Gallimore et al., 1977). Considering FN role in cell adhesion and the inherent invasive properties of cancer cell makes it a plausible argument for the loss of FN from cell surface during oncogenic transformation. The reason for the loss appears to vary from reduced production (Hynes et al., 1978, Olden and Yamada, 1977), reduced binding (Wagner et al., 1981, Olden and Yamada, 1977, Hayman et al., 1981, Hynes et al., 1978, Vaheri and Ruoslahti, 1975) and increased degradation (Olden and Yamada, 1977). However it is yet to be consistently validated in animal or human cancer tissues.

Both the cellular binding domain (CBD) and the N-terminal domain of FN play an important role in angiogenesis, with the former playing a more critical role (Kim et al., 2000). The angiogenic potential of FN depends on the presence of an intact CBD which can bind to $\alpha_5\beta_1$ integrin and growth factor stimulation (Kim et al., 2000). The $\alpha_5\beta_1$ integrin is very selective for fibronectin as it requires the specific peptide sequences on the ninth (PHSRN) and tenth (RGDS) type III repeats of fibronectin for ligand recognition (Aota et al., 1994, Hynes, 1992). The selective inhibition of $\alpha_5\beta_1$ integrin can lead to inhibition of tumour growth and angiogenesis (Kim et al., 2000).

Fibronectin is thought to be up-regulated during the EMT process (especially type 2 and type 3 EMT) and considered as a marker for detection of EMT phenotype of cancer (Foroni et al., 2012, Sung et al., 2011, Yang et al., 2007, Strutz et al., 2002). There are two proposed mechanisms through which FN promotes changes during EMT. The first one is through activation of Stat3 (Signal transducer and activator of transcription 3), which regulates cell proliferation, growth and survival (Idowu et al., 2012, Marotta et al., 2011). FN activates Stat3 in an Epidermal Growth Factor Receptor (EGFR) independent manner through FAK/Pyk2:Jak2 pathway in breast cancer cell lines (MDA-MB21) (Balanis et al., 2013). In MCF-10A cell lines, FN was shown to activate Smad2 independently as well as through TGFβ1 (Transforming Growth Factor β1) to promote epithelial changes during EMT. It is entirely possible that, the up-regulation of FN observed during EMT could be both a cause and result of tumour progression.

In the normal breast tissue there is conflicting reports within the literature about the expression of FN. Fibronectin is detected in small quantities or absent in the normal mammary tissue (Ioachim et al., 2002, Helleman et al., 2008, Kadar et al., 2002, Koukoulis et al., 1993, Christensen, 1992). The studies in which FN expression was seen in normal mammary tissue it was confined mainly to the stroma and also in the basement membrane region of the epithelial cells (Friedman et al., 1984, Stampfer et al., 1981). Gould et al (1990) showed that, FN expression was faint or mostly absent in the intra-lobular stroma and seen mainly in the inter-lobular stroma (Gould et al., 1990) . Amongst the various splice variants of FN, only the EDA of FN was found in adult breast ducts and acini (Koukoulis et al., 1993).

However more extensive expression of FN was observed in breast cancer. The distribution of FN was mainly around the peri-tumoural stroma (Gould et al., 1990, Friedman et al., 1984, Stampfer et al., 1981) and intra-tumoural stroma (Gould et al., 1990). The breast cancer epithelial cells were mostly negative (Gould et al., 1990, Friedman et al., 1984), however some have reported minimal staining within the cytoplasmic region of the epithelial cells (Christensen et al., 1985, Stampfer et al., 1981). All three known splice variants of FN are expressed in breast cancer, with expression of EDB FN and oncofoetal FN restricted to breast cancer and foetal breast tissue (Koukoulis et al., 1993). The FN expression was stronger and more extensive in invasive ductal cancer compared to invasive lobular cancer (Gould et al., 1990). In summary, the expression of FN is minimal in normal breast tissue, but enhances through the spectrum of fibrocystic and hyperplastic breast tissue with most extensive staining noted in the stromal elements of breast cancer (Ioachim et al., 2002, Gould et al., 1990).

Even though there is emerging evidence regarding epithelial FN in FBC, only the stromal expression was evaluated in the MBC using IHC. This was due to the following reasons, a) IHC was performed to evaluate the stromal expression FN and relate it to the GEP data mining results identified in matched male and female breast cancer samples b) there is ambiguity about the role of epithelial FN in breast cancer c) the FN expression was patchy within the tumour epithelial cells and was absent in most of the cases reviewed in this cohort.

The FN immunoreactivity was specific to stroma and around 60% of the cases were positive for fibronectin. FN did not show any association with proliferative

markers (like Ki67 and Survivin) or tumour grade in this cohort. Similarly other than a weak association with AR, there was no significant association with known prognostic biomarkers (ER, CK5/6, CK18 and HER2).

There was a strong positive association between FN and nodal status. Helleman et al (2008) also found a similar strong association between nodal status and FN gene expression in FBC patients. Similarly in paired breast tumour and lymph node metastatic samples, the FN protein expression was elevated in the ECM of the lymph node metastatic samples (Hao et al., 2004). However, the FN gene was under expressed in the lymph node metastasis (Hao et al., 2004). Hence, Hao et al (2004) proposed that "there might be a negative feedback of FN gene expression by the secreted protein or the FN protein being secreted by other cells into the ECM". Ioachim et al (2002) showed that there was cytoplasmic staining of FN in only 10% of the primary tumour cells compared to 50% in the involved lymph node metastasis. Extrapolation of the findings by Hao et al (2004) and Ioachim et al (2002) would suggest a paracrine mechanism leading to increased expression of FN in metastatic lymph nodes.

The prognostic role of FN thus far has been evaluated only in FBC. Generally in FBC, FN expression (gene or protein) has been associated with a worse clinical outcome (Ioachim et al., 2002, Helleman et al., 2008, Bae et al., 2013, Yao et al., 2007). In this MBC cohort there was a definitive trend towards poor 5 year OS amongst FN positive patients. However on multivariate analysis, FN expression was not found to be an independent predictor of survival in the MBC cohort. Similarly even in FBC, stromal expression of FN was not proven to be

an independent predictor of survival. Conversely epithelial FN expression was found to be an independent predictor of worse survival in FBC patients (Bae et al., 2013). The other published studies on the prognostic role of FN protein expression were biased as in one a less reliable IHC technique (immunoperoxidase method) was used (Christensen et al., 1988) and in the other plasma FN was analysed instead of cellular FN (Takei et al., 1995).

4.6 CONCLUSION

The gene expression profiling in matched male and female breast cancer samples showed up-regulation of genes involved in ECM synthesis (COL12A1, COL1A1, COL1A2, COL3A1, COL6A3, COL10A1, COL11A1, COL8A1 and FBN1) and degradation (SPARC, MMP11, MMP13 and SERPINA), a trans-membrane receptor (ITGB8) and a gene involved in ECM re-modelling (FN1). The result was validated by evaluating fibronectin expression in MBC tissue microarrays. There was high stromal expression of fibronectin in 60% of the MBC cases. High fibronectin expression was also found to have a strong association with nodal status and a trend towards poor survival in the MBC cohort.

Changes occurring in the ECM during tumour progression is complex with interaction between various ECM molecules and their receptors, as well as ECM based recruitment and organisation of other molecules. Hence the role or changes induced by a single gene or protein may not be substantial. However, single gene studies are essential in bridging the gap in knowledge to better understand the larger role played by ECM during tumour progression. Examining the wider role of ECM genes identified in this study is on-going within the research group.

4.7 LIMITATIONS AND FUTURE DIRECTION

A few limitations of GEP of matched male and female breast cancer samples and validation of the results in MBC TMA's should be considered while interpreting and extrapolating the results of this research.

The mRNA was extracted from sections made from FFPE tumour rich MBC samples. In this study, neither macro nor micro-dissection was performed to extract tissues separately for GEP from cancer epithelial cells and intra or inter-lobular stroma. This would have been ideal to establish true changes taking place in each compartment of the MBC. However H&E staining of the section was evaluated by a breast pathologist (AMH) to ensure that each FFPE section used for mRNA extraction contained sufficient amount of cancer and stromal elements.

The GEP was performed with an aim to identify differentially expressed genes in matched male and female breast cancer patients. Hence differentially expressed genes identified might not represent the molecular changes occurring in male breast during progression from normal to malignancy. An attempt was made to extrapolate this finding by comparing with publically available FBC data sets to identify common differentially expressed genes during cancer development and progression. Hence the results derived from data mining needs further independent validation to improve its external validity.

The use of tissue microarray for studying stromal expression of FN may not be ideal. A tissue microarray would be inadequate to study the advancing edge of a tumour, where FN expression is proposed to be highest due to its role in ECM

remodelling (Ioachim et al., 2002). Similarly it would be difficult to quantify and evaluate the expression of FN in inter and intra-lobular stroma in a TMA to confirm or refute the findings drawn from studies conducted in FBC (Gould et al., 1990). However there are many advantages for using TMAs in MBC, including evaluation of large number of samples in a systematic way by using standardised IHC methods (Ilyas et al., 2013).

There is a vast amount of data generated through the GEP analysis. The differentially expressed genes with a role in ECM remodelling (Table 16; Chapter 4, Section 4.4, Page no. 133) would be an ideal start point for further translational research. The protein end products of the genes identified can be evaluated for prognostic or predictive role in MBC cohort using immunohistochemistry. The GO and IPA results provide another platform for further research. The GO analysis has shown (Table 11; Chapter 4, Result section 4.4, Page no. 125) enrichment of cellular adhesion and collagen fibril organization amongst the top 10 enriched biological process. The role of these biological process and pathways in the development and progression of MBC could be evaluated in future studies. The top 10 up and down regulated protein encoding non-sex linked genes identified during the data mining provides another platform for pursuing further research (Table 12 and 13, Chapter 4, Section 4.4, Page no. 129). Each of these genes and their end products could unveil details that could explain the natural progression of MBC.

A future aim should be to perform GEP in matched male normal and breast cancer tissue samples. However this will be challenging due to the difficulties in obtaining normal male breast tissue for research purpose. A compromise may

be to use breast tissue with gynaecomastia changes, although it might also not be truly representative of normal breast milieu. However gynaecomastoid tissue has been used as a surrogate for normal male breast tissue for evaluating the microRNA expression (Lehmann et al., 2010, Fassan et al., 2009). Careful consideration should also be given regarding the proliferative changes occurring in the glandular and/or connective tissue of the gynaecomastoid breast while interpreting GEP results if used as surrogate for normal male breast tissue. Another option would be to collect blood samples from healthy male volunteers as controls (germline DNA) and compare it with tumour DNA to identify the differential expression of genes.

Future studies should also aim to micro-dissect the tumour to better understand the role played by tumour microenvironment in the development and progression of MBC. This will enable us to understand the changes that is taking place in the epithelial and stromal compartments of MBC compared to normal or gynaecomastoid breast tissue. This will also facilitate direct comparison with FBC gene expression profiles available through the public repositories.

The external validity of the study was compromised due to the missing data and lack of information about disease free survival (DFS). The lack of adequate power could have led to type II error. This may have been one of the contributing factors for the inability to detect independent predictability amongst various biomarkers tested in the MBC cohort. In 2012, there were only 353 cases of MBC diagnosed in the UK, which accounts for less than 1% of all the cancers diagnosed in men (CRUK, 2012). When dealing with a disease of such

low incidence, it is often difficult to achieve sufficient statistical power to provide answers to all relevant research questions. Despite best effort by the team, there was significant amount of missing clinical, pathological and survival data (40%) in this cohort. However even after exclusion of cases without adequate clinical, pathological and follow-up data, there was reasonable sample size (n = 238) available for statistical analysis. The sample size of the current study remains to be the one of the largest amongst those published within the literature.

The survival analysis was performed using OS in this cohort with a median age of 68 years (IQR = 17 years). Hence the results of survival analysis might have been affected by analytical bias. The information about DFS could have been ideal; however it was difficult to obtain as the samples were collected through collaboration from multiple international centres over many years with some of the data originating as far as back in the early 90's. There was also lack of adequate and consistent adjuvant treatment data from most of the centres. Hence it has not been possible to account for the influence of various adjuvant treatments during the statistical analysis. Even though no major breakthroughs or innovative therapeutic advances have emerged in adjuvant therapy of MBC over recent years, accounting for this confounding factor could have improved the internal and external validity of this study.

In spite of the limitations discussed, this study is one of the largest cohort of MBC evaluated for the expression and prognostic role of various biomarkers. The results have shown promising insight into MBC biology at molecular and protein levels. It is essential to further pursue both the biomarker and gene

expression findings deduced from this study to better understand MBC biology. The limitations and areas for improvement identified from this study should lay foundations for conducting future studies in this field. The rarity of MBC necessitates the need for seeking an international collaboration or consortium (Korde et al., 2010) to collect MBC samples in a systematic way to facilitate future research.

REFERENCES

ABREU, M. H., AFONSO, N., ABREU, P. H., MENEZES, F., LOPES, P., HENRIQUE, R., PEREIRA, D. & LOPES, C. 2016. Male breast cancer: Looking for better prognostic subgroups. *Breast,* 26**,** 18-24.

ADAMI, H. O., HOLMBERG, L., MALKER, B. & RIES, L. 1985. Long-term survival in 406 males with breast cancer. *Br J Cancer,* 52**,** 99-103.

AHN, S. H., HWANG, U. K., KWAK, B. S., YOON, H. S., KU, B. K., KANG, H. J., KIM, J. S., KO, B. K., KO, C. D., YOON, K. S., CHO, D. Y., KIM, J. S. & SON, B. H. 2004. Prevalence of BRCA1 and BRCA2 mutations in Korean breast cancer patients. *J Korean Med Sci,* 19**,** 269-74.

AKBAR, S., JORDAN, L. B., PURDIE, C. A., THOMPSON, A. M. & MCKENNA, S. J. 2015. Comparing computer-generated and pathologist-generated tumour segmentations for immunohistochemical scoring of breast tissue microarrays. *Br J Cancer,* 113**,** 1075-80.

AL-JOUDI, F. S., ISKANDAR, Z. A., HASNAN, J., RUSLI, J., KAMAL, Y., IMRAN, A. K., AHMED, M. & ZAKARIA, J. 2007. Expression of survivin and its clinicopathological correlations in invasive ductal carcinoma of the breast. *Singapore Med J,* 48**,** 607-14.

ALBO, D., AMES, F. C., HUNT, K. K., ROSS, M. I., SINGLETARY, S. E. & KUERER, H. M. 2003. Evaluation of lymph node status in male breast cancer patients: a role for sentinel lymph node biopsy. *Breast Cancer Res Treat,* 77**,** 9-14.

ALLRED, D. C., HARVEY, J. M., BERARDO, M. & CLARK, G. M. 1998. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod Pathol,* 11**,** 155-68.

ANDERSEN, J. A. & GRAM, J. B. 1982. Male breast at autopsy. *Acta Pathol Microbiol Immunol Scand A,* 90**,** 191-7.

ANDERSON, J., REDDY, V. B., GREEN, L., BITTERMAN, P., BOROK, R., MAGGI-GALLUZZI, C., MONTIRONI, R., WICK, M., GOULD, V. E. & GATTUSO, P. 2002. Role of expression of cell cycle inhibitor p27 and MIB-1 in predicting lymph node metastasis in male breast carcinoma. *Breast J,* 8**,** 101-7.

ANDERSON, W. F., ALTHUIS, M. D., BRINTON, L. A. & DEVESA, S. S. 2004. Is male breast cancer similar or different than female breast cancer? *Breast Cancer Res Treat,* 83**,** 77-86.

ANELLI, A., ANELLI, T. F., YOUNGSON, B., ROSEN, P. P. & BORGEN, P. I. 1995. Mutations of the p53 gene in male breast cancer. *Cancer,* 75**,** 2233-8.

AOTA, S., NOMIZU, M. & YAMADA, K. M. 1994. The short amino acid sequence Pro-His-Ser-Arg-Asn in human fibronectin enhances cell-adhesive function. *J Biol Chem,* 269**,** 24756-61.

BADVE, S., TURBIN, D., THORAT, M. A., MORIMIYA, A., NIELSEN, T. O., PEROU, C. M., DUNN, S., HUNTSMAN, D. G. & NAKSHATRI, H. 2007. FOXA1 expression in breast cancer--correlation with luminal subtype A and survival. *Clin Cancer Res,* 13**,** 4415-21.

BAE, Y. K., KIM, A., KIM, M. K., CHOI, J. E., KANG, S. H. & LEE, S. J. 2013. Fibronectin expression in carcinoma cells correlates with tumor

aggressiveness and poor clinical outcome in patients with invasive breast cancer. *Hum Pathol,* 44**,** 2028-37.

BAGLEY, C. S., WESLEY, M. N., YOUNG, R. C. & LIPPMAN, M. E. 1987. Adjuvant chemotherapy in males with cancer of the breast. *Am J Clin Oncol,* 10**,** 55-60.

BALANIS, N., WENDT, M. K., SCHIEMANN, B. J., WANG, Z., SCHIEMANN, W. P. & CARLIN, C. R. 2013. Epithelial to mesenchymal transition promotes breast cancer progression via a fibronectin-dependent STAT3 signaling pathway. *J Biol Chem,* 288**,** 17954-67.

BALLERINI, P., RECCHIONE, C., CAVALLERI, A., MONETA, R., SACCOZZI, R. & SECRETO, G. 1990. Hormones in male breast cancer. *Tumori,* 76**,** 26-8.

BASHAM, V. M., LIPSCOMBE, J. M., WARD, J. M., GAYTHER, S. A., PONDER, B. A., EASTON, D. F. & PHAROAH, P. D. 2002. BRCA1 and BRCA2 mutations in a population-based study of male breast cancer. *Breast Cancer Res,* 4**,** R2.

BERGAMASCHI, A., TAGLIABUE, E., SORLIE, T., NAUME, B., TRIULZI, T., ORLANDI, R., RUSSNES, H. G., NESLAND, J. M., TAMMI, R., AUVINEN, P., KOSMA, V. M., MENARD, S. & BORRESEN-DALE, A. L. 2008. Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome. *J Pathol,* 214**,** 357-67.

BESIC, N., CERNIVC, B., DE GREVE, J., LOKAR, K., KRAJC, M., NOVAKOVIC, S., ZGAJNAR, J. & TEUGELS, E. 2008. BRCA2 gene mutations in Slovenian male breast cancer patients. *Genet Test,* 12**,** 203-9.

BLOOM, K. J., GOVIL, H., GATTUSO, P., REDDY, V. & FRANCESCATTI, D. 2001. Status of HER-2 in male and female breast carcinoma. *Am J Surg,* 182**,** 389-92.

BOLTON, K. L., GARCIA-CLOSAS, M., PFEIFFER, R. M., DUGGAN, M. A., HOWAT, W. J., HEWITT, S. M., YANG, X. R., CORNELISON, R., ANZICK, S. L., MELTZER, P., DAVIS, S., LENZ, P., FIGUEROA, J. D., PHAROAH, P. D. & SHERMAN, M. E. 2010. Assessment of automated image analysis of breast cancer tissue microarrays for epidemiologic studies. *Cancer Epidemiol Biomarkers Prev,* 19**,** 992-9.

BRATTHAUER, G. L., STAMATAKOS, M. D. & VINH, T. N. 2010. Cells with minimal expression of the JAK/STAT pathway related proteins STAT5a and the prolactin receptor: evidence of an alternate prolactin receptor isoform in breast disease. *Protein Pept Lett,* 17**,** 104-8.

BRENNAN, D. J., REXHEPAJ, E., O'BRIEN, S. L., MCSHERRY, E., O'CONNOR, D. P., FAGAN, A., CULHANE, A. C., HIGGINS, D. G., JIRSTROM, K., MILLIKAN, R. C., LANDBERG, G., DUFFY, M. J., HEWITT, S. M. & GALLAGHER, W. M. 2008. Altered cytoplasmic-to-nuclear ratio of survivin is a prognostic indicator in breast cancer. *Clin Cancer Res,* 14**,** 2681-9.

BRENNER, B., FRIED, G., LEVITZKI, P., RAKOWSKY, E., LURIE, H., IDELEVICH, E., NEUMAN, A., KAUFMAN, B., SULKES, J. & SULKES, A. 2002. Male breast carcinoma in Israel: higher incident but possibly prognosis in Ashkenazi Jews. *Cancer,* 94**,** 2128-33.

BRINTON, L. A., COOK, M. B., MCCORMACK, V., JOHNSON, K. C., OLSSON, H., CASAGRANDE, J. T., COOKE, R., FALK, R. T.,

GAPSTUR, S. M., GAUDET, M. M., GAZIANO, J. M., GKIOKAS, G., GUENEL, P., HENDERSON, B. E., HOLLENBECK, A., HSING, A. W., KOLONEL, L. N., ISAACS, C., LUBIN, J. H., MICHELS, K. B., NEGRI, E., PARISI, D., PETRIDOU, E. T., PIKE, M. C., RIBOLI, E., SESSO, H. D., SNYDER, K., SWERDLOW, A. J., EUROPEAN RARE CANCER STUDY, G., TRICHOPOULOS, D., URSIN, G., VAN DEN BRANDT, P. A., VAN DEN EEDEN, S. K., WEIDERPASS, E., WILLETT, W. C., EWERTZ, M. & THOMAS, D. B. 2014. Anthropometric and hormonal risk factors for male breast cancer: male breast cancer pooling project results. *J Natl Cancer Inst,* 106**,** djt465.

BRINTON, L. A., KEY, T. J., KOLONEL, L. N., MICHELS, K. B., SESSO, H. D., URSIN, G., VAN DEN EEDEN, S. K., WOOD, S. N., FALK, R. T., PARISI, D., GUILLEMETTE, C., CARON, P., TURCOTTE, V., HABEL, L. A., ISAACS, C. J., RIBOLI, E., WEIDERPASS, E. & COOK, M. B. 2015. Prediagnostic Sex Steroid Hormones in Relation to Male Breast Cancer Risk. *J Clin Oncol,* 33**,** 2041-50.

BRINTON, L. A., RICHESSON, D. A., GIERACH, G. L., LACEY, J. V., JR., PARK, Y., HOLLENBECK, A. R. & SCHATZKIN, A. 2008. Prospective evaluation of risk factors for male breast cancer. *J Natl Cancer Inst,* 100**,** 1477-81.

CALLAGY, G. M., PHAROAH, P. D., PINDER, S. E., HSU, F. D., NIELSEN, T. O., RAGAZ, J., ELLIS, I. O., HUNTSMAN, D. & CALDAS, C. 2006. Bcl-2 is a prognostic marker in breast cancer independently of the Nottingham Prognostic Index. *Clin Cancer Res,* 12**,** 2468-75.

CALLARI, M., CAPPELLETTI, V., DE CECCO, L., MUSELLA, V., MIODINI, P., VENERONI, S., GARIBOLDI, M., PIEROTTI, M. A. & DAIDONE, M. G. 2011. Gene expression analysis reveals a different transcriptomic landscape in female and male breast cancer. *Breast Cancer Res Treat,* 127**,** 601-10.

CARDOSO, F. Characterization of male breast cancer: First results of the EORTC 10085/TBCRC/BIG/NABCG International Male BC Program. San Antonio Breast Cancer Symposium, 2014 San Antonio, USA.

CAREY, A. H., WATERWORTH, D., PATEL, K., WHITE, D., LITTLE, J., NOVELLI, P., FRANKS, S. & WILLIAMSON, R. 1994. Polycystic ovaries and premature male pattern baldness are associated with one allele of the steroid metabolism gene CYP17. *Hum Mol Genet,* 3**,** 1873-6.

CAREY, L. A., PEROU, C. M., LIVASY, C. A., DRESSLER, L. G., COWAN, D., CONWAY, K., KARACA, G., TROESTER, M. A., TSE, C. K., EDMISTON, S., DEMING, S. L., GERADTS, J., CHEANG, M. C., NIELSEN, T. O., MOORMAN, P. G., EARP, H. S. & MILLIKAN, R. C. 2006. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA,* 295**,** 2492-502.

CARLSSON, G., HAFSTROM, L. & JONSSON, P. E. 1981. Male breast cancer. *Clin Oncol,* 7**,** 149-55.

CARNEMOLLA, B., BALZA, E., SIRI, A., ZARDI, L., NICOTRA, M. R., BIGOTTI, A. & NATALI, P. G. 1989. A tumor-associated fibronectin isoform generated by alternative splicing of messenger RNA precursors. *J Cell Biol,* 108**,** 1139-48.

CASAGRANDE, J. T., HANISCH, R., PIKE, M. C., ROSS, R. K., BROWN, J. B. & HENDERSON, B. E. 1988. A case-control study of male breast cancer. *Cancer Res,* 48**,** 1326-30.

CHAKRAVARTHY, A. & KIM, C. R. 2002. Post-mastectomy radiation in male breast cancer. *Radiother Oncol,* 65**,** 99-103.

CHEANG, M. C., CHIA, S. K., VODUC, D., GAO, D., LEUNG, S., SNIDER, J., WATSON, M., DAVIES, S., BERNARD, P. S., PARKER, J. S., PEROU, C. M., ELLIS, M. J. & NIELSEN, T. O. 2009. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst,* 101**,** 736-50.

CHEN, L. B., SUMMERHAYES, I., HSIEH, P. & GALLIMORE, P. H. 1979. Possible role of fibronectin in malignancy. *J Supramol Struct,* 12**,** 139-50.

CHIN, K., DEVRIES, S., FRIDLYAND, J., SPELLMAN, P. T., ROYDASGUPTA, R., KUO, W. L., LAPUK, A., NEVE, R. M., QIAN, Z., RYDER, T., CHEN, F., FEILER, H., TOKUYASU, T., KINGSLEY, C., DAIRKEE, S., MENG, Z., CHEW, K., PINKEL, D., JAIN, A., LJUNG, B. M., ESSERMAN, L., ALBERTSON, D. G., WALDMAN, F. M. & GRAY, J. W. 2006. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell,* 10**,** 529-41.

CHRISTENSEN, L. 1992. The distribution of fibronectin, laminin and tetranectin in human breast cancer with special attention to the extracellular matrix. *APMIS Suppl,* 26**,** 1-39.

CHRISTENSEN, L., NIELSEN, M., ANDERSEN, J. & CLEMMENSEN, I. 1988. Stromal fibronectin staining pattern and metastasizing ability of human breast carcinoma. *Cancer Res,* 48**,** 6227-33.

CHRISTENSEN, L., NIELSEN, M., HOLUND, B. & CLEMMENSEN, I. 1985. In vivo demonstration of cytoplasmic fibronectin in human breast carcinomas. *Virchows Arch A Pathol Anat Histopathol,* 407**,** 337-46.

CHUNG, H. C., KOH, E. H., ROH, J. K., MIN, J. S., LEE, K. S., SUH, C. O., KIM, K. E., LOH, J. J., LEE, K. B. & KIM, B. S. 1990. Male breast cancer--a 20-year review of 16 cases at Yonsei University. *Yonsei Med J,* 31**,** 242-50.

CIMMINO, V. M., DEGNIM, A. C., SABEL, M. S., DIEHL, K. M., NEWMAN, L. A. & CHANG, A. E. 2004. Efficacy of sentinel lymph node biopsy in male breast cancer. *J Surg Oncol,* 86**,** 74-7.

CIOCCA, V., BOMBONATI, A., GATALICA, Z., DI PASQUALE, M., MILOS, A., RUIZ-ORRICO, A., DREHER, D., FOLCH, N., MONZON, F., SANTEUSANIO, G., PEROU, C. M., BERNARD, P. S. & PALAZZO, J. P. 2006. Cytokeratin profiles of male breast cancers. *Histopathology,* 49**,** 365-70.

CLARK, J. L., NGUYEN, P. L., JASZCZ, W. B., JATOI, A. & NIEHANS, G. A. 2000. Prognostic variables in male breast cancer. *Am Surg,* 66**,** 502-11.

CONTRACTOR, K. B., KAUR, K., RODRIGUES, G. S., KULKARNI, D. M. & SINGHAL, H. 2008. Male breast cancer: is the scenario changing. *World J Surg Oncol,* 6**,** 58.

CREW, K. D., NEUGUT, A. I., WANG, X., JACOBSON, J. S., GRANN, V. R., RAPTIS, G. & HERSHMAN, D. L. 2007. Racial disparities in treatment and survival of male breast cancer. *J Clin Oncol,* 25**,** 1089-98.

CRUK. 2012. *Breast Cancer Incidence Statistics* [Online]. London. Available: http://www.cancerresearchuk.org/content/breast-cancer-incidence-statistics.

CUTULI, B., LACROZE, M., DILHUYDY, J. M., VELTEN, M., DE LAFONTAN, B., MARCHAL, C., RESBEUT, M., GRAIC, Y., CAMPANA, F., MONCHO-BERNIER, V. & ET AL. 1995. Male breast cancer: results of the treatments and prognostic factors in 397 cases. *Eur J Cancer,* 31A**,** 1960-4.

DELGALLO, W. D., RODRIGUES, J. R., BUENO, S. P., VIERO, R. M. & SOARES, C. T. 2010. Cell blocks allow reliable evaluation of expression of basal (CK5/6) and luminal (CK8/18) cytokeratins and smooth muscle actin (SMA) in breast carcinoma. *Cytopathology,* 21**,** 259-66.

DEVESA, S. S., BLOT, W. J., STONE, B. J., MILLER, B. A., TARONE, R. E. & FRAUMENI, J. F., JR. 1995. Recent cancer trends in the United States. *J Natl Cancer Inst,* 87**,** 175-82.

DOBERAUER, C., NIEDERLE, N. & SCHMIDT, C. G. 1988. Advanced male breast cancer treatment with the LH-RH analogue buserelin alone or in combination with the antiandrogen flutamide. *Cancer,* 62**,** 474-8.

DONEGAN, W. L., REDLICH, P. N., LANG, P. J. & GALL, M. T. 1998. Carcinoma of the breast in males: a multiinstitutional survey. *Cancer,* 83**,** 498-509.

DOWSETT, M., NIELSEN, T. O., A'HERN, R., BARTLETT, J., COOMBES, R. C., CUZICK, J., ELLIS, M., HENRY, N. L., HUGH, J. C., LIVELY, T., MCSHANE, L., PAIK, S., PENAULT-LLORCA, F., PRUDKIN, L., REGAN, M., SALTER, J., SOTIRIOU, C., SMITH, I. E., VIALE, G., ZUJEWSKI, J. A. & HAYES, D. F. 2011. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst,* 103**,** 1656-64.

DUNNING, A. M., HEALEY, C. S., PHAROAH, P. D., FOSTER, N. A., LIPSCOMBE, J. M., REDMAN, K. L., EASTON, D. F., DAY, N. E. & PONDER, B. A. 1998. No association between a polymorphism in the steroid metabolism gene CYP17 and risk of breast cancer. *Br J Cancer,* 77**,** 2045-7.

EFRON, B. 1987. Better Bootstrap Confidence Intervals. *Journal of the American Statistical Association,* 82**,** 171-185.

EL-TAMER, M. B., KOMENAKA, I. K., TROXEL, A., LI, H., JOSEPH, K. A., DITKOFF, B. A., SCHNABEL, F. R. & KINNE, D. W. 2004. Men with breast cancer have better disease-specific survival than women. *Arch Surg,* 139**,** 1079-82.

ERREN, T. C. 2001. A meta-analysis of epidemiologic studies of electric and magnetic fields and breast cancer in women and men. *Bioelectromagnetics,* Suppl 5**,** S105-19.

EVANS, D. B. & CRICHLOW, R. W. 1987. Carcinoma of the male breast and Klinefelter's syndrome: is there an association? *CA Cancer J Clin,* 37**,** 246-51.

EVANS, D. G., SUSNERWALA, I., DAWSON, J., WOODWARD, E., MAHER, E. R. & LALLOO, F. 2010. Risk of breast cancer in male BRCA2 carriers. *J Med Genet,* 47**,** 710-1.

EVANS, G. F., ANTHONY, T., TURNAGE, R. H., SCHUMPERT, T. D., LEVY, K. R., AMIRKHAN, R. H., CAMPBELL, T. J., LOPEZ, J. & APPELBAUM,

A. H. 2001. The diagnostic accuracy of mammography in the evaluation of male breast disease. *Am J Surg,* 181**,** 96-100.
EWERTZ, M. 1996. Epidemiology of breast cancer: the Nordic contribution. *Eur J Surg,* 162**,** 97-9.
FACKENTHAL, J. D., MARSH, D. J., RICHARDSON, A. L., CUMMINGS, S. A., ENG, C., ROBINSON, B. G. & OLOPADE, O. I. 2001. Male breast cancer in Cowden syndrome patients with germline PTEN mutations. *J Med Genet,* 38**,** 159-64.
FALCHETTI, M., LUPI, R., RIZZOLO, P., CECCARELLI, K., ZANNA, I., CALO, V., TOMMASI, S., MASALA, G., PARADISO, A., GULINO, A., GIANNINI, G., RUSSO, A., PALLI, D. & OTTINI, L. 2008. BRCA1/BRCA2 rearrangements and CHEK2 common mutations are infrequent in Italian male breast cancer cases. *Breast Cancer Res Treat,* 110**,** 161-7.
FASSAN, M., BAFFA, R., PALAZZO, J. P., LLOYD, J., CROSARIOL, M., LIU, C. G., VOLINIA, S., ALDER, H., RUGGE, M., CROCE, C. M. & ROSENBERG, A. 2009. MicroRNA expression profiling of male breast cancer. *Breast Cancer Res,* 11**,** R58.
FERGUSON, J. E., SCHOR, A. M., HOWELL, A. & FERGUSON, M. W. 1992. Changes in the extracellular matrix of the normal human breast during the menstrual cycle. *Cell Tissue Res,* 268**,** 167-77.
FERREIRA, M., MESQUITA, M., QUARESMA, M. & ANDRE, S. 2008. Prolactin receptor expression in gynaecomastia and male breast carcinoma. *Histopathology,* 53**,** 56-61.
FFRENCH-CONSTANT, C., VAN DE WATER, L., DVORAK, H. F. & HYNES, R. O. 1989. Reappearance of an embryonic pattern of fibronectin splicing during wound healing in the adult rat. *J Cell Biol,* 109**,** 903-14.
FINAK, G., BERTOS, N., PEPIN, F., SADEKOVA, S., SOULEIMANOVA, M., ZHAO, H., CHEN, H., OMEROGLU, G., METERISSIAN, S., OMEROGLU, A., HALLETT, M. & PARK, M. 2008. Stromal gene expression predicts clinical outcome in breast cancer. *Nat Med,* 14**,** 518-27.
FONSECA, R. R., TOMAS, A. R., ANDRE, S. & SOARES, J. 2006. Evaluation of ERBB2 gene status and chromosome 17 anomalies in male breast cancer. *Am J Surg Pathol,* 30**,** 1292-8.
FORD, D., EASTON, D. F., STRATTON, M., NAROD, S., GOLDGAR, D., DEVILEE, P., BISHOP, D. T., WEBER, B., LENOIR, G., CHANG-CLAUDE, J., SOBOL, H., TEARE, M. D., STRUEWING, J., ARASON, A., SCHERNECK, S., PETO, J., REBBECK, T. R., TONIN, P., NEUHAUSEN, S., BARKARDOTTIR, R., EYFJORD, J., LYNCH, H., PONDER, B. A., GAYTHER, S. A., ZELADA-HEDMAN, M. & ET AL. 1998. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet,* 62**,** 676-89.
FORONI, C., BROGGINI, M., GENERALI, D. & DAMIA, G. 2012. Epithelial-mesenchymal transition and breast cancer: role, molecular mechanisms and clinical impact. *Cancer Treat Rev,* 38**,** 689-97.
FRIDLYAND, J., SNIJDERS, A. M., YLSTRA, B., LI, H., OLSHEN, A., SEGRAVES, R., DAIRKEE, S., TOKUYASU, T., LJUNG, B. M., JAIN, A. N., MCLENNAN, J., ZIEGLER, J., CHIN, K., DEVRIES, S., FEILER, H., GRAY, J. W., WALDMAN, F., PINKEL, D. & ALBERTSON, D. G. 2006.

Breast tumor copy number aberration phenotypes and genomic instability. *BMC Cancer,* 6**,** 96.
FRIEDMAN, L. S., GAYTHER, S. A., KUROSAKI, T., GORDON, D., NOBLE, B., CASEY, G., PONDER, B. A. & ANTON-CULVER, H. 1997. Mutation analysis of BRCA1 and BRCA2 in a male breast cancer population. *Am J Hum Genet,* 60**,** 313-9.
FRIEDMAN, R., GELFAND, T., WEISS, D. W. & DOLJANSKI, F. 1984. Patterns of fibronectin deposition in normal and neoplastic fibroblasts and mammary tissue. *Int J Tissue React,* 6**,** 291-301.
FUJISUE, M., NISHIMURA, R., OKUMURA, Y., TASHIMA, R., NISHIYAMA, Y., OSAKO, T., TOYOZUMI, Y. & ARIMA, N. 2012. Clinical Significance of CK19 Negative Breast Cancer. *Cancers (Basel),* 5**,** 1-11.
GALLIMORE, P. H., MCDOUGALL, J. K. & CHEN, L. B. 1977. In vitro traits of adenovirus-transformed cell lines and their relevance to tumorigenicity in nude mice. *Cell,* 10**,** 669-78.
GENNARI, R., CURIGLIANO, G., JERECZEK-FOSSA, B. A., ZURRIDA, S., RENNE, G., INTRA, M., GALIMBERTI, V., LUINI, A., ORECCHIA, R., VIALE, G., GOLDHRISCH, A. & VERONESI, U. 2004. Male breast cancer: a special therapeutic problem. Anything new? (Review). *Int J Oncol,* 24**,** 663-70.
GENTILINI, O., CHAGAS, E., ZURRIDA, S., INTRA, M., DE CICCO, C., GATTI, G., SILVA, L., RENNE, G., CASSANO, E. & VERONESI, U. 2007. Sentinel lymph node biopsy in male patients with early breast cancer. *Oncologist,* 12**,** 512-5.
GIORDANO, S. H., BUZDAR, A. U. & HORTOBAGYI, G. N. 2002a. Breast cancer in men. *Ann Intern Med,* 137**,** 678-87.
GIORDANO, S. H., COHEN, D. S., BUZDAR, A. U., PERKINS, G. & HORTOBAGYI, G. N. 2004. Breast carcinoma in men: a population-based study. *Cancer,* 101**,** 51-7.
GIORDANO, S. H., PERKINS, G. H., BROGLIO, K., GARCIA, S. G., MIDDLETON, L. P., BUZDAR, A. U. & HORTOBAGYI, G. N. 2005. Adjuvant systemic therapy for male breast carcinoma. *Cancer,* 104**,** 2359-64.
GIORDANO, S. H., VALERO, V., BUZDAR, A. U. & HORTOBAGYI, G. N. 2002b. Efficacy of anastrozole in male breast cancer. *Am J Clin Oncol,* 25**,** 235-7.
GOSS, P. E., REID, C., PINTILIE, M., LIM, R. & MILLER, N. 1999. Male breast carcinoma: a review of 229 patients who presented to the Princess Margaret Hospital during 40 years: 1955-1996. *Cancer,* 85**,** 629-39.
GOULD, V. E., KOUKOULIS, G. K. & VIRTANEN, I. 1990. Extracellular matrix proteins and their receptors in the normal, hyperplastic and neoplastic breast. *Cell Differ Dev,* 32**,** 409-16.
GRAFF, J. R., KONICEK, B. W., CARTER, J. H. & MARCUSSON, E. G. 2008. Targeting the eukaryotic translation initiation factor 4E for cancer therapy. *Cancer Res,* 68**,** 631-4.
GRINNELL, F. 1978. Cellular adhesiveness and extracellular substrata. *Int Rev Cytol,* 53**,** 65-144.
GRINNELL, F., FELD, M. & MINTER, D. 1980. Fibroblast adhesion to fibrinogen and fibrin substrata: requirement for cold-insoluble globulin (plasma fibronectin). *Cell,* 19**,** 517-25.

GUDMUNDSDOTTIR, K., THORLACIUS, S., JONASSON, J. G., SIGFUSSON, B. F., TRYGGVADOTTIR, L. & EYFJORD, J. E. 2003. CYP17 promoter polymorphism and breast cancer risk in males and females in relation to BRCA2 status. *Br J Cancer,* 88**,** 933-6.

GUINEE, V. F., OLSSON, H., MOLLER, T., SHALLENBERGER, R. C., VAN DEN BLINK, J. W., PETER, Z., DURAND, M., DISCHE, S., CLETON, F. J., ZEWUSTER, R. & ET AL. 1993. The prognosis of breast cancer in males. A report of 335 cases. *Cancer,* 71**,** 154-61.

GURCAN, M. N., BOUCHERON, L. E., CAN, A., MADABHUSHI, A., RAJPOOT, N. M. & YENER, B. 2009. Histopathological image analysis: a review. *IEEE Rev Biomed Eng,* 2**,** 147-71.

GUTMAN, A. & KORNBLIHTT, A. R. 1987. Identification of a third region of cell-specific alternative splicing in human fibronectin mRNA. *Proc Natl Acad Sci U S A,* 84**,** 7179-82.

HAMILTON-BURKE, W., COLEMAN, L., CUMMINGS, M., GREEN, C. A., HOLLIDAY, D. L., HORGAN, K., MARAQA, L., PETER, M. B., POLLOCK, S., SHAABAN, A. M., SMITH, L. & SPEIRS, V. 2010. Phosphorylation of estrogen receptor beta at serine 105 is associated with good prognosis in breast cancer. *Am J Pathol,* 177**,** 1079-86.

HANAHAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: the next generation. *Cell,* 144**,** 646-74.

HAO, X., SUN, B., HU, L., LAHDESMAKI, H., DUNMIRE, V., FENG, Y., ZHANG, S. W., WANG, H., WU, C., WANG, H., FULLER, G. N., SYMMANS, W. F., SHMULEVICH, I. & ZHANG, W. 2004. Differential gene and protein expression in primary breast malignancies and their lymph node metastases as revealed by combined cDNA microarray and tissue microarray analysis. *Cancer,* 100**,** 1110-22.

HARALDSSON, K., LOMAN, N., ZHANG, Q. X., JOHANNSSON, O., OLSSON, H. & BORG, A. 1998. BRCA2 germ-line mutations are frequent in male breast cancer patients without a family history of the disease. *Cancer Res,* 58**,** 1367-71.

HARRIS, A. L., DOWSETT, M., STUART-HARRIS, R. & SMITH, I. E. 1986. Role of aminoglutethimide in male breast cancer. *Br J Cancer,* 54**,** 657-60.

HARVEY, J. M., CLARK, G. M., OSBORNE, C. K. & ALLRED, D. C. 1999. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. *J Clin Oncol,* 17**,** 1474-81.

HAYMAN, E. G., ENGVALL, E. & RUOSLAHTI, E. 1981. Concomitant loss of cell surface fibronectin and laminin from transformed rat kidney cells. *J Cell Biol,* 88**,** 352-57.

HELLEMAN, J., JANSEN, M. P., RUIGROK-RITSTIER, K., VAN STAVEREN, I. L., LOOK, M. P., MEIJER-VAN GELDER, M. E., SIEUWERTS, A. M., KLIJN, J. G., SLEIJFER, S., FOEKENS, J. A. & BERNS, E. M. 2008. Association of an extracellular matrix gene cluster with breast cancer prognosis and endocrine therapy response. *Clin Cancer Res,* 14**,** 5555-64.

HILL, T. D., KHAMIS, H. J., TYCZYNSKI, J. E. & BERKEL, H. J. 2005. Comparison of male and female breast cancer incidence trends, tumor characteristics, and survival. *Ann Epidemiol,* 15**,** 773-80.

HINNIS, A. R., LUCKETT, J. C. & WALKER, R. A. 2007. Survivin is an independent predictor of short-term survival in poor prognostic breast cancer patients. *Br J Cancer,* 96**,** 639-45.
HITTMAIR, A. P., LININGER, R. A. & TAVASSOLI, F. A. 1998. Ductal carcinoma in situ (DCIS) in the male breast: a morphologic study of 84 cases of pure DCIS and 30 cases of DCIS associated with invasive carcinoma--a preliminary report. *Cancer,* 83**,** 2139-49.
HOWAT, W. J., LEWIS, A., JONES, P., KAMPF, C., PONTEN, F., VAN DER LOOS, C. M., GRAY, N., WOMACK, C. & WARFORD, A. 2014. Antibody validation of immunohistochemistry for biomarker discovery: recommendations of a consortium of academic and pharmaceutical based histopathology researchers. *Methods,* 70**,** 34-8.
HOWELL, A., SIMS, A. H., ONG, K. R., HARVIE, M. N., EVANS, D. G. & CLARKE, R. B. 2005. Mechanisms of Disease: prediction and prevention of breast cancer--cellular and molecular interactions. *Nat Clin Pract Oncol,* 2**,** 635-46.
HSING, A. W., MCLAUGHLIN, J. K., COCCO, P., CO CHIEN, H. T. & FRAUMENI, J. F., JR. 1998. Risk factors for male breast cancer (United States). *Cancer Causes Control,* 9**,** 269-75.
HU, Z., FAN, C., OH, D. S., MARRON, J. S., HE, X., QAQISH, B. F., LIVASY, C., CAREY, L. A., REYNOLDS, E., DRESSLER, L., NOBEL, A., PARKER, J., EWEND, M. G., SAWYER, L. R., WU, J., LIU, Y., NANDA, R., TRETIAKOVA, M., RUIZ ORRICO, A., DREHER, D., PALAZZO, J. P., PERREARD, L., NELSON, E., MONE, M., HANSEN, H., MULLINS, M., QUACKENBUSH, J. F., ELLIS, M. J., OLOPADE, O. I., BERNARD, P. S. & PEROU, C. M. 2006. The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics,* 7**,** 96.
HULTBORN, R., HANSON, C., KOPF, I., VERBIENE, I., WARNHAMMAR, E. & WEIMARCK, A. 1997. Prevalence of Klinefelter's syndrome in male breast cancer patients. *Anticancer Res,* 17**,** 4293-7.
HUMPHRIES, M. P., JORDAN, V. C. & SPEIRS, V. 2015. Obesity and male breast cancer: provocative parallels? *BMC Med,* 13**,** 134.
HYNES, R. O. 1992. Integrins: versatility, modulation, and signaling in cell adhesion. *Cell,* 69**,** 11-25.
HYNES, R. O., ALI, I. U., DESTREE, A. T., MAUTNER, V., PERKINS, M. E., SENGER, D. R., WAGNER, D. D. & SMITH, K. K. 1978. A large glycoprotein lost from the surfaces of transformed cells. *Ann N Y Acad Sci,* 312**,** 317-42.
HYNES, R. O. & YAMADA, K. M. 1982. Fibronectins: multifunctional modular glycoproteins. *J Cell Biol,* 95**,** 369-77.
IDOWU, M. O., KMIECIAK, M., DUMUR, C., BURTON, R. S., GRIMES, M. M., POWERS, C. N. & MANJILI, M. H. 2012. CD44(+)/CD24(-/low) cancer stem/progenitor cells are more abundant in triple-negative invasive breast carcinoma phenotype and are associated with poor outcome. *Hum Pathol,* 43**,** 364-73.
ILYAS, M., GRABSCH, H., ELLIS, I. O., WOMACK, C., BROWN, R., BERNEY, D., FENNELL, D., SALTO-TELLEZ, M., JENKINS, M., LANDBERG, G., BYERS, R., TREANOR, D., HARRISON, D., GREEN, A. R., BALL, G., HAMILTON, P., NATIONAL CANCER RESEARCH INSTITUTE, B. & IMAGING CLINICAL STUDIES, G. 2013. Guidelines and considerations

for conducting experiments using tissue microarrays. *Histopathology,* 62**,** 827-39.
IOACHIM, E., CHARCHANTI, A., BRIASOULIS, E., KARAVASILIS, V., TSANOU, H., ARVANITIS, D. L., AGNANTIS, N. J. & PAVLIDIS, N. 2002. Immunohistochemical expression of extracellular matrix components tenascin, fibronectin, collagen type IV and laminin in breast cancer: their prognostic value and role in tumour invasion and progression. *Eur J Cancer,* 38**,** 2362-70.
ITALIANO, A., LARGILLIER, R., MARCY, P. Y., FOA, C., FERRERO, J. M., HARTMANN, M. T. & NAMER, M. 2004. [Complete remission obtained with letrozole in a man with metastatic breast cancer]. *Rev Med Interne,* 25**,** 323-4.
JEMAL, A., SIEGEL, R., WARD, E., HAO, Y., XU, J., MURRAY, T. & THUN, M. J. 2008. Cancer statistics, 2008. *CA Cancer J Clin,* 58**,** 71-96.
JOHANSSON, I., NILSSON, C., BERGLUND, P., LAUSS, M., RINGNER, M., OLSSON, H., LUTS, L., SIM, E., THORSTENSSON, S., FJALLSKOG, M. L. & HEDENFALK, I. 2012. Gene expression profiling of primary male breast cancers reveals two unique subgroups and identifies N-acetyltransferase-1 (NAT1) as a novel prognostic biomarker. *Breast Cancer Res,* 14**,** R31.
JOHANSSON, I., NILSSON, C., BERGLUND, P., STRAND, C., JONSSON, G., STAAF, J., RINGNER, M., NEVANLINNA, H., BARKARDOTTIR, R. B., BORG, A., OLSSON, H., LUTS, L., FJALLSKOG, M. L. & HEDENFALK, I. 2011. High-resolution genomic profiling of male breast cancer reveals differences hidden behind the similarities with female breast cancer. *Breast Cancer Res Treat,* 129**,** 747-60.
JOHANSSON, I., RINGNER, M. & HEDENFALK, I. 2013. The landscape of candidate driver genes differs between male and female breast cancer. *PLoS One,* 8**,** e78299.
JONSSON, G., STAAF, J., VALLON-CHRISTERSSON, J., RINGNER, M., HOLM, K., HEGARDT, C., GUNNARSSON, H., FAGERHOLM, R., STRAND, C., AGNARSSON, B. A., KILPIVAARA, O., LUTS, L., HEIKKILA, P., AITTOMAKI, K., BLOMQVIST, C., LOMAN, N., MALMSTROM, P., OLSSON, H., JOHANNSSON, O. T., ARASON, A., NEVANLINNA, H., BARKARDOTTIR, R. B. & BORG, A. 2010. Genomic subtypes of breast cancer identified by array-comparative genomic hybridization display distinct molecular and clinical characteristics. *Breast Cancer Res,* 12**,** R42.
JOSHI, M. G., LEE, A. K., LODA, M., CAMUS, M. G., PEDERSEN, C., HEATLEY, G. J. & HUGHES, K. S. 1996. Male breast carcinoma: an evaluation of prognostic factors contributing to a poorer outcome. *Cancer,* 77**,** 490-8.
KACZMAREK, J., CASTELLANI, P., NICOLO, G., SPINA, B., ALLEMANNI, G. & ZARDI, L. 1994. Distribution of oncofetal fibronectin isoforms in normal, hyperplastic and neoplastic human breast tissues. *Int J Cancer,* 59**,** 11-6.
KADAR, A., TOKES, A. M., KULKA, J. & ROBERT, L. 2002. Extracellular matrix components in breast carcinomas. *Semin Cancer Biol,* 12**,** 243-57.

KANTARJIAN, H., YAP, H. Y., HORTOBAGYI, G., BUZDAR, A. & BLUMENSCHEIN, G. 1983. Hormonal therapy for metastatic male breast cancer. *Arch Intern Med,* 143**,** 237-40.

KARNOUB, A. E., DASH, A. B., VO, A. P., SULLIVAN, A., BROOKS, M. W., BELL, G. W., RICHARDSON, A. L., POLYAK, K., TUBO, R. & WEINBERG, R. A. 2007. Mesenchymal stem cells within tumour stroma promote breast cancer metastasis. *Nature,* 449**,** 557-63.

KENNEDY, S. M., O'DRISCOLL, L., PURCELL, R., FITZ-SIMONS, N., MCDERMOTT, E. W., HILL, A. D., O'HIGGINS, N. J., PARKINSON, M., LINEHAN, R. & CLYNES, M. 2003. Prognostic importance of survivin in breast cancer. *Br J Cancer,* 88**,** 1077-83.

KIM, S., BELL, K., MOUSA, S. A. & VARNER, J. A. 2000. Regulation of angiogenesis in vivo by ligation of integrin alpha5beta1 with the central cell-binding domain of fibronectin. *Am J Pathol,* 156**,** 1345-62.

KIRKEGAARD, T., EDWARDS, J., TOVEY, S., MCGLYNN, L. M., KRISHNA, S. N., MUKHERJEE, R., TAM, L., MUNRO, A. F., DUNNE, B. & BARTLETT, J. M. 2006. Observer variation in immunohistochemical analysis of protein expression, time for a change? *Histopathology,* 48**,** 787-94.

KNUDSEN, E. S., ERTEL, A., DAVICIONI, E., KLINE, J., SCHWARTZ, G. F. & WITKIEWICZ, A. K. 2012. Progression of ductal carcinoma in situ to invasive breast cancer is associated with gene expression programs of EMT and myoepithelia. *Breast Cancer Res Treat,* 133**,** 1009-24.

KORDE, L. A., ZUJEWSKI, J. A., KAMIN, L., GIORDANO, S., DOMCHEK, S., ANDERSON, W. F., BARTLETT, J. M., GELMON, K., NAHLEH, Z., BERGH, J., CUTULI, B., PRUNERI, G., MCCASKILL-STEVENS, W., GRALOW, J., HORTOBAGYI, G. & CARDOSO, F. 2010. Multidisciplinary meeting on male breast cancer: summary and research recommendations. *J Clin Oncol,* 28**,** 2114-22.

KORNBLIHTT, A. R., UMEZAWA, K., VIBE-PEDERSEN, K. & BARALLE, F. E. 1985. Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene. *EMBO J,* 4**,** 1755-9.

KORNBLIHTT, A. R., VIBE-PEDERSEN, K. & BARALLE, F. E. 1984. Human fibronectin: molecular cloning evidence for two mRNA species differing by an internal segment coding for a structural domain. *EMBO J,* 3**,** 221-6.

KORNEGOOR, R., VERSCHUUR-MAES, A. H., BUERGER, H., HOGENES, M. C., DE BRUIN, P. C., OUDEJANS, J. J., VAN DER GROEP, P., HINRICHS, B. & VAN DIEST, P. J. 2012. Molecular subtyping of male breast cancer by immunohistochemistry. *Mod Pathol,* 25**,** 398-404.

KORSCHING, E., PACKEISEN, J., AGELOPOULOS, K., EISENACHER, M., VOSS, R., ISOLA, J., VAN DIEST, P. J., BRANDT, B., BOECKER, W. & BUERGER, H. 2002. Cytogenetic alterations and cytokeratin expression patterns in breast cancer: integrating a new model of breast differentiation into cytogenetic pathways of breast carcinogenesis. *Lab Invest,* 82**,** 1525-33.

KOUKOULIS, G. K., HOWEEDY, A. A., KORHONEN, M., VIRTANEN, I. & GOULD, V. E. 1993. Distribution of tenascin, cellular fibronectins and integrins in the normal, hyperplastic and neoplastic breast. *J Submicrosc Cytol Pathol,* 25**,** 285-95.

KWIATKOWSKA, E., TERESIAK, M., FILAS, V., KARCZEWSKA, A., BREBOROWICZ, D. & MACKIEWICZ, A. 2003. BRCA2 mutations and androgen receptor expression as independent predictors of outcome of male breast cancer patients. *Clin Cancer Res,* 9**,** 4452-9.

LA VECCHIA, C., LEVI, F. & LUCCHINI, F. 1992. Descriptive epidemiology of male breast cancer in Europe. *Int J Cancer,* 51**,** 62-6.

LANDIS, J. R. & KOCH, G. G. 1977. The measurement of observer agreement for categorical data. *Biometrics,* 33**,** 159-74.

LEEK, R. D., KAKLAMANIS, L., PEZZELLA, F., GATTER, K. C. & HARRIS, A. L. 1994. bcl-2 in normal human breast and carcinoma, association with oestrogen receptor-positive, epidermal growth factor receptor-negative tumours and in situ cancer. *Br J Cancer,* 69**,** 135-9.

LEFOR, A. T. & NUMANN, P. J. 1988. Carcinoma of the breast in men. *N Y State J Med,* 88**,** 293-6.

LEHMANN, U., STREICHERT, T., OTTO, B., ALBAT, C., HASEMEIER, B., CHRISTGEN, H., SCHIPPER, E., HILLE, U., KREIPE, H. H. & LANGER, F. 2010. Identification of differentially expressed microRNAs in human male breast cancer. *BMC Cancer,* 10**,** 109.

LI, F., YANG, J., RAMNATH, N., JAVLE, M. M. & TAN, D. 2005. Nuclear or cytoplasmic expression of survivin: what is the significance? *Int J Cancer,* 114**,** 509-12.

LOPEZ, M., DI LAURO, L., LAZZARO, B. & PAPALDO, P. 1985. Hormonal treatment of disseminated male breast cancer. *Oncology,* 42**,** 345-9.

LYNCH, H. T., WATSON, P. & NAROD, S. A. 1999. The genetic epidemiology of male breast carcinoma. *Cancer,* 86**,** 744-6.

MA, X. J., DAHIYA, S., RICHARDSON, E., ERLANDER, M. & SGROI, D. C. 2009. Gene expression profiling of the tumor microenvironment during breast cancer progression. *Breast Cancer Res,* 11**,** R7.

MABUCHI, K., BROSS, D. S. & KESSLER, II 1985. Risk factors for male breast cancer. *J Natl Cancer Inst,* 74**,** 371-5.

MALZAHN, K., MITZE, M., THOENES, M. & MOLL, R. 1998. Biological and prognostic significance of stratified epithelial cytokeratins in infiltrating ductal breast carcinomas. *Virchows Arch,* 433**,** 119-29.

MARGARIA, E., CHIUSA, L., FERRARI, L., DAL CANTON, O. & PICH, A. 2000. Therapy and survival in male breast carcinoma: A retrospective analysis of 50 cases. *Oncol Rep,* 7**,** 1035-9.

MAROTTA, L. L., ALMENDRO, V., MARUSYK, A., SHIPITSIN, M., SCHEMME, J., WALKER, S. R., BLOUSHTAIN-QIMRON, N., KIM, J. J., CHOUDHURY, S. A., MARUYAMA, R., WU, Z., GONEN, M., MULVEY, L. A., BESSARABOVA, M. O., HUH, S. J., SILVER, S. J., KIM, S. Y., PARK, S. Y., LEE, H. E., ANDERSON, K. S., RICHARDSON, A. L., NIKOLSKAYA, T., NIKOLSKY, Y., LIU, X. S., ROOT, D. E., HAHN, W. C., FRANK, D. A. & POLYAK, K. 2011. The JAK2/STAT3 signaling pathway is required for growth of CD44(+)CD24(-) stem cell-like breast cancer cells in human tumors. *J Clin Invest,* 121**,** 2723-35.

MEIJERS-HEIJBOER, H., VAN DEN OUWELAND, A., KLIJN, J., WASIELEWSKI, M., DE SNOO, A., OLDENBURG, R., HOLLESTELLE, A., HOUBEN, M., CREPIN, E., VAN VEGHEL-PLANDSOEN, M., ELSTRODT, F., VAN DUIJN, C., BARTELS, C., MEIJERS, C., SCHUTTE, M., MCGUFFOG, L., THOMPSON, D., EASTON, D.,

SODHA, N., SEAL, S., BARFOOT, R., MANGION, J., CHANG-CLAUDE, J., ECCLES, D., EELES, R., EVANS, D. G., HOULSTON, R., MURDAY, V., NAROD, S., PERETZ, T., PETO, J., PHELAN, C., ZHANG, H. X., SZABO, C., DEVILEE, P., GOLDGAR, D., FUTREAL, P. A., NATHANSON, K. L., WEBER, B., RAHMAN, N. & STRATTON, M. R. 2002. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet,* 31**,** 55-9.

MOHAMMED, Z. M., MCMILLAN, D. C., ELSBERGER, B., GOING, J. J., ORANGE, C., MALLON, E., DOUGHTY, J. C. & EDWARDS, J. 2012. Comparison of visual and automated assessment of Ki-67 proliferative activity and their impact on outcome in primary operable invasive ductal breast cancer. *Br J Cancer,* 106**,** 383-8.

MUIR, D., KANTHAN, R. & KANTHAN, S. C. 2003. Male versus female breast cancers. A population-based comparative immunohistochemical analysis. *Arch Pathol Lab Med,* 127**,** 36-41.

MUNOZ DE TORO, M. M., MAFFINI, M. V., KASS, L. & LUQUE, E. H. 1998. Proliferative activity and steroid hormone receptor status in male breast carcinoma. *The Journal of steroid biochemistry and molecular biology,* 67**,** 333-9.

MURPHY, L. C. & WATSON, P. 2002. Steroid receptors in human breast tumorigenesis and breast cancer progression. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie,* 56**,** 65-77.

NAHLEH, Z. & GIRNIUS, S. 2006. Male breast cancer: a gender issue. *Nature clinical practice. Oncology,* 3**,** 428-37.

NAHLEH, Z. A., SRIKANTIAH, R., SAFA, M., JAZIEH, A. R., MUHLEMAN, A. & KOMROKJI, R. 2007. Male breast cancer in the veterans affairs population: a comparative analysis. *Cancer,* 109**,** 1471-7.

NEUHAUSEN, S., DUNNING, A., STEELE, L., YAKUMO, K., HOFFMAN, M., SZABO, C., TEE, L., BAINES, C., PHAROAH, P., GOLDGAR, D. & EASTON, D. 2004. Role of CHEK2*1100delC in unselected series of non-BRCA1/2 male breast cancers. *Int J Cancer,* 108**,** 477-8.

NEWMAN, J. 1997. Breast cancer in men and mammography of the male breast. *Radiol Technol,* 69**,** 17-28; quiz 29-36.

NIELSEN, T. O., HSU, F. D., JENSEN, K., CHEANG, M., KARACA, G., HU, Z., HERNANDEZ-BOUSSARD, T., LIVASY, C., COWAN, D., DRESSLER, L., AKSLEN, L. A., RAGAZ, J., GOWN, A. M., GILKS, C. B., VAN DE RIJN, M. & PEROU, C. M. 2004. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res,* 10**,** 5367-74.

OHAYON, T., GAL, I., BARUCH, R. G., SZABO, C. & FRIEDMAN, E. 2004. CHEK2*1100delC and male breast cancer risk in Israel. *Int J Cancer,* 108**,** 479-80.

OJOPI, E. P., CAVALLI, L. R., CAVALIERI, L. M., SQUIRE, J. A. & ROGATTO, S. R. 2002. Comparative genomic hybridization analysis of benign and invasive male breast neoplasms. *Cancer Genet Cytogenet,* 134**,** 123-6.

OLDEN, K. & YAMADA, K. M. 1977. Mechanism of the decrease in the major cell surface protein of chick embryo fibroblasts after transformation. *Cell,* 11**,** 957-69.

OSBORNE, C. K. 1998. Steroid hormone receptors in breast cancer management. *Breast Cancer Res Treat,* 51**,** 227-38.
OTTINI, L., RIZZOLO, P., ZANNA, I., FALCHETTI, M., MASALA, G., CECCARELLI, K., VEZZOSI, V., GULINO, A., GIANNINI, G., BIANCHI, S., SERA, F. & PALLI, D. 2009. BRCA1/BRCA2 mutation status and clinical-pathologic features of 108 male breast cancer cases from Tuscany: a population-based study in central Italy. *Breast Cancer Res Treat,* 116**,** 577-86.
PANNUTI, F., MARTONI, A., BUSUTTI, L., GRAMELLINI, M. & PIANA, E. 1982. High-dose medroxyprogesterone acetate in advanced male breast cancer. *Cancer Treat Rep,* 66**,** 1763-5.
PARIKH, R. R., YANG, Q., HIGGINS, S. A. & HAFFTY, B. G. 2008. Outcomes in young women with breast cancer of triple-negative phenotype: the prognostic significance of CK19 expression. *Int J Radiat Oncol Biol Phys,* 70**,** 35-42.
PARK, J. & SCHWARZBAUER, J. E. 2014. Mammary epithelial cell interactions with fibronectin stimulate epithelial-mesenchymal transition. *Oncogene,* 33**,** 1649-57.
PATEL, H. Z., 2ND, BUZDAR, A. U. & HORTOBAGYI, G. N. 1989. Role of adjuvant chemotherapy in male breast cancer. *Cancer,* 64**,** 1583-5.
PATTERSON, J. S., BATTERSBY, L. A. & BACH, B. K. 1980. Use of tamoxifen in advanced male breast cancer. *Cancer Treat Rep,* 64**,** 801-4.
PEARLSTEIN, E. 1976. Plasma membrane glycoprotein which mediates adhesion of fibroblasts to collagen. *Nature,* 262**,** 497-500.
PELEKANOU, V. & LECLERCQ, G. 2011. Recent insights into the effect of natural and environmental estrogens on mammary development and carcinogenesis. *Int J Dev Biol,* 55**,** 869-78.
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., FLUGE, O., PERGAMENSCHIKOV, A., WILLIAMS, C., ZHU, S. X., LONNING, P. E., BORRESEN-DALE, A. L., BROWN, P. O. & BOTSTEIN, D. 2000. Molecular portraits of human breast tumours. *Nature,* 406**,** 747-52.
PERTSCHUK, L. P., FELDMAN, J. G., KIM, Y. D., BRAITHWAITE, L., SCHNEIDER, F., BRAVERMAN, A. S. & AXIOTIS, C. 1996. Estrogen receptor immunocytochemistry in paraffin embedded tissues with ER1D5 predicts breast cancer endocrine response more accurately than H222Sp gamma in frozen sections or cytosol-based ligand-binding assays. *Cancer,* 77**,** 2514-9.
PICH, A., MARGARIA, E., CHIUSA, L., CANDELARESI, G. & DAL CANTON, O. 1999. Androgen receptor expression in male breast carcinoma: lack of clinicopathological association. *Br J Cancer,* 79**,** 959-64.
PICH, A., MARGARIA, E., CHIUSA, L., PONTI, R. & GEUNA, M. 1996. DNA ploidy and p53 expression correlate with survival and cell proliferative activity in male breast carcinoma. *Hum Pathol,* 27**,** 676-82.
PIKE, M. C., KRAILO, M. D., HENDERSON, B. E., CASAGRANDE, J. T. & HOEL, D. G. 1983. 'Hormonal' risk factors, 'breast tissue age' and the age-incidence of breast cancer. *Nature,* 303**,** 767-70.
PINDER, S. E., WENCYK, P., SIBBERING, D. M., BELL, J. A., ELSTON, C. W., NICHOLSON, R., ROBERTSON, J. F., BLAMEY, R. W. & ELLIS, I. O.

1995. Assessment of the new proliferation marker MIB1 in breast carcinoma using image analysis: associations with other prognostic factors and survival. *Br J Cancer,* 71**,** 146-9.

PISCUOGLIO, S., NG, C. K., MURRAY, M. P., GUERINI-ROCCO, E., MARTELOTTO, L. G., GEYER, F. C., BIDARD, F. C., BERMAN, S., FUSCO, N., SAKR, R. A., EBERLE, C., DE MATTOS-ARRUDA, L., MACEDO, G. S., AKRAM, M., BASLAN, T., HICKS, J., KING, T. A., BROGI, E., NORTON, L., WEIGELT, B., HUDIS, C. A. & REIS-FILHO, J. S. 2016. The Genomic Landscape of Male Breast Cancers. *Clin Cancer Res*.

PLESAN, D. M., GEORGESCU, C. V., PATRANA, N., PLESAN, C. & STOICA, D. 2010. Immunohistochemical study of p53 and Ki67 in a group of patients with mammary carcinoma. *Rom J Morphol Embryol,* 51**,** 459-65.

POLYAK, K. & KALLURI, R. 2010. The role of the microenvironment in mammary gland development and cancer. *Cold Spring Harb Perspect Biol,* 2**,** a003244.

RAVANDI-KASHANI, F. & HAYES, T. G. 1998. Male breast cancer: a review of the literature. *Eur J Cancer,* 34**,** 1341-7.

RAYSON, D., ERLICHMAN, C., SUMAN, V. J., ROCHE, P. C., WOLD, L. E., INGLE, J. N. & DONOHUE, J. H. 1998. Molecular markers in male breast carcinoma. *Cancer,* 83**,** 1947-55.

REED, W., HANNISDAL, E., BOEHLER, P. J., GUNDERSEN, S., HOST, H. & MARTHIN, J. 2000. The prognostic value of p53 and c-erb B-2 immunostaining is overrated for patients with lymph node negative breast carcinoma: a multivariate analysis of prognostic factors in 613 patients with a follow-up of 14-30 years. *Cancer,* 88**,** 804-13.

REXHEPAJ, E., BRENNAN, D. J., HOLLOWAY, P., KAY, E. W., MCCANN, A. H., LANDBERG, G., DUFFY, M. J., JIRSTROM, K. & GALLAGHER, W. M. 2008. Novel image analysis approach for quantifying expression of nuclear proteins assessed by immunohistochemistry: application to measurement of oestrogen and progesterone receptor levels in breast cancer. *Breast Cancer Res,* 10**,** R89.

REXHEPAJ, E., JIRSTROM, K., O'CONNOR, D. P., O'BRIEN, S. L., LANDBERG, G., DUFFY, M. J., BRENNAN, D. J. & GALLAGHER, W. M. 2010. Validation of cytoplasmic-to-nuclear ratio of survivin as an indicator of improved prognosis in breast cancer. *BMC Cancer,* 10**,** 639.

RIBEIRO, G. G. 1976. The results of diethylstilboestrol therapy for recurrent and metastatic carcinoma of the male breast. *Br J Cancer,* 33**,** 465-7.

ROSENBAUM, P. F., VENA, J. E., ZIELEZNY, M. A. & MICHALEK, A. M. 1994. Occupational exposures associated with male breast cancer. *Am J Epidemiol,* 139**,** 30-6.

RUDLOWSKI, C., SCHULTEN, H. J., GOLAS, M. M., SANDER, B., BARWING, R., PALANDT, J. E., SCHLEHE, B., LINDENFELSER, R., MOLL, R., LIERSCH, T., SCHUMPELICK, V., GUNAWAN, B. & FUZESI, L. 2006. Comparative genomic hybridization analysis on male breast cancer. *Int J Cancer,* 118**,** 2455-60.

RYAN, B. M., KONECNY, G. E., KAHLERT, S., WANG, H. J., UNTCH, M., MENG, G., PEGRAM, M. D., PODRATZ, K. C., CROWN, J., SLAMON, D. J. & DUFFY, M. J. 2006. Survivin expression in breast cancer predicts

clinical outcome and is associated with HER2, VEGF, urokinase plasminogen activator and PAI-1. *Ann Oncol,* 17**,** 597-604.
SAKAI, T., LARSEN, M. & YAMADA, K. M. 2003. Fibronectin requirement in branching morphogenesis. *Nature,* 423**,** 876-81.
SALVADORI, B., SACCOZZI, R., MANZARI, A., ANDREOLA, S., CONTI, R. A., CUSUMANO, F. & GRASSI, M. 1994. Prognosis of breast cancer in males: an analysis of 170 cases. *Eur J Cancer,* 30A**,** 930-5.
SASCO, A. J., LOWENFELS, A. B. & PASKER-DE JONG, P. 1993. Review article: epidemiology of male breast cancer. A meta-analysis of published case-control studies and discussion of selected aetiological factors. *Int J Cancer,* 53**,** 538-49.
SCHEDIN, P. & HOVEY, R. C. 2010. Editorial: The mammary stroma in normal development and function. *J Mammary Gland Biol Neoplasia,* 15**,** 275-7.
SCHUCHARDT, U., SEEGENSCHMIEDT, M. H., KIRSCHNER, M. J., RENNER, H. & SAUER, R. 1996. Adjuvant radiotherapy for breast carcinoma in men: a 20-year clinical experience. *Am J Clin Oncol,* 19**,** 330-6.
SCHWARZBAUER, J. E., PATEL, R. S., FONDA, D. & HYNES, R. O. 1987. Multiple sites of alternative splicing of the rat fibronectin gene transcript. *EMBO J,* 6**,** 2573-80.
SCHWARZBAUER, J. E., TAMKUN, J. W., LEMISCHKA, I. R. & HYNES, R. O. 1983. Three different fibronectin mRNAs arise by alternative splicing within the coding region. *Cell,* 35**,** 421-31.
SCOTT-CONNER, C. E., JOCHIMSEN, P. R., MENCK, H. R. & WINCHESTER, D. J. 1999. An analysis of male and female breast cancer treatment and survival among demographically identical pairs of patients. *Surgery,* 126**,** 775-80; discussion 780-1.
SHAABAN, A. M., BALL, G. R., BRANNAN, R. A., CSERNI, G., DI BENEDETTO, A., DENT, J., FULFORD, L., HONARPISHEH, H., JORDAN, L., JONES, J. L., KANTHAN, R., MARAQA, L., LITWINIUK, M., MOTTOLESE, M., POLLOCK, S., PROVENZANO, E., QUINLAN, P. R., REALL, G., SHOUSHA, S., STEPHENS, M., VERGHESE, E. T., WALKER, R. A., HANBY, A. M. & SPEIRS, V. 2012. A comparative biomarker study of 514 matched cases of male and female breast cancer reveals gender-specific biological differences. *Breast Cancer Res Treat,* 133**,** 949-58.
SHAO, M. M., CHAN, S. K., YU, A. M., LAM, C. C., TSANG, J. Y., LUI, P. C., LAW, B. K., TAN, P. H. & TSE, G. M. 2012. Keratin expression in breast cancers. *Virchows Arch,* 461**,** 313-22.
SHAVERS, V. L., HARLAN, L. C. & STEVENS, J. L. 2003. Racial/ethnic variation in clinical presentation, treatment, and survival among breast cancer patients under age 35. *Cancer,* 97**,** 134-47.
SHPITZ, B., BOMSTEIN, Y., STERNBERG, A., KLEIN, E., LIVERANT, S., GROISMAN, G. & BERNHEIM, J. 2000. Angiogenesis, p53, and c-erbB-2 immunoreactivity and clinicopathological features in male breast cancer. *J Surg Oncol,* 75**,** 252-7.
SIMPSON, P. T., GALE, T., REIS-FILHO, J. S., JONES, C., PARRY, S., SLOANE, J. P., HANBY, A., PINDER, S. E., LEE, A. H., HUMPHREYS, S., ELLIS, I. O. & LAKHANI, S. R. 2005. Columnar cell lesions of the

breast: the missing link in breast cancer progression? A morphological and molecular analysis. *Am J Surg Pathol,* 29**,** 734-46.

SORLIE, T., TIBSHIRANI, R., PARKER, J., HASTIE, T., MARRON, J. S., NOBEL, A., DENG, S., JOHNSEN, H., PESICH, R., GEISLER, S., DEMETER, J., PEROU, C. M., LONNING, P. E., BROWN, P. O., BORRESEN-DALE, A. L. & BOTSTEIN, D. 2003. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A,* 100**,** 8418-23.

SPEIRS, V. & WALKER, R. A. 2007. New perspectives into the biological and clinical relevance of oestrogen receptors in the human breast. *The Journal of pathology,* 211**,** 499-506.

STAMPFER, M. R., VLODAVSKY, I., SMITH, H. S., FORD, R., BECKER, F. F. & RIGGS, J. 1981. Fibronectin production by human mammary cells. *J Natl Cancer Inst,* 67**,** 253-61.

STEINITZ, R., KATZ, L. & BEN-HUR, M. 1981. Male breast cancer in Israel: selected epidemiological aspects. *Isr J Med Sci,* 17**,** 816-21.

STENLUND, C. & FLODERUS, B. 1997. Occupational exposure to magnetic fields in relation to male breast cancer and testicular cancer: a Swedish case-control study. *Cancer Causes Control,* 8**,** 184-91.

STIERER, M., ROSEN, H., WEITENSFELDER, W., HAUSMANINGER, H., TELEKY, B., JAKESZ, R., FRUHWIRTH, H., DUNSER, M., BELLER, S., HAID, A. & ET AL. 1995. Male breast cancer: Austrian experience. *World J Surg,* 19**,** 687-92; discussion 692-3.

STRUTZ, F., ZEISBERG, M., ZIYADEH, F. N., YANG, C. Q., KALLURI, R., MULLER, G. A. & NEILSON, E. G. 2002. Role of basic fibroblast growth factor-2 in epithelial-mesenchymal transformation. *Kidney Int,* 61**,** 1714-28.

SUDO, T., IWAYA, T., NISHIDA, N., SAWADA, G., TAKAHASHI, Y., ISHIBASHI, M., SHIBATA, K., FUJITA, H., SHIROUZU, K., MORI, M. & MIMORI, K. 2013. Expression of mesenchymal markers vimentin and fibronectin: the clinical significance in esophageal squamous cell carcinoma. *Ann Surg Oncol,* 20 Suppl 3**,** S324-35.

SUNG, C. O., PARK, C. K. & KIM, S. H. 2011. Classification of epithelial-mesenchymal transition phenotypes in esophageal squamous cell carcinoma is strongly associated with patient prognosis. *Mod Pathol,* 24**,** 1060-8.

SWIATONIOWSKI, G., MATKOWSKI, R., SUDER, E., BRUZEWICZ, S., SETTA, M., KORNAFEL, J., POLOZOWSKI, A. & SUROWIAK, P. 2005. E-cadherin and fibronectin expressions have no prognostic role in stage II ductal breast cancer. *Anticancer Res,* 25**,** 2879-83.

SYRJAKOSKI, K., HYYTINEN, E. R., KUUKASJARVI, T., AUVINEN, A., KALLIONIEMI, O. P., KAINU, T. & KOIVISTO, P. A. 2003. Androgen receptor gene alterations in Finnish male breast cancer. *Breast Cancer Res Treat,* 77**,** 167-70.

SYRJAKOSKI, K., KUUKASJARVI, T., AUVINEN, A. & KALLIONIEMI, O. P. 2004. CHEK2 1100delC is not a risk factor for male breast cancer population. *Int J Cancer,* 108**,** 475-6.

TAKEI, H., IINO, Y., HORIGUCHI, J. & YOKOE, T. 1995. Immunohistochemical fibronectin staining pattern and prognosis in invasive breast carcinoma. *Oncology,* 52**,** 106-11.

TANAKA, K., IWAMOTO, S., GON, G., NOHARA, T., IWAMOTO, M. & TANIGAWA, N. 2000. Expression of survivin and its relationship to loss of apoptosis in breast carcinomas. *Clin Cancer Res,* 6**,** 127-34.

TAVAKOLI, M., BATENI, E., ATTARBASHI-MOGHADAM, F., TALEBI, A., YAGHINI, J. & MOGHAREHABED, A. 2011. Comparison of fibronectin in human marginal gingiva and interdental papilla using immunohistochemistry. *Dent Res J (Isfahan),* 8**,** S109-13.

TAYLOR-PAPADIMITRIOU, J., BURCHELL, J. & HURST, J. 1981. Production of fibronectin by normal and malignant human mammary epithelial cells. *Cancer Res,* 41**,** 2491-500.

TERRANOVA, V. P., AUMAILLEY, M., SULTAN, L. H., MARTIN, G. R. & KLEINMAN, H. K. 1986. Regulation of cell attachment and cell number by fibronectin and laminin. *J Cell Physiol,* 127**,** 473-9.

THOMAS, D. B. 1993. Breast cancer in men. *Epidemiol Rev,* 15**,** 220-31.

THOMAS, D. B., JIMENEZ, L. M., MCTIERNAN, A., ROSENBLATT, K., STALSBERG, H., STEMHAGEN, A., THOMPSON, W. D., CURNEN, M. G., SATARIANO, W., AUSTIN, D. F. & ET AL. 1992. Breast cancer in men: risk factors with hormonal implications. *Am J Epidemiol,* 135**,** 734-48.

THORLACIUS, S., OLAFSDOTTIR, G., TRYGGVADOTTIR, L., NEUHAUSEN, S., JONASSON, J. G., TAVTIGIAN, S. V., TULINIUS, H., OGMUNDSDOTTIR, H. M. & EYFJORD, J. E. 1996. A single BRCA2 mutation in male and female breast cancer families from Iceland with varied cancer phenotypes. *Nat Genet,* 13**,** 117-9.

THORLACIUS, S., SIGURDSSON, S., BJARNADOTTIR, H., OLAFSDOTTIR, G., JONASSON, J. G., TRYGGVADOTTIR, L., TULINIUS, H. & EYFJORD, J. E. 1997. Study of a single BRCA2 mutation with high carrier frequency in a small population. *Am J Hum Genet,* 60**,** 1079-84.

TIRKKONEN, M., KAINU, T., LOMAN, N., JOHANNSSON, O. T., OLSSON, H., BARKARDOTTIR, R. B., KALLIONIEMI, O. P. & BORG, A. 1999. Somatic genetic alterations in BRCA2-associated and sporadic male breast cancer. *Genes, chromosomes & cancer,* 24**,** 56-61.

TISCHKOWITZ, M. D., HODGSON, S. V. & FENTIMAN, I. S. 2002. 19. Male breast cancer: aetiology, genetics and clinical management. *Int J Clin Pract,* 56**,** 750-4.

TOMMASI, S., MANGIA, A., IANNELLI, G., CHIARAPPA, P., ROSSI, E., OTTINI, L., MOTTOLESE, M., ZOLI, W., ZUFFARDI, O. & PARADISO, A. 2010. Gene copy number variation in male breast cancer by aCGH. *Anal Cell Pathol (Amst),* 33**,** 113-9.

TRAN, J., RAK, J., SHEEHAN, C., SAIBIL, S. D., LACASSE, E., KORNELUK, R. G. & KERBEL, R. S. 1999. Marked induction of the IAP family antiapoptotic proteins survivin and XIAP by VEGF in vascular endothelial cells. *Biochem Biophys Res Commun,* 264**,** 781-8.

TRUONG, P. T., BERTHELET, E., LEE, J., KADER, H. A. & 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**,** 2006-14.

TUOMINEN, V. J., RUOTOISTENMAKI, S., VIITANEN, A., JUMPPANEN, M. & ISOLA, J. 2010. ImmunoRatio: a publicly available web application for

quantitative image analysis of estrogen receptor (ER), progesterone receptor (PR), and Ki-67. *Breast Cancer Res,* 12**,** R56.
TWOROGER, S. S. & HANKINSON, S. E. 2008. Prolactin and breast cancer etiology: an epidemiologic perspective. *J Mammary Gland Biol Neoplasia,* 13**,** 41-53.
VAHERI, A. & RUOSLAHTI, E. 1975. Fibroblast surface antigen produced but not retained by virus-transformed human cells. *J Exp Med,* 142**,** 530-5.
VAN DE RIJN, M., PEROU, C. M., TIBSHIRANI, R., HAAS, P., KALLIONIEMI, O., KONONEN, J., TORHORST, J., SAUTER, G., ZUBER, M., KOCHLI, O. R., MROSS, F., DIETERICH, H., SEITZ, R., ROSS, D., BOTSTEIN, D. & BROWN, P. 2002. Expression of cytokeratins 17 and 5 identifies a group of breast carcinomas with poor clinical outcome. *Am J Pathol,* 161**,** 1991-6.
VARGAS, A. C., MCCART REED, A. E., WADDELL, N., LANE, A., REID, L. E., SMART, C. E., COCCIARDI, S., DA SILVA, L., SONG, S., CHENEVIX-TRENCH, G., SIMPSON, P. T. & LAKHANI, S. R. 2012. Gene expression profiling of tumour epithelial and stromal compartments during breast cancer progression. *Breast Cancer Res Treat,* 135**,** 153-65.
VELCULESCU, V. E., MADDEN, S. L., ZHANG, L., LASH, A. E., YU, J., RAGO, C., LAL, A., WANG, C. J., BEAUDRY, G. A., CIRIELLO, K. M., COOK, B. P., DUFAULT, M. R., FERGUSON, A. T., GAO, Y., HE, T. C., HERMEKING, H., HIRALDO, S. K., HWANG, P. M., LOPEZ, M. A., LUDERER, H. F., MATHEWS, B., PETROZIELLO, J. M., POLYAK, K., ZAWEL, L., KINZLER, K. W. & ET AL. 1999. Analysis of human transcriptomes. *Nat Genet,* 23**,** 387-8.
VERGHESE, E. T., SHENOY, H., COOKSON, V. J., GREEN, C. A., HOWARTH, J., PARTANEN, R. H., POLLOCK, S., WATERWORTH, A., SPEIRS, V., HUGHES, T. A. & HANBY, A. M. 2011. Epithelial-mesenchymal interactions in breast cancer: evidence for a role of nuclear localized beta-catenin in carcinoma-associated fibroblasts. *Histopathology,* 59**,** 609-18.
VETTO, J., JUN, S. Y., PADUCH, D., EPPICH, H. & SHIH, R. 1999. Stages at presentation, prognostic factors, and outcome of breast cancer in males. *Am J Surg,* 177**,** 379-83.
VOLM, M. D. 2003. Male breast cancer. *Curr Treat Options Oncol,* 4**,** 159-64.
VOROBIOF, D. A. & FALKSON, G. 1987. Nasally administered buserelin inducing complete remission of lung metastases in male breast cancer. *Cancer,* 59**,** 688-9.
WAGNER, D. D., IVATT, R., DESTREE, A. T. & HYNES, R. O. 1981. Similarities and differences between the fibronectins of normal and transformed hamster cells. *J Biol Chem,* 256**,** 11708-15.
WAGNER, J. L., THOMAS, C. R., JR., KOH, W. J. & RUDOLPH, R. H. 1995. Carcinoma of the male breast: update 1994. *Med Pediatr Oncol,* 24**,** 123-32.
WANG-RODRIGUEZ, J., CROSS, K., GALLAGHER, S., DJAHANBAN, M., ARMSTRONG, J. M., WIEDNER, N. & SHAPIRO, D. H. 2002. Male breast carcinoma: correlation of ER, PR, Ki-67, Her2-Neu, and p53 with treatment and survival, a study of 65 cases. *Mod Pathol,* 15**,** 853-61.
WASIELEWSKI, M., DEN BAKKER, M. A., VAN DEN OUWELAND, A., MEIJER-VAN GELDER, M. E., PORTENGEN, H., KLIJN, J. G.,

MEIJERS-HEIJBOER, H., FOEKENS, J. A. & SCHUTTE, M. 2009. CHEK2 1100delC and male breast cancer in the Netherlands. *Breast Cancer Res Treat,* 116**,** 397-400.
WEBER-CHAPPUIS, K., BIERI-BURGER, S. & HURLIMANN, J. 1996a. Comparison of prognostic markers detected by immunohistochemistry in male and female breast carcinomas. *European journal of cancer,* 32A**,** 1686-92.
WEBER-CHAPPUIS, K., BIERI-BURGER, S. & HURLIMANN, J. 1996b. Comparison of prognostic markers detected by immunohistochemistry in male and female breast carcinomas. *Eur J Cancer,* 32A**,** 1686-92.
WEISS, H. A., DEVESA, S. S. & BRINTON, L. A. 1996. Laterality of breast cancer in the United States. *Cancer Causes Control,* 7**,** 539-43.
WEISS, J. R., MOYSICH, K. B. & SWEDE, H. 2005. Epidemiology of male breast cancer. *Cancer Epidemiol Biomarkers Prev,* 14**,** 20-6.
WHITE, J., KEARINS, O., DODWELL, D., HORGAN, K., HANBY, A. M. & SPEIRS, V. 2011. Male breast carcinoma: increased awareness needed. *Breast Cancer Res,* 13**,** 219.
WILLIAMS, C. M., ENGLER, A. J., SLONE, R. D., GALANTE, L. L. & SCHWARZBAUER, J. E. 2008. Fibronectin expression modulates mammary epithelial cell proliferation during acinar differentiation. *Cancer Res,* 68**,** 3185-92.
WILLSHER, P. C., LEACH, I. H., ELLIS, I. O., BELL, J. A., ELSTON, C. W., BOURKE, J. B., BLAMEY, R. W. & ROBERTSON, J. F. 1997a. Male breast cancer: pathological and immunohistochemical features. *Anticancer Res,* 17**,** 2335-8.
WILLSHER, P. C., LEACH, I. H., ELLIS, I. O., BOURKE, J. B., BLAMEY, R. W. & ROBERTSON, J. F. 1997b. A comparison outcome of male breast cancer with female breast cancer. *Am J Surg,* 173**,** 185-8.
WINCHESTER, D. J. 1996. Male breast cancer. *Semin Surg Oncol,* 12**,** 364-9.
WISEMAN, B. S. & WERB, Z. 2002. Stromal effects on mammary gland development and breast cancer. *Science,* 296**,** 1046-9.
WOODWARD, T. L., MIENALTOWSKI, A. S., MODI, R. R., BENNETT, J. M. & HASLAM, S. Z. 2001. Fibronectin and the alpha(5)beta(1) integrin are under developmental and ovarian steroid regulation in the normal mouse mammary gland. *Endocrinology,* 142**,** 3214-22.
YAMADA, K. M. 1991. Fibronectin and Other Cell Interactive Glycoproteins. *Cell Biology of the Extracellular Matrix.* 2nd ed. New York: Plenum press.
YAMADA, K. M. & KENNEDY, D. W. 1979. Fibroblast cellular and plasma fibronectins are similar but not identical. *J Cell Biol,* 80**,** 492-8.
YANG, J., REN, Y., WANG, L., LI, B., CHEN, Y., ZHAO, W., XU, W., LI, T. & DAI, F. 2010. PTEN mutation spectrum in breast cancers and breast hyperplasia. *J Cancer Res Clin Oncol,* 136**,** 1303-11.
YANG, Z., ZHANG, X., GANG, H., LI, X., LI, Z., WANG, T., HAN, J., LUO, T., WEN, F. & WU, X. 2007. Up-regulation of gastric cancer cell invasion by Twist is accompanied by N-cadherin and fibronectin expression. *Biochem Biophys Res Commun,* 358**,** 925-30.
YAO, E. S., ZHANG, H., CHEN, Y. Y., LEE, B., CHEW, K., MOORE, D. & PARK, C. 2007. Increased beta1 integrin is associated with decreased survival in invasive breast cancer. *Cancer Res,* 67**,** 659-64.

YOUNG, I. E., KURIAN, K. M., MACKENZIE, M. A., KUNKLER, I. H., COHEN, B. B., HOOPER, M. L., WYLLIE, A. H. & STEEL, C. M. 2000. The CAG repeat within the androgen receptor gene in male breast cancer patients. *J Med Genet,* 37**,** 139-40.

YOUNIS, T., HACHE, K. D., RAYSON, D., DEWAR, R., GRAY, S. & BARNES, P. J. 2009. Survivin and COX-2 expression in male breast carcinoma. *Breast,* 18**,** 228-32.

ZARDI, L., CARNEMOLLA, B., SIRI, A., PETERSEN, T. E., PAOLELLA, G., SEBASTIO, G. & BARALLE, F. E. 1987. Transformed human cells produce a new fibronectin isoform by preferential alternative splicing of a previously unobserved exon. *EMBO J,* 6**,** 2337-42.

APPENDICES

6.1.1 Appendix 1: Ethics approval

NHS

Leeds (West) Research Ethics Committee
6th Floor, Wellcome Wing
Leeds General Infirmary
Great George Street
Leeds
LS1 3EX

Telephone: 0113 3923181
Facsimile: 0113 392 2863

24 July 2006

Dr Valerie Speirs
Senior Lecturer
University of Leeds
Leeds Institute of Molecular Medicine
Wellcome Trust Brenner Building
St James's University Hospital, Leeds
LS9 7TF

Dear Dr Speirs

Full title of study: **Tissue microarray construction for use in identifying prognostic and predictive factors in male breast disease**

REC reference number: **06/Q1205/156**

The Research Ethics Committee reviewed the above application at the meeting held on 14 July 2006.

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation.

Ethical review of research sites

The favourable opinion applies to the research sites listed on the attached form.

Conditions of approval

The favourable opinion is given provided that you comply with the conditions set out in the attached document. You are advised to study the conditions carefully.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Application	1	23 June 2006
Investigator CV		21 June 2006
Protocol	1	
Covering Letter		21 June 2006
Summary/Synopsis		

Research governance approval

The study should not commence at any NHS site until the local Principal Investigator has obtained final research governance approval from the R&D Department for the relevant NHS care organisation.

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

06/Q1205/156	**Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project

Yours sincerely

Laura Sawiuk
REC Co-ordinator
On Behalf of
Dr John Puntis
Chair

Email: laura.sawiuk@leedsth.nhs.uk

Enclosures: *List of names and professions of members who were present at the meeting and those who submitted written comments*

Standard approval conditions

Site approval form (SF1)

Copy to: *C Skinner*
University of Leeds
Research Office
Room 7.11 Worsley Building
University of Leeds
LS2 9JT

Research and Development Department
Leeds Teaching Hospitals NHS Trust

6.1.2 Appendix 2: Breakdown of patients selected for statistical analysis

Table 18: Selection of patients from MBC TMA cores for final statistical analysis

Characteristics	Frequency	Details
Total Number of patients in whom TMA cores were available	**428**	
Missing core	19	MBC 067, 087, 089, 091, 216,217, 227, 228, 231, 236, 240, 241, 252, 260, 261, 301, 317, 429, 486
Total number of patients available for analysis	**409**	
Clinical and/or Pathological data missing	28	
Survival data missing	37	
Both survival and Clinical and/or Pathological data missing	143	
Total number analysed for descriptive/association analysis	**238**	**409 - 171 (143+28) = 238**
Total number analysed for survival analysis	**229**	**409 - 180 (143+37) = 229**

6.1.3 Appendix 3: RNA quality control standards

Table 19: ALMAC diagnostic expression quality control acceptance criteria

Sample type	Almac diagnostic QC criteria	
	Spectrophotometer A260-280	Bio-analyser
Total RNA*	1.68 – 2.08 (1.38 -2.18)**	Two distinct peaks (18S and 28S)
NuGEN FFPE processing		
cDNA	> 1.70	NA
Fragmented DNA	NA	≥ 80% of product should be 200 bases or less

*Not applicable to RNA extracted from FFPE material. **Tolerance limits: Almac diagnostics processed samples within the tolerance limits if accompanied with a satisfactory bio-analyser profile.

Table 20: RNA quality control

RNA Sample Quality

Sample Passes
Within Tolerance
Sample Fails

Biomaterial	FFPE	material
Eluate volume:	23.5	µl
Recommended Concentration for Process	9.5	ng/µl

† Not applicable to FFPE Material

Client Sample ID	RNA Sample Name	Spectrophotometer QC				Comments	Batch Number	Operator Number
		Conc (ng/µl)	Yield µg	A260/280	Pass / Fail			
Almac 21	S0712R0021	240.72	5.66	2.03	Pass	N/A	1	1
Almac 1	S0712R0001	11.52	0.27	1.77	Pass	N/A	1	1
Almac 17	S0712R0017	19.93	0.47	1.87	Pass	N/A	1	1
Almac 7	S0712R0007	38.92	0.91	1.87	Pass	N/A	1	1
Almac 20	S0712R0020	90.47	2.13	2.05	Pass	N/A	1	1
Almac 3	S0712R0003	7.76	0.18	2.1	Fail	Sample Vacufuged	1	1
Almac 13	S0712R0013	43.24	1.02	1.85	Pass	N/A	1	1
Almac 6	S0712R0006	36.57	0.86	1.75	Pass	N/A	1	1
Almac 25	S0712R0025	89.75	2.11	2.03	Pass	N/A	1	1
Almac 14	S0712R0014	9.12	0.21	2.04	Fail	Sample Vacufuged	1	1
Almac 4	S0712R0004	66.61	1.57	2	Pass	N/A	1	1
Almac 12	S0712R0012	141.44	3.32	2.09	Pass	N/A	1	1
Almac 9	S0712R0009	65.31	1.53	1.98	Pass	N/A	1	1
Almac 23	S0712R0023	132.48	3.11	2	Pass	N/A	1	1
Almac 24	S0712R0024	77.9	1.83	1.96	Pass	N/A	1	1
Almac 16	S0712R0016	15.12	0.36	2.08	Pass	N/A	1	1
Almac 10	S0712R0010	70.21	1.65	1.98	Pass	N/A	1	1
Almac 8	S0712R0008	35.93	0.84	1.98	Pass	N/A	1	1
Almac 15	S0712R0015	99.87	2.35	2.06	Pass	N/A	1	1
Almac 5	S0712R0005	37.81	0.89	1.82	Pass	N/A	1	1
Almac 22	S0712R0022	238.02	5.59	2.04	Pass	N/A	1	1
Almac 11	S0712R0011	10.07	0.24	1.88	Pass	N/A	1	1
Almac 18	S0712R0018	19.85	0.47	2	Pass	N/A	1	1
Almac 2	S0712R0002	24.73	0.58	1.85	Pass	N/A	1	1
Almac 19	S0712R0019	62.68	1.47	2	Pass	N/A	1	1

Vacufuged RNA Sample Quality

Sample Passes
Sample is within tolerance limts
Sample Fails

Eluate volume:	13	µl
Recommended Concentration for Process	9.5	ng/µl

† Not applicable to FFPE Material

Client Sample ID	Pool & VacSample Name	Spectrophotometer QC				
		Conc (ng/µl)	Yield ng	A260/280	Pass / Fail	Comments
Almac 3	S0712R0003_vac	10.72	139.36	1.87	Pass	N/A
Almac 14	S0712R0014_vac	16.86	219.18	1.76	Pass	N/A

6.1.4 Appendix 4: cDNA quality controls

Table 21: cDNA quality control

cDNA Sample Quality

Sample Passes
Sample Fails

Input to Frag \ Exon:	3.5	µg
Eluate volume:	25	µl

† Not applicable to FFPE Material

Client Sample ID	cDNA Sample Name	Spectrophotometer QC						Bioanalyzer QC †		Batch Number	Operator Number
		Conc (ng/µl)	Yield µg	Pass / Fail	A260/280	Pass / Fail	Comments	Pass / Fail / NA	Comments		
Almac 6	S0712A0006a	204.70	5.12	Pass	1.95	Pass		NA		1	1
Almac 13	S0712A0013a	169.50	4.24	Pass	1.96	Pass		NA		1	1
Almac 3	S0712A0003a	111.10	2.78	Fail	1.93	Pass	Sample processed at lower input for fragmentation and labelling due to limited RNA available for re-amplification	NA		1	1
Almac 20	S0712A0020a	186.30	4.66	Pass	1.89	Pass		NA		1	1
Almac 7	S0712A0007a	140.10	3.50	Pass	1.97	Pass		NA		1	1
Almac 17	S0712A0017a	159.30	3.98	Pass	1.95	Pass		NA		1	1
Almac 1	S0712A0001a	146.00	3.65	Pass	1.88	Pass		NA		1	1
Almac 21	S0712A0021a	175.80	4.40	Pass	1.94	Pass		NA		1	1
Almac 16	S0712A0016a	156.70	3.92	Pass	1.84	Pass		NA		1	1
Almac 24	S0712A0024a	186.00	4.65	Pass	1.90	Pass		NA		1	1
Almac 23	S0712A0023a	167.90	4.20	Pass	1.84	Pass		NA		1	1
Almac 9	S0712A0009a	173.50	4.34	Pass	1.87	Pass		NA		1	1
Almac 12	S0712A0012a	293.80	7.35	Pass	1.84	Pass		NA		1	1
Almac 4	S0712A0004a	217.60	5.44	Pass	1.84	Pass		NA		1	1
Almac 14	S0712A0014a	201.80	5.05	Pass	1.91	Pass		NA		1	1
Almac 25	S0712A0025a	205.20	5.13	Pass	1.89	Pass		NA		1	1
Almac 2	S0712A0002a	149.50	3.74	Pass	1.86	Pass		NA		1	1
Almac 18	S0712A0018a	194.70	4.87	Pass	1.85	Pass		NA		1	1
Almac 11	S0712A0011a	102.80	2.57	Fail	1.87	Pass	Sample re-amplified and pooled, see S0712A0011b	NA		1	1
Almac 22	S0712A0022a	207.40	5.19	Pass	1.85	Pass		NA		1	1
Almac 5	S0712A0005a	162.00	4.05	Pass	1.85	Pass		NA		1	1
Almac 15	S0712A0015a	219.80	5.50	Pass	1.84	Pass		NA		1	1
Almac 8	S0712A0008a	130.20	3.26	Fail	1.85	Pass	Sample re-amplified and pooled, see S0712A0008b	NA		1	1
Almac 10	S0712A0010a	222.40	5.56	Pass	1.86	Pass		NA		1	1
Almac 19	S0712A0019a	169.90	4.25	Pass	1.85	Pass		NA		1	1
Almac 8	S0712A0008b	97.18	2.43	Fail	1.73	Pass	Re-amplifed QC result	NA		2	1
Almac 11	S0712A0011b	69.48	1.74	Fail	1.73	Pass	Re-amplifed QC result	NA		2	1

Pooled & Vacufuged cDNA Sample Quality

Sample Passes
Sample Fails

Input to Frag \ Exon:	3.5	µg
Elution volume:	25	µl

† Not applicable to FFPE Material

Client Sample	cDNA Sample Name	Spectrophotometer QC					
		Conc (ng/µl)	Yield µg	Pass / Fail	A260/280	Pass / Fail	Comments
Almac 8	S0712A0008_vac	198.23	4.96	Pass	1.93	Pass	
Almac 11	S0712A0011_vac	150.76	3.77	Pass	1.95	Pass	

6.1.5 Appendix 5: GeneChip quality control (QC) and data integrity assessment

The variability in the quality of GeneChip profile QC results is typical when the material is derived from FFPE. Hence in order to improve the analysis, data transformation was performed after removing the array quality variable. Principal component analysis (PCA) (Figure 33a and 33b) and hierarchical clustering (Figure 34) was performed for data integrity analysis. In our cohort, 2 samples failed the analysis for background intensity assessed using mean absolute deviation (mad) (Table 22). However, both these samples were included in the final analysis as they passed all other GeneChip QC measures (Table 23 and 24) and data integrity analysis (Figure 33 and 34). Conversely, 3 MBC samples failed the percentage of present call (less than 15%), scaling factor and data integrity tests (Table 23 and 24). Hence these samples were excluded from the final analysis.

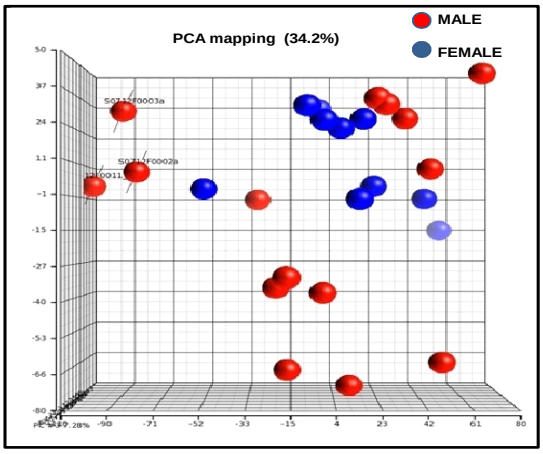


Figure. 33a

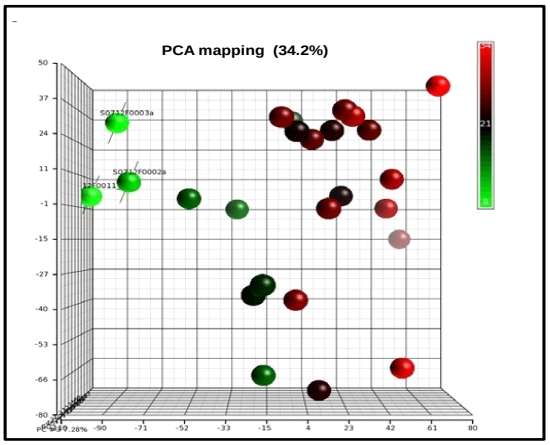


Figure. 33b

Figure 33: PCA plot showing sample groups coloured by gender and by percentage present calls

Data integrity analysis was performed using principal component analysis (PCA). (a) PCA based on gender groups - Red circle = Male, Blue circle = Female. (b) PCA based on percentage present calls -The intensity of the colour depends on the gene expression, Red = Over-expressed and Green = Under-expressed.

Figure 34: Hierarchical clustering showing the relationship between the samples

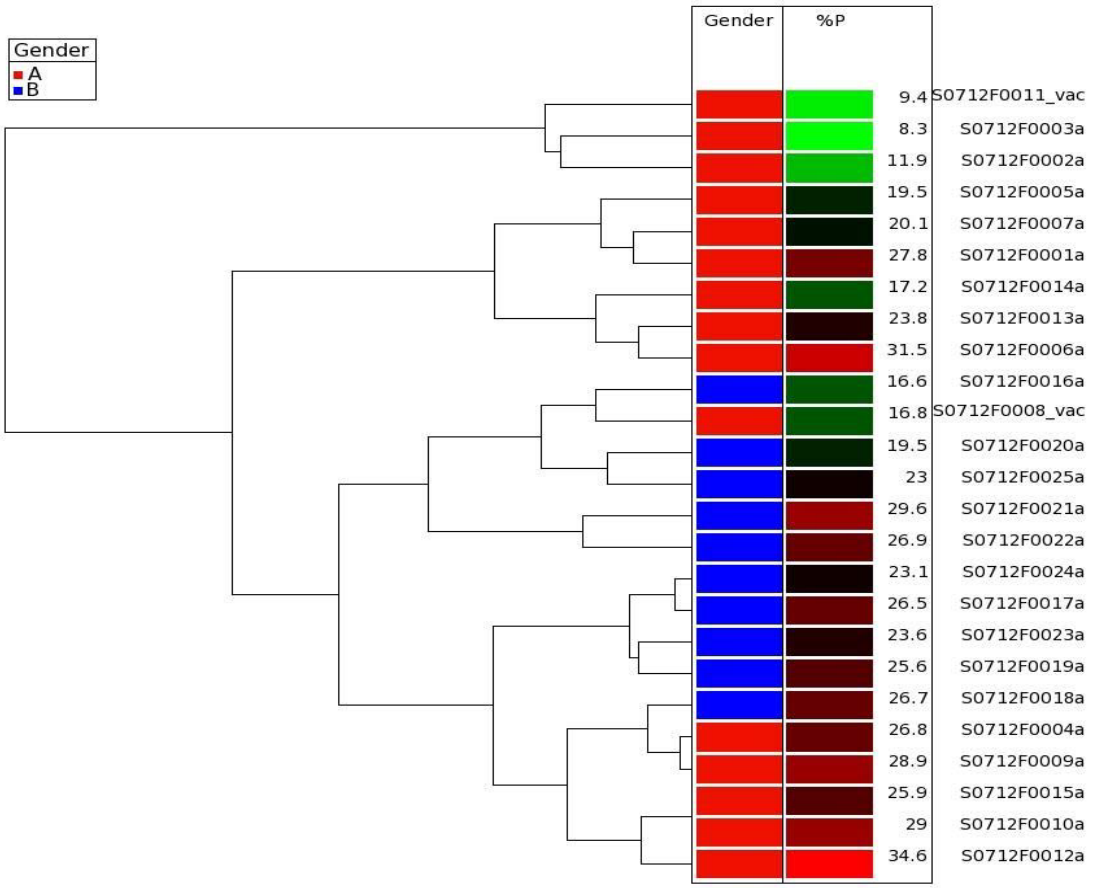


Table 22: Average background intensity value for each chip used in the experiment

Sample Name	Background	Sample Passed?
S0712F0001a	26.46	Yes
S0712F0002a	24.51	Yes
S0712F0003a	26.12	Yes
S0712F0004a	24.95	Yes
S0712F0005a	23.15	Yes
S0712F0006a	23.77	Yes
S0712F0007a	23.68	Yes
S0712F0008_vac	24.29	Yes
S0712F0009a	26.71	Yes
S0712F0010a	26.13	Yes
S0712F0011_vac	24.52	Yes
S0712F0012a	**29.64**	**No**
S0712F0013a	24.82	Yes
S0712F0014a	24.56	Yes
S0712F0015a	24.63	Yes
S0712F0016a	26.08	Yes
S0712F0017a	24.78	Yes
S0712F0018a	25.30	Yes
S0712F0019a	24.59	Yes
S0712F0020a	23.63	Yes
S0712F0021a	23.92	Yes
S0712F0022a	22.70	Yes
S0712F0023a	23.23	Yes
S0712F0024a	23.17	Yes
S0712F0025a	**30.91**	**No**
Median	**24.59**	
Std(mad)=1.4826*mad	**1.349166**	

Table 23: Percentage of present call for the arrays

Sample Name	% Present Call	Sample Passed?
S0712F0001a	27.80	Yes
S0712F0002a	**11.90**	**No**
S0712F0003a	**8.30**	**No**
S0712F0004a	26.80	Yes
S0712F0005a	19.50	Yes
S0712F0006a	31.50	Yes
S0712F0007a	20.10	Yes
S0712F0008_vac	16.80	Yes
S0712F0009a	28.90	Yes
S0712F0010a	29.00	Yes
S0712F0011_vac	**9.40**	**No**
S0712F0012a	34.60	Yes
S0712F0013a	23.80	Yes
S0712F0014a	17.20	Yes

S0712F0015a	25.90	Yes
S0712F0016a	16.60	Yes
S0712F0017a	26.50	Yes
S0712F0018a	26.70	Yes
S0712F0019a	25.60	Yes
S0712F0020a	19.50	Yes
S0712F0021a	29.60	Yes
S0712F0022a	26.90	Yes
S0712F0023a	23.60	Yes
S0712F0024a	23.10	Yes
S0712F0025a	23.00	Yes
Median	**23.80**	
Std(mad)=1.4826*mad	**6.37518**	
Max	**34.60**	
Min	**8.30**	
Range	**26.30**	

Table 24: Scaling factor for the arrays

Sample Name	Scaling Factor	Sample Passed?
S0712F0001a	13.348	Yes
S0712F0002a	**42.061**	**No**
S0712F0003a	**71.461**	**No**
S0712F0004a	24.520	Yes
S0712F0005a	33.118	Yes
S0712F0006a	14.788	Yes
S0712F0007a	31.476	Yes
S0712F0008_vac	39.227	Yes
S0712F0009a	20.923	Yes
S0712F0010a	20.537	Yes
S0712F0011_vac	**51.182**	**No**
S0712F0012a	14.627	Yes
S0712F0013a	20.730	Yes
S0712F0014a	30.940	Yes
S0712F0015a	28.412	Yes
S0712F0016a	33.992	Yes
S0712F0017a	23.976	Yes
S0712F0018a	19.633	Yes
S0712F0019a	24.277	Yes
S0712F0020a	37.876	Yes
S0712F0021a	18.801	Yes
S0712F0022a	22.301	Yes
S0712F0023a	26.298	Yes
S0712F0024a	28.803	Yes
S0712F0025a	21.579	Yes
Median	**24.520**	
Std(mad)=1.4826*mad	**8.4789894**	

6.1.6 Appendix 6: Bioinformatics

The general filtering was carried out using background expression and variance filters to identify a list of 14, 949 probe sets. The criteria used was to remove any probe set with an expression on the background level (p = 0.3). Similarly, any probe set with a variance below the mean global variance was removed in an intensity dependent manner (α-value = 0.9). The data was then subjected to advanced fold change filtering based on the intensity and variation of probe sets between the groups. Student's t-test was performed to determine differentially expressed probe sets statistically and false discovery rate (pFDR) was calculated to account for multiple test correction. Hierarchical agglomerative clustering was undertaken amongst differentially expressed genes, using both less stringent (Significance level for log fold change = 0.05) and stringent criteria (Significance level for log fold change = 0.01). The agglomerative method starts with each sample/gene as a separate cluster and merges them into successive larger clusters. The analysis was performed in Partek GS v6.5.

Functional enrichment analysis of differentially expressed genes identified with the less stringent criteria (n=735) was undertaken using Ingenuity Pathway Analysis (IPA v9.0 build 116623, Ingenuity® Systems, content version 3211) and Gene Ontology (GO) analysis using Functional Enrichment Tool (FET version 1.0). The significance of the association between differentially expressed genes to the total number of genes in a canonical pathway/GO term was expressed as a ratio and Fisher's exact test was performed to establish statistical significance. Benjamini & Hochberg adjusted p-value was also calculated to account for multiple testing and represented as pFDR.

6.1.7 Appendix 7: Publications

1) Sundara Rajan S, Horgan K, Speirs V, Hanby AM. External validation of the ImmunoRatio image analysis application for ERα determination in breast cancer. J Clin Pathol. 2014; 67(1):72-5.
2) Sundara Rajan S, Hanby AM, Horgan K, et al. The potential untility of geminin as a predictive biomarker in breast cancer. Breast Cancer Res Treat. 2014; 143(1):91-8.
3) Peter MB, Shaaban AM, Sundara Rajan S, et al. Investigating and critically appraising the expression and potential role of androgen receptor in breast carcinoma. Hormone Molecular Biology and Clinical Investigation. 2011; 7(1): 273-278
4) Millican-Slater R, Good R, Nash C, Heads JA, Pollock S, Chalkley R, Gomm J, Jones JL, Sundara-Rajan S, Horgan K, Hanby AM, Speirs V. Adding value to rare tissue samples donated to biobanks: Characterisation of breast tissue and primary cell cultures obtained from a female-to-male transgender patient. Cell Tissue Bank. 2015; 16(1):27-34.
5) Sundara Rajan S, Hanby AM, Horgan K, Speirs V. Survivin expression in male breast cancer. SARS oral presentation abstract. BJS 2013; 100 (S4): 2-49 (Page14).