

University of Nottingham

UK | CHINA | MALAYSIA

Deciphering Invasive Breast Cancer Heterogeneity for Personalised Therapy

Maryam Musleh Althobiti

(BSc, MSc)

Thesis submitted to the University of Nottingham for The Degree of Doctorate of Philosophy

January 2021

Abstract

***Background and Aims**:* Characterisation of molecular and cellular heterogeneity in breast cancer (BC) is crucial to refine the prognostic classification of BC and understand the development of treatment resistance. This study investigated the hypotheses that: (i) molecular heterogeneity in terms of expression of BC stem cell (BCSC) markers, oestrogen receptor alpha (ER) and ER-related biomarkers, and the density of tumour-infiltrating lymphocytes (TILs) exists within individual tumours and also between different tumours, and (ii) assessment of such heterogeneity using high-throughput pathological techniques may help to predict tumour progression and patient outcomes.

Materials and Methods: Primary invasive BC tissue specimens (*n* = 1948), including needle core biopsies (NCBs) and surgical excision specimens (SESs), lymph node (LN) metastases and recurrent tumour tissue samples were obtained for well-characterised cohorts of patients with primary invasive BC treated at Nottingham City Hospital between 1989 and 2018. Tissue microarrays were created and single and double immunohistochemistry, RNA *in situ* hybridisation (RNAscope), and haematoxylin and eosin staining were used to investigate inter-tumour and intra-tumour heterogeneity. The Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) cohort was used to assess mRNA and gene copy number changes. BC heterogeneity was quantified using a variety of pathologic assessments.

Results: Analysis of cancer stem cell markers revealed BCSC markers exert different functions in BC and are co-expressed in varied combinations. High ALDH1A1 protein expression was associated with aggressive tumour features and

poorer BC-specific survival (BCSS), especially in the luminal B and triple negative (TN) intrinsic subtypes. ALDH1A1 expression correlated positively with expression of other BCSC markers, including EpCAM and SOX9. High BMI-1 expression was associated with less aggressive tumour features and longer BCSS in ER+ tumours, but shorter BCSS in TNBC. BMI-1 expression correlated negatively with the BCSC markers ALDH1A1 and CD133 in ER+ tumour. Cells co-expressing ALDH1A1+, SOX9+ and EpCAM+ were observed in 12% of BC cases, in 1-10% of tumour cells. The ALDH1A1+SOX9+EpCAM+ BCSC phenotype was associated with advanced nodal stage, aggressive tumour features, ER- and HER2+ tumours, and poorer outcomes.

Examination of ER expression in BC showed high concordance between ER and PgR expression at both the mRNA and protein levels. Approximately half of low ER (1-9%)-expressing NCBs were reclassified as ER-negative on the SESs (<1%); the NCBs of these cases were considered as ER false positives and repeat staining of low ER-expressing tumours is recommended. A high degree of intratumour ER heterogeneity was observed, which reflects the clinical significance of this receptor. Intra-tumoral heterogeneity in ER expression was observed between different regions of the same tumour, with varied patterns of ER staining intensity and distribution. A low ER heterogeneity score was associated with more aggressive tumour features and a shorter disease-free interval; a high ER heterogeneity score and the diffuse distribution pattern was associated with less aggressive features and a longer disease-free interval. The percentage of ER+ cells was higher in LN metastases compared to the paired primary tumours. High expression of the ER+ related marker, SLC39A6, was associated with good prognostic features, longer BCSS—especially in ER+ tumours—and expression of

the ER-related proteins PgR, GATA3 and TFF1. Furthermore, high SLC39A6 expression was associated with longer BCSS and distant metastasis-free survival in patients who did not receive endocrine treatment, but not in patients who received endocrine treatment.

The density of TILs did not vary across tumour blocks from the same case, but differed significantly between tumour regions of the same block. A high percentage average TILs was associated with longer BCSS in TNBC, regardless of whether the patients received adjuvant chemotherapy or not. A high percentage average TILs in recurrent tumours was also associated with shorter post-recurrence survival.

Conclusions: This study confirms the hypothesis that significant heterogeneity exists in BC and can be linked to tumour behaviour. Combined morphological and molecular analyses provided new insight into the characteristics and prognostic and predictive value of heterogeneity in BC. Clinical assessment of heterogeneity in terms of BCSC markers and phenotypes, expression of ER and the density of TILs may help improve prognostication, prediction of response to treatment and clinical decision-making for patients with BC, though validation of these results is necessary before these strategies can be adopted in the clinic.

Abstracts and Publications

Presented abstracts at national and international conferences

1-Zinc Finger Protein 22: a Biomarker Associated with Breast Cancer Metastatic Progression. Maryam Althobiti, Mohammed A Aleskandarany, Andrew R Green, Maria Diez-Rodriguez, Christopher C Nolan, Graham R Ball*, Ian O Ellis, Emad A Rakha. *7th Scientific Meeting of the Pathological Society of Great Britain & Ireland, January 19-20, 2017, London, United Kingdom.* **Poster**

2-The Heterogeneity of Tumour Infiltrating Lymphocytes and its Effect on Patient Outcome. Maryam Althobiti, Mohammed A Aleskandarany, Chitra Joseph, Michael Toss, Nigel Mongan, Maria Diez-Rodriguez[1], Christopher C Nolan, Ibraheem Ashankyty, Ian O Ellis , Andrew R Green, Emad A Rakha. *8th Scientific Meeting of the Pathological Society of Great Britain & Ireland, **20-23 2017**, Belfast, United Kingdom.* **Oral presentation**

3- Heterogeneity of Tumour Infiltrating Lymphocytes in Invasive Breast Cancer. Maryam Althobiti, Mohammed A Aleskandarany, Chitra Joseph, Michael Toss, Nigel Mongan, Maria Diez-Rodriguez[1], Christopher C Nolan, Ibraheem Ashankyty, Ian O Ellis , Andrew R Green, Emad A Rakha.*UK Interdisciplinary Breast Cancer Symposium on 15th-16th January 2018 at Manchester Central Conference Centre.* **Oral presentation**

4- The Prognostic Significant of ALDH1A1 in Breast Cancer Molecular Subtypes. Maryam Althobiti, Rokaya El Ansari, Chitra Joseph, Ian O Ellis, Andrew R Green, Emad A Rakha. *12 Scientific Meeting of the Pathological Society of Great Britain & Ireland, July **2-4 2019**, Belfast, United Kingdom*. **Poster**

5- Expression and Co-Expression of Stem Cell Markers in Breast Cancer: An Immunohistochemical Study. Maryam Althobiti, Michael Toss, Chitra Joesph, Mohammed Alskandarany, Ian O Ellis, Andrew R Green, Emad A Rakha. CRN *(Cancer Research Nottingham) Symposium, on 16th September 2019*. **Poster**

Publications

1-The role of PIP5K1α/pAKT and targeted inhibition of growth of subtypes of breast cancer using PIP5K1α inhibitor. Martuza Sarwar, Azharuddin Sajid, Syed Khaja, Mohammed Aleskandarany, Richard Karlsson, Maryam Althobiti, Niels Ødum, Nigel P. Mongan, Nisthman Dizeyi, Heather Johnson, Andrew R. Green, Ian O. Ellis, Emad A. Rakha, Jenny L Persson.

2- Breast cancer intratumour heterogeneity: current status and clinical implications. Chitra Joseph, Athanasia Papadaki, Maryam Althobiti, Mansour Alsaleem, Mohammed A Aleskandarany, Emad A Rakha.

3- Heterogeneity of tumour-infiltrating lymphocytes in breast cancer and its prognostic significance. Maryam Althobiti, Mohammed A Aleskandarany, Chitra Joseph, Michael Toss, Nigel Mongan, Maria Diez-Rodriguez, Christopher C Nolan, Ibraheem Ashankyty, Ian O Ellis, Andrew R Green, Emad A Rakha.

4- Overexpression of the cancer stem cell marker CD133 confers a poor prognosis in invasive breast cancer. Chitra Joseph, Maariya Arshad, Sasagu Kurozomi,Maryam Althobiti, Islam M. Miligy, Sara Al-izzi, Michael S. Toss, Fang Qin Goh, Simon J. Johnston, Stewart G. Martin, Ian O. Ellis, Nigel P. Mongan, Andrew R. Green, Emad A. Rakha.

5-Enhanced glutamine uptake influences composition of immune cell infiltrates in breast cancer. Rokaya El Ansari, Madeleine L. Craze, Maryam Althobiti, Lutfi Alfarsi, Ian O. Ellis, Emad A. Rakha and Andrew R. Green.

6- The prognostic significance of BMI-1 expression in invasive breast cancer is dependent on its molecular subtypes. Maryam Althobiti, Abir A Muftah, Mohammed A Aleskandarany, Chitra Joseph, Michael S Toss, Andrew Green, Emad Rakha.

7- The prognostic significance of ALDH1A1 expression in early invasive breast cancer. Maryam Althobiti, Rokaya El Ansari, Mohammed Aleskandarany, Chitra Joseph, Michael S Toss, Andrew R Green, Emad A Rakha.

This thesis is dedicated

To the memory of my father

And to my mother

With love and eternal appreciation

Acknowledgements

I would like to take this opportunity to express my heartfelt thanks to my supervisors, Prof Emad Rakha, Dr Andrew Green and Prof Ian Ellis, for their time, guidance and support throughout this project.

I would also like to thank all members of the breast cancer pathology group for their assistance with all aspects of my lab work. Special thanks go to Mr. Chris Nolan, for his valuable technical support, encouraging words and friendly days we had together.

I am also grateful to have been surrounded by intelligent, friendly and supportive colleagues.

Most importantly, none of this work could have happened without my family, especially my mother, who offered her encouragement through daily phone calls and messages and her unconditional love and prayer.

My sincerest appreciation is extended to my sponsor in the Kingdom of Saudi Arabia, Shaqra University, and the Saudi Arabian Cultural Bureau in the United Kingdom for their generous support.

Abbreviations

A	
AML	Acute myeloid leukaemia
ALDH	Aldehyde dehydrogenase activity
ALDH1A1	Aldehyde dehydrogenase 1 family, member A1
ALDH1A2	Aldehyde dehydrogenase 1 family, member A2
ALDH1A3	Aldehyde dehydrogenase 1 family, member A3
AP	Alkaline phosphatase
B	
BC	breast cancer
BCSCs	breast cancer stem cell
BCSS	BC-specific survival
BCSS	breast cancer specific survival
BMI-1	B Lymphoma Mo-MLV Insertion Region 1 Homolog
BRCA1	breast cancer type-1 susceptibility protein
BRCA2	breast cancer type-2 susceptibility protein
C	
CAFs	cancer-associated fibroblasts
CAP	College of American Pathologists
CD133	cluster of differentiation 13
CD20	cluster of differentiation 20
CD24	cluster of differentiation 24
CD3	cluster of differentiation 3
CD44	cluster of differentiation 44
CD68	cluster of differentiation 68
CD8	cluster of differentiation 8
CDH1	downregulate E-cadherin
CGH	comparative genomic hybridization
CI	confidence intervals
CISH	chromogenic *in situ* hybridisation
CK	cytokeratin
cm	centimeter
CN	Copy number
CSC	CSC cancer stem cells
CTLs	cytotoxic T lymphocytes
D	
DAB	3,3'-diaminobenzidine
DBD	DNA-binding domain
DCIS	Ductal carcinoma in situ
DFS	disease-free survival
DHH	Desert Hedgehog
Dll	Delta-like
DMFS	Distant metastasis-free survival
E	
E2	oestrogen 17β-estradiol
E-cadherin	epithelial cadherin

ECM	extracellular matrix
EDI	ecosystem diversity index
EGF	Epidermal growth factor
EGFR	epidermal growth factor receptor
EMT	epithelial-mesenchymal transition
EpCAM	Epithelial cell adhesion molecule
ER	Oestrogen receptor
ER+	ER positive
F	
FAK	focal adhesion kinase
FBS	fetal bovine serum
FISH	fluorescence *in situ* hybridisation
FOXA1	Forkhead box protein A1
FOXP3	Forkhead box P3
frizzled receptors	FZD
G	
GATA3	GATA binding protein 3
GSK-3β	glycogen synthase kinase 3 beta
H	
H&E	haematoxylin and eosin
H_2O	Water
HER2	human epidermal growth factor receptor-2
HER2	Human epidermal growth factor receptor 2
HIF	hypoxia-inducible factor
HR	hazard ratio
HRP	Horse-radish peroxide
HRT	hormone-replacement therapy
H-score	Histological score
I	
IBC	Invasive breast cancer
IHC	Immunohistochemistry
IHH	Indian Hedgehog
ILC	Invasive lobular carcinoma
IMS	industrial methylated spirits
J	
JAG	Jagged
L	
LBD	ligand-binding domain
LN	lymph nodes
LVI	lymphovascular invasion
M	
MAPK	mitogen-activated protein kinase
MAPK	mitogen-activated protein kinase
METABRIC	Molecular Taxonomy of Breast Cancer International Consortium
mg/ml	milligram/milliliter
MHC1	major Histocompatibility complex class I
N	
NCT	Neoadjuvant chemotherapy

NK	natural killer
NPI	Nottingham Prognostic Index
Nrf2a	erythroid-2-related factor 2
NST	no special type
NTD	N-terminal domain
NCB	Needle core biopsy
P	
PAM50	prediction analysis of microarray of 50 genes
PBS	phosphate-buffered saline
pCR	pathological complete response
PD-1	Programmed cell death protein 1
PDL-1	Programmed death-ligand 1
PgR	Progesterone receptor
PPARβ/δ	peroxisome proliferator activated receptor β/δ
PTCH1	Patched 1
PTCH2	Patched 2
R	
RA	retinoic acid
RARs	retinoid receptors
RARα	retinoic acid receptor alpha
RARβ	retinoic acid receptor beta
rpm	Rotation Per Minute
RXRs	retinoid X receptors
S	
SHH	Sonic Hedgehog
SIRT2	Silent Information Regulator 2
SLC39A6	solute carrier family 39 member 6
SMO	Smoothened
SOX10	SRY-Box Transcription Factor 10
SOX9	SRY-Box Transcription Factor 9
SES	surgical excision specimen
T	
T reg	T regulatory cells
TBS	Tris Buffered Saline
TFF1	Trefoil factor 1
TGF-β	Transforming growth factor beta
Th1	T helper 1
Th2	T helper 2
TILs	tumour infiltrating lymphocytes
TMA	tissue microarray
TNBC	Triple negative breast cancer
TNF-α	tumour necrosis factor alpha
V	
V	Voltage
M	
µl	microliters
µm	micrometers

Table of Contents

Abstract ii

Abstracts and Publications v

Acknowledgements viii

Abbreviations ix

Table of Contents xii

List of Tables xx

List of Figures xxv

Chapter 1: Introduction 1

1.1 Introduction to breast cancer 1

1.1.1 Histology of the breast 1

1.1.2 Epidemiology of breast cancer 2

1.1.3 Factors associated with the risk of BC 3

1.1.3.1 Age 3

1.1.3.2 Family history and genetic factors 4

1.1.3.3 Breast tissue density 4

1.1.3.4 Hormone-related factors 5

1.1.3.5 Lifestyle factors 5

1.2 Prognostic and predictive tools in BC 6

1.2.1 Tumour size 6

1.2.2 Histologic grade 7

1.2.3 Nodal status and tumour stage 8

1.2.4 Nottingham Prognostic Index 8

1.2.5 Lymphovascular invasion 10

1.2.6 ER/PgR status 10

1.2.7 Oestrogen receptor in BC 10

1.2.8 HER2/neu 13

1.3 Classification of BC 14

1.3.1 Molecular classifications 14

1.3.2 Histological subtypes 15

1.4 The tumour microenvironment in IBC 17

1.5 Role of immune cells in BC 19

1.5.1 Prognostic significance of TIL subtypes in IBC 20

1.6 Intratumour heterogeneity in IBC 24

1.7 Cancer stem cell hypothesis 26
1.7.1 Cancer stem cell phenotype in BC 27
1.7.2 Signalling pathways that regulate BCSCs 28
1.7.2.1 Hedgehog signalling 29
1.7.2.2 Wnt signalling 29
1.7.2.3 Notch signalling 30
1.7.3 BCSCs and metastasis 32
1.8 Clonal evolution theory and clinical implications 34
1.9 Genomic instability and heterogeneity 35
1.10 Spatial and temporal heterogeneity in BC 37
1.11 Hypothesis 41
1.12 Aims 42

Chapter 2: Materials and Methods 43
2.1 Study cohorts 43
2.1.1 Nottingham patient cohorts 43
2.1.1.1 Confirmation of ER, PgR and HER2 status for the Nottingham cohorts 45
2.1.1.2. Availability of data on biomarker expression for the primary Nottingham cohort 49
2.1.2 Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) cohort 50
2.2 Assessment of tumour burden for tissue specimens from the Nottingham cohort 53
2.2.1 Haematoxylin and eosin (H&E) staining 53
2.3 Immunohistochemical staining (IHC) of tumour sections from the Nottingham cohorts 55
2.3.1 Construction of tissue microarrays 55
2.3.2 Validation of antibody specificity by Western blotting 56
2.3.2.1 Preparation of cell lysates for Western blotting 56
2.3.2.2 Electrophoresis and transfer 57
2.3.2.3 Alternative enhanced chemiluminescence (ECL) detection method 59
2.3.3. Automated IHC using the Ventana Autostainer 61
2.3.4 Manual immunohistochemistry 62
2.3.4.1 Antigen retrieval 62
2.3.4.2 Manual IHC process 63
2.3.5 Optimisation of the antibodies for IHC 65
2.3.6 Assessment of IHC staining using the H-score 65
2.3.7 Double IHC staining 66
2.3.7.1 Experimental steps 66
2.3.7.2 Optimisation of antibodies for double IHC 68

2.4 Chromogenic *in situ* hybridization assay 69

2.5 Statistical analysis 70

2.5.1 Dichotomisation of data 71

2.5.2 Analysis of the associations between clinicopathological parameters and biomarker expression 71

2.5.3 Univariate analysis and multivariate analysis 71

Chapter 3: Prognostic significance of ALDH1A1 expression in invasive breast cancer ... 73

3.1 Introduction 73

3.1.1 Hypotheses 77

3.1.2 Aims of the study 77

3.1.3 Objectives of the study 77

3.2 Materials and Methods 78

3.2.3 Characteristics of the study cohorts 78

3.2.4 Availability of biomarker data 78

3.2.1 ALDH1A1 antibody validation 80

3.2.2 Tissue microarrays and immunohistochemical analysis 80

3.2.5 Statistical analysis 81

3.3 Results 82

3.3.1 ALDH1A1 expression in BC 82

3.3.2 ALDH1A1 protein expression in BC 83

3.3.3 ALDH1A1 mRNA expression and CN changes 85

3.3.4 Associations between ALDH1A1 expression and clinicopathological features 85

3.3.6 Correlations between ALDH1A1 expression and the expression of other BCSC markers 87

3.3.7 Associations between ALDH1A1 expression and patient outcomes 88

3.3.8 Multivariate analysis 90

3.4 Discussion 92

3.5 Conclusion 97

Chapter 4: The prognostic significance of BMI-1 expression varies in different subtypes of invasive breast cancer 98

4.1 Introduction 98

4.1.1 Hypotheses 100

4.1.2 Aims 100

4.1.3 Objectives 100

4.2 Materials and Methods 102

4.2.1 Study cohort for analysis of BMI-1 expression 102

4.2.2 Availability of data on biomarker expression 102

4.2.3 Antibody validation 102

4.2.4 IHC of BMI-1 protein expression 103

4.2.5 Assessment of BMI-1 protein expression 103

4.2.6 Statistical analysis 103

4.3 Results 104

4.3.1 Validation of the BMI-1 antibody 104

4.3.2 BMI-1 protein expression 105

4.3.3 BMI-1 mRNA expression 106

4.3.4. BMI-1 expression in BC molecular subtypes 106

4.3.5 Correlations between expression of BMI-1 and expression other biomarkers at the protein and mRNA levels 107

4.3.6 Associations between BMI-1 expression and the clinicopathological features of BC 109

4.3.7 Associations between BMI-1 protein expression and patient outcomes 112

4.3.8 Multivariate analysis 116

4.4 Discussion 118

4.5 Conclusion 121

Chapter 5: Identification of a breast cancer stem cell phenotype: an immunohistochemical study 122

5.1 Introduction 122

5.1.1 Hypothesis 123

5.1.2 Study aims 124

5.1.3 Objectives 124

5.2 Materials and Methods 125
5.2.1 Patient cohort 125
5.2.2 Availability of biomarker data for the cohort 125
5.2.3 Validation of antibody specificity by western blotting 127
5.2.4 Double IHC staining 128
5.2.5 Double IHC scoring 128
5.2.6 Statistical analyses 130
5.3 Results 131
5.3.1 BCSC marker expression in BC tissues 131
5.3.2 Correlations between expression of the BCSC markers 133
5.3.3 Associations between EpCAM and SOX9 expression and the clinicopathological parameters of BC 135
5.3.4 Survival analysis for SOX9 and EpCAM expression 137
5.3.4.1 Breast cancer-specific survival 137
5.3.4.2 Disease-free interval 137
5.3.4.3 Distant metastasis-free survival 142
5.3.5 Co-expression of the biomarkers ALDH1A1, SOX9 and EpCAM in BC 145
5.3.6 Co-localisation of ALDH1A1+, SOX9+ and EpCAM+ expression in BC tissue samples using double IHC 148
5.3.7 Associations between the ALDH1A1+SOX9+EpCAM+ phenotype and the clinicopathological features of BC 151
5.3.8 Associations between the ALDH1A1+SOX9+EpCAM+ BCSC phenotype and patient outcomes in BC 154
5.4 Discussion 159
5.5 Conclusion 163

Chapter 6: Refining Oestrogen Receptor (ER) Positivity in Invasive Breast Cancer 164
6.1 Introduction 164
6.1.1 Hypothesis 165
6.1.2 Aims of the chapter 166
6.1.3 Objectives 166
6.2 Materials and Methods 167
6.2.1 Patients and data collection 167
6.2.2 Pathological evaluation of ER expression 167
6.2.3 IHC staining 168
6.2.4 IHC assessment 169
6.2.5 RNAscope in situ hybridisation assay 169
6.2.6 RNAscope assessment 170

6.2.7 Statistical analysis **171**

6.3 Results **172**

6.3.1 ER expression in NCBs and SESs **172**

6.3.2 Interobserver agreement in ER scoring between pathologists **173**

6.3.3 Characteristics of IHC staining **173**

6.3.4 Comparison of ER antibody clones **176**

6.3.5 PgR expression in different ER+ subgroups **178**

6.3.6 Association between the RNAscope and IHC results for ER and PgR **179**

6.3.7 Comparison of the clinicopathological characteristics and outcomes of the low ER+ subgroup and other subgroups **179**

6.3.8 Patient outcome analysis **183**

6.4 Discussion **184**

6.5 Conclusion **188**

Chapter 7: Quantification and clinical impacts of intratumor ER heterogeneity in breast cancer **189**

7.1 Introduction **189**

7.1.2 Hypothesis **191**

7.1.3 Aims of the study **191**

7.1.4 Objectives of the study **191**

7.2 Materials and Methods **193**

7.2.1. Study cohort **193**

7.2.2 IHC staining **193**

7.2.3 Assessment of intratumoural ER heterogeneity **194**

7.2.4 Intratumoral ER heterogeneity index **196**

7.2.5 Statistical analysis **196**

7.3.1 Intra-tumour heterogeneity of ER expression in BC **198**

7.3.2 Co-expression of ER and PgR in primary IBC **203**

7.3.3 ER heterogeneity in primary tumours and the paired lymph node metastases **205**

7.3.4 Intratumor ER heterogeneity index **206**

7.3.5 Clinical significance of intratumour ER heterogeneity **208**

7.4 Discussion **215**

Chapter 8: Oestrogen-regulated protein SLC39A6 is a prognostic marker associated with good outcome in luminal breast cancer 219

8.1 Introduction 219

8.1 1 Hypothesis 221

8.1.2 Aims 221

8.1.3 Objectives 221

8.2 Materials and Methods 222

8.2.1 Study cohorts 222

8.2.2 Assessment of SLC39A6 protein expression in the Nottingham cohort 224

8.2.3 Existing immunohistochemical analysis for the Nottingham cohort 224

8.2.4 Statistical analysis 224

8.3 Results 226

8.3.1 SLC39A6 protein expression in BC 226

8.3.2 SLC39A6 mRNA expression in BC 228

8.3.3 Associations between SLC39A6 expression and the clinicopathological features of BC 229

8.3.4 Associations between expression of SLC39A6 and other ER-related biomarkers 233

8.3.5 Outcome analysis 235

8.3.6 Multivariate analysis 239

8.3.7 Associations between SLC39A6 and the response to endocrine therapy 242

8.5 Conclusion 248

Chapter 9: Heterogeneity and prognostic significance of tumour infiltrating lymphocytes (TILs) in invasive breast cancer 249

9.1 Introduction 249

9.1.1 Hypothesis 251

9.1.2 Aims 251

9.1.3 Objectives 251

9.2 Materials and Methods 253

9.2.1 Characteristics of the study cohorts 253

9.2.2 Availability of other data on immune markers for the cohort 253

9.2.3 Assessment of TILs in H&E-stained sections 254

9.2.4 Statistical analysis 255

9.3 Results 257

9.3.1 Heterogeneity of TILs at the morphological level 257

9.3.2 Immune cell subset counts and overall TILs in BC 261

9.3.3 PDL-1 and PD1 expression and TILs in BC 264

9.3.4 Prognostic significance of TILs at the morphological level in BC 267

9.3.5 Prognostic significance of immune cell subsets in BC 272

9.3.6 Prognostic significance of PDL-1 and PD1 expression in BC 276

9.4 Discussion 281

9.5 Conclusion 287

Chapter 10: General Discussion 288

10.1 General Discussion 288

10.2 Conclusions 297

10.3 Limitations 298

10.4 Future recommendations 299

10.5 10.5 Ethical approval 285

References 301

List of Tables

Table 2.1: Clinicopathological features of the primary, secondary and low oestrogen receptor (ER) Nottingham cohorts. 44

Table 2.2: Clinicopathological features of the Nottingham oestrogen receptor (ER) heterogeneity cohort. 45

Table 2.3: Antibodies used for immunohistochemical analysis of the Nottingham cohort. 46

Table 2.4: Clinicopathological features of the METABRIC cohort 48

Table 2.5. Details of the antibodies used in this thesis. 56

Table 3.1: Summary of the patient demographics of the cohorts used to analyse ALDH1A1 protein and mRNA expression. 74

Table 3.2: Associations between ALDH1A1 protein and mRNA expression and clinicopathological features in breast cancer* 81

Table 3.3: Spearman's rank analysis of the correlations between ALDH1A1 protein and mRNA expression and the expression of other breast cancer stem cell (BCSC) markers. 83

Table 3.4: Multivariate Cox regression hazard model of the associations between ALDH1A1 protein and mRNA levels and other prognostic clinicopathological features and breast cancer-specific survival in the entire breast cancer cohorts. 86

Table 3.5: Multivariate Cox regression hazard model of the associations between ALDH1A1 protein levels and other prognostic clinicopathological features and breast cancer-specific survival in the luminal B and triple negative (TN) breast cancer subtypes. 86

Table 4.10: Multivariate Cox regression hazard model including other prognostic clinicopathological parameters shows that high BMI1 (mRNA) provided a dependent prognostic factor in all BC and ER+ tumour. 110

Table 4.3: BMI-1 protein expression H-scores and mRNA log fold changes in different molecular subtypes of breast cancer. 100

Table 4.4: Correlations between BMI-1 protein expression and the protein expression of other breast cancer stem cell markers in the entire Nottingham cohort and luminal ER+ subtype....101

Table 4.5: Correlations between BMI-1 mRNA expression and the mRNA expression of other breast cancer stem cell markers in the entire METABRIC cohort and luminal ER+ BC. 102

Table 4.6: Associations between BMI-1 protein expression and the clinicopathological features of breast cancer. 103

Table 4.7: Associations between BMI-1 mRNA expression and the clinicopathological features of breast cancer. 104

Table 4.8: Multivariate Cox regression hazard model of prognostic clinicopathological parameters for the entire Nottingham BC cohort and the subgroup of patients with ER+ tumours. 109

Table 4.9: Multivariate Cox regression hazard model of prognostic clinicopathological features in TNBC. 110

Table 5.1: Demographic features of the Nottingham patient cohort 118

Table 5.2: Mean and median percentages of cells expressing the tested breast cancer stem cell (BCSC) markers in the entire breast cancer (BC) cohort and oestrogen receptor negative (ER-), oestrogen receptor positive (ER+) and human epidermal growth factor receptor 2 positive (HER2+).....124

Table 5.3. Associations between ALDH1A1, SOX9 and EpCAM expression in breast cancer (BC) (chi squared test) 126

Table 5.4: Associations between SOX9 and EpCAM expression and the clinicopathological features of the entire breast cancer cohort. 128

Table 5.5: Associations between ALDH1A1+SOX9+EpCAM+ co-expression and the clinicopathological features of breast cancer (BC) in the entire cohort. 136

Table 5.6: Percentages of cells co-expressing ALDH1A1+SOX9+ and ALDH1A1+EpCAM+ in the entire breast cancer (BC) cohort and ER-, ER+ and HER2+ subtypes. 139

Table 5.7: Associations between the ALDH1A1+SOX9+EpCAM+ breast cancer stem cell (BCSC) phenotype and the clinicopathological features of breast cancer (BC) in the entire cohort and human epidermal growth factor receptor 2 positive (HER+) subgroup. 141

Table 5.8: Associations between the ALDH1A1+SOX9+EpCAM+ breast cancer stem cell (BCSC) phenotype and the clinicopathological features of oestrogen receptor negative (ER-) and oestrogen receptor positive (ER+) tumours.....142

Table 5.9: Multivariate Cox regression hazard model of the associations between the ALDH1A1+SOX9+EpCAM+ cell phenotype and other prognostic clinicopathological features in the entire cohort and oestrogen receptor negative (ER-) and human epidermal growth factor receptor 2 positive (HER2+) subgroups.....146

Table 6.1: Criteria for RNAscope scoring 160

Table 6.2: Oestrogen expression (ER) (percentage) in needle core biopsies (NCBs) and the matched surgical excision specimens (SESs). 161

Table 6.3: Number of cases in each subgroup reclassified to 0 or < 1% based on the surgical excision specimens (SESs).....163

Table 6.4: Comparison of the numbers of ER+ cells for needle core biopsies (NCBs) and the matched surgical excision specimens (SESs) stained using the EP1 antibody and the SP1 antibody with/without a protein block.....165

Table 6.5: Progesterone receptor (PgR) expression in different oestrogen receptor positive (ER+) subgroups. 167

Table 6.6. Comparison of the immunohistochemistry analysis of oestrogen receptor expression and RNAscope assay of ESR1 status for surgical excision specimens (SESs). 168

Table 6.7: Comparison of the clinicopathological features of the < 1% Oestrogen positive (ER+) versus 1-9% Oestrogen positive (ER+) subgroups. 171

Table 6.8: Comparison of the clinicopathological features of the 1-9% Oestrogen positive (ER+) versus 10-100% Oestrogen positive (ER+) subgroups. 172

Table 7.1 Spearman's rank analysis of the correlation between the percentages of cells with various ER+ staining intensities at the regional level in primary breast cancer. 188

Table 7.2. Percentages of cells (median, mean and range) exhibiting strong, weak, very weak and negative oestrogen receptor (ER) and progesterone receptors (PgR) staining in the 111 primary breast cancer (BC) tumours. 192

Table 7.3. Clinicopathological features of the tumours stratified by the patterns of oestrogen receptor (ER) distribution in 111 cases of primary breast cancer (BC). 197

Table 7.4: The association between intratumor oestrogen receptor (ER) heterogeneity and clinicopathological features in 111 cases of primary breast cancer (BC)* 198

Table 7.5. Association between intratumor oestrogen receptor (ER) heterogeneity and ER homogeneity expression with the clinicopathological features of 502 cases of primary breast cancer (BC). 199

Table 8.1: Summary of the demographic features of the cohorts assessed in this chapter. 210

Table 8.10: Multivariate Cox regression hazard model for for breast cancer specific survival including other ER-related markers and SLC39A6 protein expression in the oestrogen receptor positive subgroup. 229

Table 8.11: Multivariate Cox regression hazard model for breast cancer specific survival including other prognostic clinicopathological parameters and SLC39A6 mRNA expression in the entire cohort and oestrogen receptor positive subgroup. 229

Table 8.12: Multivariate Cox regression hazard model for breast cancer specific survival including other oestrogen receptor related markers and SLC39A6 mRNA expression in the oestrogen receptor positive subgroup. 230

Table 8.2: Mean, median and ranges of SLC39A6 mRNA and protein expression in the ER+ and ER- subgroups of BC. 216

Table 8.3: Associations between SLC39A6 cytoplasmic and nuclear protein expression and the clinicopathological features of breast cancer. 218

Table 8.4: Associations between nuclear and cytoplasmic SLC39A6 protein expression and the clinicopathological features of oestrogen receptor positive tumours. 219

Table 8.5: Associations between SLC39A6 mRNA expression and the clinicopathological features of BC in the entire METABRIC cohort and oestrogen receptor positive tumours. 220

Table 8.6: Associations between cytoplasmic and nuclear SLC39A6 protein expression and expression of other oestrogen receptor-related markers. 221

Table 8.7: Associations between SLC39A6 mRNA expression and the mRNA expression levels of other oestrogen receptor-related markers 222

Table 8.8: Multivariate Cox regression hazard model for breast cancer specific survival including prognostic clinicopathological parameters and SLC39A6 nuclear protein expression in the entire cohort. 228

Table 8.9: Multivariate Cox regression hazard model of for breast cancer specific survival including prognostic clinicopathological features and SLC39A6 protein expression in the oestrogen receptor positive subgroup. 228

Table 9.1: Mean, median and range of percentages of TILs in primary BC and the corresponding recurrent tumours for the subgroup of the study cohort with primary and recurrent tumour blocks (n = 44). 247

Table 9.10: Associations between percentage PD1+ TILs and percentage PDL-1+ TILs and the clinicopathological features of breast cancer 271

Table 9.11: Associations between the percentage tumour cells expressing PDL-1+ in the cytoplasm and the clinicopathological features of breast cancer. 272

Table 9.2: Mean, median and range of percentages of TILs in different molecular subtypes of BC. 247

Table 9.3: Mean, median and range of immune cell subsets in core diameters, assessed as absolute numbers of cells in each subset, for the entire cohort and different molecular subtypes of BC. 249

Table 9.4: Mean, median and range of percentages of PDL-1+ tumour cells in different molecular subtypes of BC. 253

Table 9.5: Mean, median and range of percentages of PDL-1+ TILs and percentages of PD1+ TILs in different molecular subtypes of BC. 253

Table 9.6: Associations between the percentage of average-tumour infiltrating lymphocytes (AV-TILs) and the clinicopathological features of primary BC. 256

Table 9.8: Multivariate Cox regression hazard model including other prognostic clinicopathological parameters of the prognostic value of immunohistochemically defined CD3+ or CD8+ T lymphocyte counts for breast cancer-specific survival in TNBC. 269

Table 9.9: Multivariate Cox regression hazard model including the numbers of various immunohistochemically defined immune cell subsets for the prognostic value of average TILs for breast cancer-specific survival in TNBC.....269

List of Figures

Figure 1.1: Schematic illustration of protein domains in the ERα nuclear receptor……11

Figure 1.2: Schematic representing the two models for the concept of intratumoural heterogeneity. ……24

Figure 2.1: Assembly of a western transfer sandwich. ……54

Figure 2.2: Schematic illustration of the single IHC assay in the NovolinkTM Max Polymer Detection System……60

Figure 2.3: Schematic illustration of the double IHC staining assay…… 64

Figure 3.1: Several pathways regulate the ability of ALDH1A1 to affect cell growth, differentiation and survival. …… 71

Figure 3.2: Western blot to confirm the specificity of the rabbit monoclonal anti-ALDH1A1 antibody.....77

Figure 3.3: Immunohistochemical (IHC) staining for ALDH1A1 in invasive BC……79

Figure 3.4: IHC analysis of ALDH1A1 expression in stromal cells……79

Figure 3.5: Kaplan–Meier survival curves of breast-cancer specific survival for the entire Nottingham cohort, and the entire METABRIC cohort stratified by ALDH1A1……84

Figure 3.6: Kaplan–Meier survival curves of breast-cancer specific survival for patients with different molecular subtypes of BC stratified by ALDH1A1 protein expression……84

Figure 3.7: Kaplan–Meier survival curves of breast-cancer specific survival for patients with different molecular subtypes of BC stratified by ALDH1A1 mRNA expression……85

Figure 4.1: Western blot to confirm the specificity of the rabbit monoclonal anti-BMI-1 antibody……98

Figure 4.2: Representative photomicrographs of BMI-1 protein expression in invasive BC……99

Figure 4.3: Kaplan–Meier survival curves for patients with different molecular subtypes of BC stratified byBMI-1 protein expression.....107

Figure 4.4: Kaplan–Meier survival curves for patients with ER+ of BC stratified by BMI-1 protein expression in ER+.....107

Figure 4.5: Kaplan–Meier survival curves for patients with ER+ of BC stratified by BMI-1 protein expression in TNBC..... 108

Figure 4.6: Kaplan-Meier plots of breast cancer-specific survival (BCSS) for patients with ER+ of BC and TNBC stratified by BMI-1 mRNA expression……108

Figure 4.7: Kaplan-Meier plots of breast cancer-specific survival (BCSS) for the entire METABRIC breast cancer (BC) cohort stratified by BMI-1 copy number……109

Figure 5.1: Western blots to validate the rabbit monoclonal anti-SOX9 antibody and EpCAM antibody. Single. ……121

Figure 5.2: Immunohistochemical double staining of a normal human breast duct for ALDH1A1, SOX9 and EpCAM……124

Figure 5.3: Correlation matrix summarising the correlations between the expressions of all tested breast cancer stem cell (BCSC) markers in the entire study cohort……..128

Figure 5.4: Associations between SOX9 and EpCAM expression and breast cancer-specific survival (BCSS) in the entire cohort and the Oestrogen receptor negative (ER-) subgroup……132

Figure 5.5: Associations between SOX9 and EpCAM expression and breast cancer-specific survival (BCSS) in the Oestrogen receptor positive (ER+) and human epidermal growth factor receptor 2 positive (HER+) subgroups. ……132

Figure 5.6: Associations between SOX9 and EpCAM expression and disease-free interval (DFI) in the entire cohort and Oestrogen receptor negative (ER-) subgroup…….. 133

Figure 5.7: Associations between SOX9 and EpCAM expression and disease-free interval (DFI) in the oestrogen receptor positive (ER+) and human epidermal growth factor receptor 2 positive (HER2+) subgroups…… 133

Figure 5.8: Associations between SOX9 and EpCAM expression and distant metastasis-free survival (DMFS) in the entire cohort and Oestrogen receptor negative (ER-) subgroup……135

Figure 5.9: Associations between SOX9 and EpCAM expression and distant metastasis-free survival (DMFS) in the ER+ and HER2+ subgroups…….. 135

Figure 5.10: Associations between co-expression of ALDH1A1, SOX9 and EpCAM and breast cancer-specific survival (BCSS), disease-free interval (DFI), and distant metastasis-free survival (DMF), respectively, in the entire cohort, and oestrogen receptor positive (ER+) subgroups……135

Figure 5.11: Double immunohistochemical staining of human breast tumours for ALDH1A1, SOX9 and EpCAM. …… 140

Figure 5.12: Associations between the ALDH1A1+SOX9+EpCAM cell phenotype and disease-free interval in the entire cohort, ER-, and HER2+. ……..146

Figure 5.13: Associations between the ALDH1A1+SOX9+EpCAM cell phenotype and disease-free interval in the entire cohort, ER-, and HER2+……147

Figure 5.14: Associations between the ALDH1A1+SOX9+EpCAM+ cell phenotype and distant metastasis-free survival in the entire cohort, ER- tumours, ER+ and HER2+ tumours……..147

Figure 6.1: Shows representative images of ESR1 expression (score 0 to score 4) in breast cancer (BC) tissue sample using RNA-scope……..161

Figure 6.2: Representative examples of nonspecific oestrogen receptor (ER) staining……..165

Figure 6.3: Representative examples of abnormal patterns of oestrogen receptor (ER) staining……..165

Figure 6.4: Representative examples of oestrogen receptor (ER) staining for different antibody clones..... 167

Figure 6.5: Kaplan-Meier survival plots of (A) breast cancer-specific survival (BCSS) and (B) recurrence free survival (DFI) for patients with 1-9% Oestrogen positive (ER+) cells and ≥ 10% Oestrogen positive (ER+) cells..... 173

Figure 7.1: shows the Intratumoral ER heterogeneity in primary BC ……185

Figure 7.2: Heatmap showing the variation in oestrogen receptor (ER) expression at the regional level……..188

Figure 7.3: Percentage of tumours exhibiting each oestrogen receptor positive (ER+) staining intensity at the tumour level. ……188

Figure 7.4: illustrates the number of cases with different patterns of ER distribution in each region……..188

Figure 7.5. Representative images of the distribution patterns in different regions of the same tumour of ER expression in primary BC.....190

Figure 7.6: Shows the high proportion of very weak and weak staining intensities in tumours with the focal and scatter distribution patterns of oestrogen receptor (ER) expression……190

Figure 7.7: Shows the same regions of an (B1) ER+ and (B2) PgR+ tumour……192

Figure 7.8 Primary BC with the different pattern of oestrogen receptor positive (ER+) expression (A1) and the paired oestrogen receptor negative (ER-) LN metastasis…..193

Figure 7.9. Representative examples of primary breast cancer (BC) specimens with high and low intratumor oestrogen receptor (ER) heterogeneity……195

Figure 7.10. Dot plot of intratumor oestrogen receptor (ER) heterogeneity versus percentage of ER+ cells for the 111 primary breast cancer (BC) specimens….. 195

Figure 7.11: Kaplan-Meier survival plots of disease-free interval (DFI) for 111 breast cancer patients stratified by high and low intratumoral oestrogen receptor (ER) heterogeneity……200

Figure 7.12: Kaplan-Meier survival plots of breast cancer specific survival (BCSS) for patients with oestrogen receptor (ER) heterogeneous tumours and patients with ER homogenous tumours who did not receive endocrine treatment and who received endocrine treatment……201

Figure 8.1: Western blot validation of the rabbit anti-SLC39A6 antibody. ……213

Figure 8.2: Representative images of immunohistochemical staining for SLC39A6 in the cytoplasm and nuclei of BC cells.....213

Figure 8.3: Kaplan-Meier survival plots of the associations between high SLC39A6 nuclear and cytoplasmic expression and breast cancer-specific survival (BCSS) in the entire breast cancer cohort..... 223

Figure 8.4: Kaplan-Meier plots of the associations between high SLC39A6 nuclear and cytoplasmic protein expression and breast cancer-specific survival (BCSS) in oestrogen receptor positive breast cancer.....223

Figure 8.5: Kaplan-Meier plots of the associations between high SLC39A6 nuclear and cytoplasmic protein expression and breast cancer-specific survival (BCSS) in oestrogen receptor -negative breast cancer.....224

Figure 8.6: Kaplan-Meier survival plots showing the association between the combination of high SLC39A6 cytoplasmic and high SLC39A6 nuclear protein expression and breast cancer-specific survival (BCSS) in the entire breast cancer cohort and oestrogen receptor positive breast cancer 224

Figure 8.7: Kaplan-Meier survival plots showing the association between high SLC39A6 mRNA expression and breast cancer-specific survival (BCSS) in the entire BC cohort and oestrogen receptor positive breast cancer.....225

Figure 8.8: Kaplan-Meier survival plots showing the association between SLC39A6 copy number loss and breast cancer-specific survival (BCSS) in the entire breast cancer cohort and oestrogen receptor positive breast cancer..... 225

Figure 8.9: (A) Kaplan-Meier survival plots showing the association between high SLC39A6 nuclear expression and longer breast cancer-specific survival (BCSS) in endocrine-therapy naïve patients with ER + BC.....230

Figure 8.10: (A) Kaplan-Meier survival plots showing the association between high SLC39A6 mRNA expression and longer BCSS in endocrine-therapy naïve patients with oestrogen receptor positive tumours..... 230

Figure 8.11: (A) Kaplan-Meier survival plots showing the association between high SLC39A6 nuclear expression and longer distant metastasis-free survival (DMFS) in endocrine-therapy naïve patients with oestrogen receptor positive breast cancer.....231

Figure 9.1: Average percentages of TILs, and hotspot TILs for multiple tumour blocks.....245

Figure 9.2: Representative photomicrographs of TILs infiltrating a primary tumour and the corresponding recurrent tumour from the same patient.....246

Figure 9.3: Representative photomicrographs of the expression of the immune cell markers CD3, CD8, FOXP3, CD20, and CD68 in BC.....250

Figure 9.4. Correlation matrix summarising the correlations between the cell counts for immune markers and percentages of AV-TILs in the entire study cohort……257

Figure 9.5: Kaplan-Meier survival curves of the associations between percentage AV-TILs and breast cancer-specific survival (BCSS) and distant metastasis-free survival (DMFS) in the entire cohort……257

Figure 9.6: Kaplan-Meier survival curves of the association between percentage AV-TILs and breast cancer-specific survival (BCSS) and distant metastasis-free survival (DMFS) in triple negative breast cancer……258

Figure 9.7: Kaplan-Meier survival curve of the association between percentage average (AV)-TILs in recurrent breast cancer and post-recurrence survival……260

Figure 9.8: Kaplan-Meier survival curves of the associations between the numbers of FOXP3+ lymphocytes and CD68+ histiocyte and breast cancer-specific survival (BCSS) in the entire cohort..... 261

Figure 9.9: Kaplan-Meier survival curves of the associations between the number of CD3+, CD8+ lymphocytes and breast cancer-specific survival (BCSS) and distant metastasis-free survival (DMFS)……261

Figure 9.10: Kaplan-Meier survival curves of the associations between the numbers of CD68+ cells and breast cancer-specific survival (BCSS) in the luminal B subtype and the numbers of CD20+ lymphocytes and BCSS in the HER2+ subtype……266

Figure 9.11: Kaplan-Meier survival curves of the associations between the combinations of percentage PD1+ TILs and percentage AV-TILs and breast cancer-specific survival and distant metastasis-free survival in the entire cohort and TNBC……267

Figure 9.12: Kaplan Meier survival curves of the association between combinations of percentage PD1+ TILs and percentage AV-TILs and breast cancer-specific survival in TNBC……267

Chapter 1: Introduction

1.1 Introduction to breast cancer

1.1.1 Histology of the breast

The breasts are basically tubuloalveolar apocrine sweat glands enveloped in fibrous fascia. The breast is composed of two types of epithelial cell (luminal and myoepithelial) and two types of stromal cells (interlobular and intralobular), which form large ducts and terminal duct lobular units. The lactiferous ducts are lined by cuboidal epithelial cells on the luminal surface and have an external myoepithelial surface. These ducts dilate to form the lactiferous sinuses, which become the lobes and lobules.

Breast adipose tissue forms a large part of the breast stroma and determines the size of the breasts [1]. The breast stroma is made up of fibrous connective tissue and adipose tissue in varying proportions. The interlobular stroma contains cellular components such as fibroblasts, myofibroblasts adipocytes, as well as the blood and lymphatic vessels. The intralobular stroma supports the acini of the terminal duct lobular units [2].

Branching morphogenesis, the process by which the breast ducts and acini begin to develop *in utero,* pauses in childhood and recommences in puberty under the influence of hormones. Branching morphogenesis defines the breast architecture and involves complex interactions between breast stromal cells, epithelial cells and the extracellular matrix (ECM) [3, 4].

The human breast is susceptible to carcinogenesis due to the changes that occur during the childhood-to-adulthood transition and the cellular plasticity of

the breast [3]. The plasticity of breast tissue is evidenced by the transition from a stem cell-derived structure with a high proliferative potential into terminally differentiated mature breast tissue that can produce milk [5]. Cancer develops when abnormal breast cell division occurs, followed by aberrant cell growth and proliferation due to variety of risk factors. Cancer cells invade the normal breast tissue and can subsequently metastasise to the lymph nodes and other organs [6].

1.1.2 Epidemiology of breast cancer

Breast cancer (BC) is the most common cancer in women. An estimated 2 million new cases of BC were diagnosed worldwide in 2018, which represents 25% of all types of cancer among females [7]. In the United States, an estimated 252,710 cases of invasive BC were diagnosed in 2017 and 42,170 women have been predicted to die of BC in 2020 (Breastcancer.org). The estimated BC incidence and mortality rates in Europe in 2012 were 464,202 and 131,257, respectively [8]. In the UK, BC is the second cause of cancer-related deaths. Approximately 55,122 new BC cases were reported in 2016 and approximately 11,000 deaths are reported each year. Women in the UK have a 1 in 7 lifetime risk of developing BC (Cancer Research UK, 2020).

The incidence of BC has been predicted to reach roughly 3.2 million cases worldwide per year by 2050 [9]. According to Cancer Research UK (2018), the incidence of BC in the UK is expected to increase to 2% (210 cases per 100,000 females) between 2014 and 2035.

Increased public awareness of BC, ageing populations and widespread availability of mammographic screening have led to increases in the BC survival rates in developed countries over the last decade [10]. Moreover, improvements

in early detection and treatment have reduced the death rates for BC in developed countries such as the UK and North America over the last 25 years [11]. However, the majority of cases (58%) of BC occur in developing countries. [12]. In Saudi Arabia, the number of female BC patients significantly increased from 1,152 in 2008 to 1,473 in 2010 [13].

1.1.3 Factors associated with the risk of BC

The risk of developing BC is influenced by various factors, including age, genetics and family history of the disease, high breast tissue density, hormone-related factors and lifestyle. Some of these common risk factors are discussed below.

1.1.3.1 Age

The majority of BC patients are diagnosed over the age of 50 years. The American Cancer Society (2019/2020) reported that half of women who developed BC were 62 years of age or younger at diagnosis. The median age at diagnosis for black women (60 years) is slightly younger than that of white women (63 years; American Cancer Society, 2019/2020). In the UK, women aged 50 to 69-years-old account for 48% of BC diagnoses (Cancer Research UK, 2018).

In the UK, around 9 in 10 women diagnosed with BC between the ages of 40 to 69-years-old survive for five years or more, compared to only 7 in 10 women diagnosed at an age of 80 or older (Cancer Research UK, 2020). Although the survival rates for younger patients (under 40 years) have doubled, younger patients characteristically have more aggressive tumour features,

including human epidermal growth factor receptor-2 (HER2)-enriched tumours and triple-negative BC [14].

1.1.3.2 Family history and genetic factors

BRCA1 and *BRCA2* gene mutations are the most well-characterised susceptibility genes and are associated with the lifetime risk of BC. BRCA1 and BRCA2 are tumour suppressors involved in DNA repair and transcriptional regulation [15, 16]. *BRCA1* and *BRCA2* gene mutations are present in between 1 in 400 and 1 in 800 women, respectively [17]. Women carrying a *BRCA1* or *BRCA2* mutation have a 45-65% risk of developing BC by the age of 70 [16, 18]. In addition, women with at least one first or second degree relative carrying the *BRCA1* or *BRCA2* gene mutations are twice as likely to develop BC before the age of 35 [19-21].

1.1.3.3 Breast tissue density

Dense breast tissue can be distinguished on a mammogram by examining the relative amounts of fat (dark areas) and fibroglandular tissue (white areas). Women with dense breast tissue have a four- to six-fold higher risk of developing BC compared to women with non-dense tissue [22]. Dense breast tissue is a strong, independent risk factor for developing BC [23]. Dense breast tissue has also been associated with larger tumours and lymph node involvement in BC [22].

1.1.3.4 Hormone-related factors

Prolonged exposure to oestrogen or progesterone due to a variety of factors, including early menarche, late menopause, low parity, nulliparity and taking hormone-replacement therapy (HRT) are associated with increased risk of developing receptor-positive BC, but not receptor-negative BC [24, 25]. The risk of developing BC is increased among women who have taken HRT (oestrogen only) for more than ten years [24]. Moreover, meta-analysis indicate that women with early and more births and longer periods of breastfeeding have lower risks of developing BC compared with nulliparous women and those who never breastfed [26].

1.1.3.5 Lifestyle factors

Several lifestyle factors have been associated with an increased risk of developing BC, including alcohol consumption, obesity and smoking. Alcohol consumption can increase the levels of oestrogen-related hormones in the blood and consequently active oestrogen receptor (ER) signalling pathways. Studies have reported that daily consumption of 35-44 grams of alcohol can raise the risk of BC by 32% [27, 28]. A high intake of fat, especially saturated fats, is also associated with an increased risk of BC in women [29]. Moreover, smoking—especially starting smoking before the age of 17 years—is associated with a higher risk of developing BC [30].

1.2 Prognostic and predictive tools in BC

Prognostic factors can be used to predict the risk of death or recurrence due to BC, independent of the treatment administered, and include tumour size, tumour grade, nodal stage, lymphovascular invasion (LVI), Oestrogen receptor (ER), Progesterone receptor (PgR) status, (human epidermal growth factor receptor 2 (HER2) status, and the Nottingham Prognostic Index (NPI).

Predictive factors can help to identify patients who are likely to respond to a specific therapy, and include ER/PgR status, HER2 status and the NPI.

1.2.1 Tumour size

Patients with larger tumours have shorter survival intervals than patients with smaller tumours [31]. Tumour size is also a very important predictor of nodal metastasis and disease-free survival in BC. Patients with tumours smaller than 3 cm are likely to have involvement of one to three axillary lymph nodes (LNs), while patients with tumours larger than 3 cm are likely to have four or more positive LNs [32]. Moreover, an analysis of 1894 patients found tumour size was a strong predictor of 15-year overall survival in women with both node-positive and node-negative BC; smaller tumour size was associated with significantly better overall survival. A decrease in tumour size of 1.0 cm was associated with 10.3% and 2.5% reductions in 15-year mortality in LN-positive BC and LN-negative BC, respectively [33].

1.2.2 Histologic grade

Histological grade is an independent prognostic factor for BC outcomes [31]. The Nottingham grading system provides a semi-quantitative assessment of three morphological features: tubule formation, nuclear pleomorphism and the number of mitotic cells.

Tubule or acinus formation is scored 1 if present in the majority of the tumour (> 75%); 2, if moderate (10–75%), and 3, if little or none is present (< 10%). Nuclear pleomorphism is scored 1 if cells are small or uniform; 2, if a moderate increase in size and variability is observed, and 3, if marked variation is present. Mitotic count is scored in the hotspot and the field diameter was 0.58 mm on high power field (HPF) (each field size, and the cut-off points are varied based on the HPF of the microscope). The three scores are as next, 1 score if the mitotic count is < 9; 2, if the mitotic count is 10 to 19, and 3, if the mitotic count is > 20. Tumours with total score of 3, 4, or 5 are classified as grade I tumours; tumours with a total score of 6 or 7, as grade II, and tumours with a total score of 8 or 9 as grade III [31, 34].

Grade I includes well-differentiated tumours, which have a good prognosis; grade II reflects moderately differentiated BC, which has a moderate prognosis; and grade III reflects poorly differentiated tumours with the poorest prognosis [31]. Histologic grade is strongly associated with both breast cancer–specific survival (BCSS) and disease-free survival (DFS) in invasive BC and is routinely assessed during tumour diagnosis [34, 35]. Histologic grade has also been demonstrated to be an independent prognostic factor for overall survival, regardless of nodal status or tumour size [36].

1.2.3 Nodal status and tumour stage

The presence of axillary lymph node (LN) metastasis is assessed during tumour staging and is the most important prognostic factor for disease-free survival. Only 20-30% of patients without nodal involvement develop recurrence within 10 years, compared to 70% of patients with axillary node involvement.

LN stage is significantly associated with mortality in BC, which increases with the extent of nodal involvement [37]. The LN staging system contains three distinct categories: 1) stage I, no LN involvement; 2) stage II, one to three LNs involved; and 3) stage III, more than three LNs involved [38, 39].

1.2.4 Nottingham Prognostic Index

The NPI, a well-established prognostication system, is widely used in the management of BC to predict patient outcome and support clinical decision-making [40]. The NPI is comprised of three major prognostic components: tumour grade, tumour size and LN stage. The NPI is calculated as follows [39, 41]:

NPI = [0.2 × tumour size (in centimetres)] + LN stage (1 to 3) + tumour grade (1 to 3)

Patients are classically divided into three groups based on NPI: excellent NPI (NPI < 3.4); moderate NPI (3.4–5.4); and poor NPI (> 5.4) [41]. The NPI can predict survival outcomes: patients with excellent, moderate, and poor NPI have 15-year survival rates of 80%, 42%, and 13%, respectively [40].

Adjuvant therapy is not recommended for patients with excellent NPI scores and ER+ tumours. Premenopausal patients with moderate and poor NPI

scores are candidates for chemotherapy regardless their ER status (patients with ER+ tumour are candidates for chemotherapy, and hormonal therapy), while only hormonal therapy is recommended for postmenopausal patients with ER+ and moderate or poor NPI scores [40].

Furthermore, the use of six subclasses—excellent (NPI, 2.02–2.4), good (2.41–3.4), moderate 1 (3.41–4.4), moderate 2 (4.41–5.4), poor (5.41–6.4) and very poor (6.41–6.8)—has been demonstrated to improve the prognostic power of the NPI [42].

1.2.5 Lymphovascular invasion

LVI refers to the presence of tumour cells within lymphatic vessels, blood vessels, or both, and is a critical stage in the formation of metastases from solid tumours. LVI is independently associated with a poor prognosis in BC [43], and an even poorer outcome in patients with hormone receptor-negative BC [44].

1.2.6 ER/PgR status

Hormone receptor status (HR), ER and PgR, is a fundamental prognostic and predictive factor in BC, and is particularly useful for identifying patients likely to benefit from adjuvant hormonal therapy. HR+ tumours exhibit a better response to tamoxifen than HR- tumours [45, 46].

There is evidence that premenopausal women with HR+ tumours have a longer disease free interval: 48% and 52% of women with HR+ BC treated with hormone therapy and chemotherapy, respectively, responded positively to treatment, compared only 6% of patients with HR- BC [47]. Patients with tumours expressing high levels of the ER have a longer disease-free interval and lower chance of recurrence after mastectomy compared to patients with ER- tumours [48].

1.2.7 Oestrogen receptor in BC

The ER are intracellular receptors that function as transcription factors to mediate the effects of oestrogen [49]. Two ER isoforms (α and β) are found in normal breast tissue and both bind the endogenous oestrogen 17β-estradiol (E2). ERα and β have 56% similarity in their ligand binding domains, but the ERβ isoform is more abundant in mammary tissues; these differences enable selective targeting of each receptor subtype. [50]. ERα is expressed in a subset

of epithelial cells and ERβ is expressed in both epithelial and stromal cells [51, 52]. ER signalling is important in the development of normal breast tissue and abnormal transcriptional activity of the ERα is strongly implicated in the development of ER+ BC [53]. ERα is upregulated in the majority (75%) of BC tumours and is a hallmark of hormone-dependent tumour growth. In contrast, ERβ is frequently downregulated in BC [54, 55].

The ERα regulates a variety of genes involved in cellular proliferation, differentiation and migration, thus aberrant ERα signalling is associated with the development of BC [53]. The risk of developing BC increases with oestrogen exposure and direct and indirect oestrogen blockade are the mainstay of treatment for ER+ BC [56].

Assessment of ERα expression is crucial to determine prognosis and predict the response to endocrine therapy in BC. Patients with high ERα expression are more likely to benefit from endocrine therapy than patients with low levels of ERα [57, 58].

The basic structure of the ER is similar to other members of the steroid hormone receptor family, and consists of six functional domains named A-F. The principal domains are the N-terminal domain (NTD), DNA-binding domain (DBD) and ligand-binding domain (LBD) [59, 60] (Figure 1.1). The transcriptional activity of the ER is regulated by the activation function domains AF1 (located in the NTD) and AF2 (located in the LBD), which are thought to act synergistically to achieve full transcriptional activity [60]. The AF1 (A/B) domain located in the C-terminal region of the ER has the highest variability of the steroid family of receptors and possesses hormone-independent activation function [61, 62]. The D region of the ER, the hinge region, connects the DBD to the LBD and contains nuclear localisation signals that are unmasked when the ligand binds. The C-

terminal F domain modulates the functions of the ER in a ligand-, gene promoter- and cell-specific manner.

In the absence of its ligand, the ER exists as a complex with HSP90 in the nucleus of the target cell. When bound to its ligand, the ER is rapidly dislodged from the complex and dimerises and interacts with the promoters of target genes via either DNA-bound transcription factors or oestrogen response elements to up- or downregulate target gene transcription [53, 63].

E2 binds to the ERα in breast cancer stem cell (BCSC) and the resultant intracellular signalling stimulates cellular proliferation and modifies cellular architecture and phenotype. E2 can also enhance the activity of c-Myc, an ERα transcriptional target, which is associated with tamoxifen resistance [64].

The precise function of the ERβ in BC remains unclear, and ERβ is not currently thought to represent a diagnostic or treatment target in BC [65].

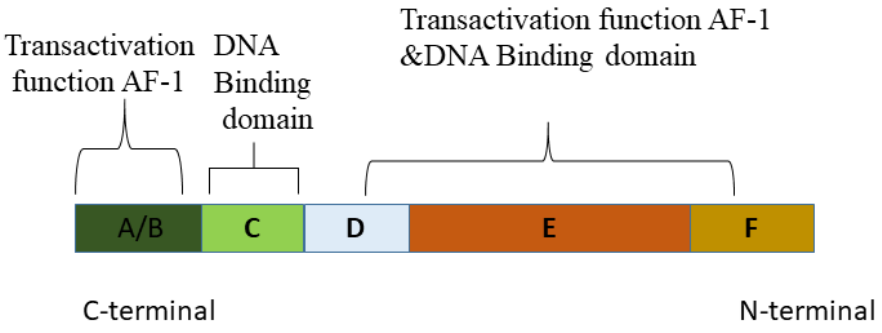


Figure 1.1: Schematic illustration the labelling of the protein domains of ERα nuclear receptor.

1.2.8 HER2/neu

HER2 is member of the epidermal growth factor (EGF) family. HER2 activates several signalling pathways to enhance cellular proliferation, promote the cell cycle, inhibit apoptosis and induce angiogenesis [66-68].

HER2 expression has been associated with expression of the ER and PgR, TNM stage, sentinel lymph node metastasis, blood vessel invasion, pathological type and histological grade in BC. High levels of HER2 expression are associated with sentinel lymph node metastasis and a poorer prognosis in BC [69].

Expression of the HER2 predicts a good treatment response in patients with HER2+ status receiving the HER2-targeted monoclonal antibodies including Herceptin (Trastuzumab). Studies have reported that HER2-positivity may also predict resistance to endocrine therapy in HR+ BC [70].

1.3 Classification of BC

1.3.1 Molecular classifications

In the last two decades, microarray-based gene profiling studies have led to the identification of several subtypes of BC. These subtypes are Luminal A, Luminal B, HER2-enriched, basal-like, and normal-like BC [71].

The Luminal A subtype has good prognosis, low rates of relapse and is the most frequently encountered type (about 60%). The Luminal A subtype is both ER+ and/or PgR+, but HER2-negative, and is characterised by low mitotic activity (low Ki67 index), a low degree of nuclear pleomorphism and low histologic grade [72]. The Luminal A subtype shows an excellent response to endocrine therapy compared to the Luminal B type [72].

Luminal B tumours account for 10-20% of all cases of BC, and are both ER+ and and/or PgR+, HER2 negative or positive, and characterised by high mitotic activity (high Ki67 index) and high histologic grade. Luminal B tumours do not respond well to hormonal therapy targeting the ER+, but respond better to neoadjuvant chemotherapy than the Luminal A subtype [61].

The HER2-enriched subtype is characterised by high expression of the HER2 gene and/or protein; 15-20% of all cases of BC are HER2-positive and about half of these tumours express the ER (though at low levels). HER2-enriched tumours are more aggressive and highly proliferative and metastatic [72].

The basal-like subtype expresses high levels of the basal myoepithelial markers cytokeratin (CK) 14 and CK 17, exhibits high levels of Ki67 expression and lacks ER, PgR and HER2 [73]. Triple-negative breast cancer (TNBC) is defined by loss of expression of ER, PR and HER2 [74]. Approximately 70-80%

of cases of TNBC are basal-like and about 70% of basal-like tumours are TNBC [75]. Around 10–15% of all BC cases are basal-like and triple negative [76]. Basal-like BC and TNBC have a poor prognosis and higher death rates compared to other subtypes of BC [77].

The normal breast-like cancer subtypes represent 5-10% of cases of BC. Normal breast-like tumours are poorly defined and have a moderate prognosis compared to the luminal and basal-like tumours [72].

Furthermore, Green and colleagues classified BC into seven subtypes using immunohistochemistry [73]. In addition to the Luminal A, Luminal B, HER2-enriched and TN subtypes, they added three more subclasses: luminal N, basal-p53-altered and basal-p53 normal [73]. The Luminal N class is defined by high ER and PgR expression and negative expression of the HER family, and has an excellent prognosis. Basal tumours are ER-, PgR- and HER2-, express high levels of basal cytokeratins and either expressed high p53 (basal–p53 altered) or low p53 (basal–p53 normal) and have a poor prognosis [73].

1.3.2 Histological subtypes

Invasive BC (IBC) can also be categorised into histological subtypes, including invasive ductal carcinoma (IDC), invasive lobular carcinoma (ILC), and tubular carcinoma.

IDC is the most common type and accounts for 70-80% of all IBC cases, while ILC is the second most common type, accounting for 10% of cases. Patients with ILC have better 5-year recurrence-free survival and overall survival than patients with IDC [78-80]. Tubular carcinoma has the best prognosis. Medullary carcinoma has an excellent prognosis and better outcome than high grade IDC [81].

Invasive ductal carcinoma of no special type (NST), is a common subtype of IDC, and has no specific defining features and poor prognosis due to the heterogeneity of this subtype [81].

Special types are tumours in which special types constitute > 90% of the tumour. Special types include infiltrating lobular carcinoma, invasive cribriform carcinoma, medullary-like carcinoma, mucinous carcinoma, infiltrating papillary and micropapillary carcinoma, and have a good prognosis compared to NST [81].

1.4 The tumour microenvironment in IBC

BC is an extraordinarily heterogonous disease that exhibits intertumour and intratumour genetic and phenotypic diversity, in terms of both tumour cells and the tumour microenvironment [82].

The BC microenvironment contains numerous cell types, such as immune cells, fibroblasts, adipocytes and vascular/lymphatic endothelial cells. The extracellular matrix (ECM) is now being increasingly recognised as a major player in disease progression, and there is considerable interest in developing treatments that may target the influence of the ECM on tumour invasion and metastatic potential [83]. Crosstalk between cells and the tumour microenvironment, as well as between tumour cells, adds more complexity to BC heterogeneity.

Cancer-associated fibroblasts (CAFs), also called peri-tumoral fibroblasts, reactive fibroblasts or myofibroblasts, are major cellular components of the BC microenvironment [84]. Cancer-associated fibroblasts secrete and help to maintain the ECM, which can either promote or suppress tumour growth in addition to supporting the cancer cells [85, 86]. CAFs also play key roles in BC development. Unlike in wound healing where fibroblasts are transiently activated, CAFs are persistently activated [87]. CAFs can promote BC cell motility and invasion by secreting the chemokine CXCL12 into the tumour microenvironment [85].

Tchou *et al.* (2012) showed that CAFs from several BC subtypes exhibit subtype-specific gene expression profiles. The gene expression profile of CAFs derived from the HER2+ subtype was significantly different to that of CAFs derived from ER+ BC or TNBC. The actin cytoskeleton, focal adhesion kinase

(FAK), Rac, Rho and mitogen-activated protein kinase (MAPK) signalling pathways are activated in CAFs isolated from HER2+ tumours. Thus, the characteristic differences in CAFs derived from HER2+ tumours may contribute to the aggressiveness of this subtype of BC, which has a high risk of local and distant recurrence [85].

Natrajan *et al.* (2016) measured microenvironmental heterogeneity along three spatial dimensions using an ecosystem diversity index (EDI) based on automated histology image analysis. Grade III tumours with a highly heterogeneous microenvironment had a significantly poorer outcome compared other grade III tumours. Heterogeneous microenvironments exhibited high levels of specific genomic alterations, indicating the clinical implications of the synergistic interaction between cancer cells and the surrounding microenvironment. Moreover, their work suggested heterogeneity in the tumour microenvironment of high-grade tumours, but not low-grade tumours, may promote tumour progression in BC [88].

1.5 Role of immune cells in BC

The immune system exerts two main roles in cancer: immune cells both promote host defences against tumour cells and facilitate the escape of tumour cells from immunosurveillance and destruction [89, 90]. The concept of immune-editing was devised to explain how cellular immunity promotes tumour growth and could also help to eradicate tumours [91, 92]. Tumour cells and immune cells interact via three cancer immunoediting phases: elimination, equilibrium and escape.

In the elimination phase (immunosurveillance), both the innate and adaptive immune systems respond to and eliminate tumour cells before a tumour manifests clinically. In the equilibrium phase, a balance exists between anti-tumour and tumour-promoting factors, so that the tumour is maintained in a state of functional dormancy. In the escape phase, the tumour cells finally overwhelm the host defences. The mechanisms of immune escape include reduced immune recognition via decreased expression of major histocompatibility complex class I (MHC I) and/or co-stimulatory molecules and increased expression of immunosuppressive factors that reduce the ability of CD8+ cytotoxic T lymphocytes to lyse tumour cells [93].

Immune cells such as T cells, B cells, natural killer (NK) cells and macrophages, in addition to other cells such as neutrophils and dendritic cells, exert different functions and can be identified in BC [94]. Importantly, the statement that the immune system promotes or prevents cancer cannot be interpreted without specifically defining which immune cell phenotypes participate in each process.

T lymphocytes account for the majority of TILs and express a large variety of activation antigens that activate host defence mechanisms [95]. The different subsets of T lymphocytes perform varied functions; for example, CD8+ T cells, also known as cytotoxic T cells, destroy tumour cells by binding to antigens presented by MHC class 1 molecules expressed on the membranes of all nucleated cells [96]. On the contrary, T regulatory cells play a vital role in immune tolerance of tumour cells via various mechanisms, which include inhibition of effector cytotoxic T lymphocytes [97]. Helper T cells, which are CD4+ T cells, are key regulators of immune system differentiation into various T helper (Th) cell subsets with distinct biological functions [98]. T helper 1 (Th1) cells and natural killer (NK) cells can protect against tumour development and progression. On the other hand, Th2 cells promote tumour development and progression [99].

In contrast to T lymphocytes, the roles of B lymphocytes are not well characterised. Resting B cells have been reported to inhibit T cells, while activated B cells facilitate T cell activity [99].

1.5.1 Prognostic significance of TIL subtypes in IBC

Different molecular subtypes of BC have distinct clinical outcomes and responses to therapy. These varied clinical behaviours could potentially be attributed, at least in part, to the variations in the density, pattern, location and type of TILs in different BC subtypes.

For instance, cytotoxic T cells (CD8+ T cells) and T regulatory cells (FOXP3+ T cells) have been associated with prognosis. A retrospective study of 3,992 patients with BC found that total intratumoural and stromal CD8+ TILs

were independently associated with better survival [100]. Another study of 1,953 cases treated at University Hospital Basel in Switzerland showed that the good independent prognostic value of total, intraepithelial, and stromal CD8+ TILs was only observed in ER- BC, whereas only intraepithelial lymphocyte (CD8+ TILs) were associated with reduced disease-specific survival in ER+ BC [101].

In contrast, univariate analysis by Lui *et al*. (2012) found that the intratumoral and stromal patterns of CD8+ TILs had no prognostic effect in ER+ BC [102]. Interestingly, in the context of intratumoral heterogeneity in TILs, the reduction in mortality associated with the presence of CD8+ T cells in ER- BC (from 57% and 21%) was reported to depend on whether the TILs were intratumoural or stromal [103].

Tumour infiltration of FOXP3+ regulatory T cells is associated with survival outcomes in some subtypes of IBC. FOXP3+ TILs are consistently detected in ER- tumours and associated with significant anti-tumour immunity and a good prognosis. On the contrary, T regs may not significantly suppress cytotoxic T cells in ER- BC [104, 105]. However, the opposite was true for ER+ BC lacking cytotoxic T cell infiltrates, as FOXP3+ TILs were associated with a poor prognosis [106]. However, FOXP3+ TILs are an independent indicator of better survival in HER2+ tumours with CD8 T cell infiltrates [107].

TNBC, one of the most aggressive subtypes of BC, exhibits a higher than average population of FOXP3+ T cells, which are associated with improved survival in TNBC [107]. However, several other studies showed FOXP3+ T cells are associated with a poor prognosis in BC [108]. Also, the relative density and tissue location of tumour-infiltrating FOXP3+ T regs have been associated with

varied prognoses in BC [109]. Based on these observations, the heterogeneity of BC at the molecular level could potentially affect the expression of the FOXP3+ TILs.

Neoadjuvant chemotherapy (NCT) is now being increasingly used as a therapeutic strategy for early stage BC. High lymphocytic infiltration is an indication of a better prognosis and better response to NCT in BC [110], although this association is not observed in all subtypes of BC [110]. In HER2+ BC, a pathological complete response (pCR) to neoadjuvant chemotherapy can be predicted based on the percentages of stromal TILs [111]. High levels of TILs were associated with higher pCR rates during anthracycline and taxane-based NCT in TNBC [112]. However, due to the absence of data from patients with TNBC who did not receive chemotherapy, TILs require further evaluation as predictive markers for chemotherapeutic response [106]. Notably, TILs were not prognostic in patients with ER+ BC receiving adjuvant chemotherapy [113]. CD8+ cytotoxic T lymphocytes (CTLs) have been associated with a favourable response to anthracycline-based chemotherapy in BC, but there is no conclusive evidence that CTLs are an independent predictor of a pCR [114].

Despite the fact that tumour immunity has a potentially protective role, immune cells (especially CD8+ CTLs) can induce tumour cells to undergo the epithelial-mesenchymal transition (EMT) and confer BCSC-like properties and aggressive features [115]. BCSCs have been shown to play an important role in chemoresistance [116]. Moreover, the EMT, an important mechanism associated with cancer cell invasion and metastasis, has been associated with chemoresistance [117, 118].

Sikandar *et al.* (2015) compared the density of inflammatory cells (both lymphocytes and monocytes) during different stages of tumorigenesis. In line with other findings [119], a higher number of TILs was observed in DCIS compared to invasive BC and normal breast tissues. Thus, TILs have been suggested to play a role in the progression of early-stage BC.

The significant heterogeneity of TILs in different histological subtypes of BC leads to significant differences in tumour behaviour and clinical outcome. The density of inflammatory cells varies between different histological types of IBC. For instance, Rakha *et al.* (2009) reported a difference in the density of inflammatory cells between the medullary histological type and no special type (NST). Medullary carcinoma (typical and atypical) and high-grade ductal carcinoma contain the highest density of inflammatory cells, and these subtypes have a better prognosis than tumours with lower densities of inflammatory cells. Interestingly, there was no difference in the prognosis of medullary carcinoma and other grade III carcinomas with high densities of inflammatory cells, however; medullary carcinoma with a low density of inflammatory cells had better prognosis than grade III ductal carcinoma with a low density of inflammatory cells [120]. Mucinous and lobular carcinoma are not highly enriched with inflammatory cells [120]. Moreover, the inflammatory cellular infiltrates vary between lobular and ductal carcinoma, which may indicate significant functional differences between the inflammatory cells in these subtypes [121]. T cells and macrophages have been reported to be the dominant inflammatory cell types in BC. Importantly, the heterogeneity of inflammatory cellular infiltrate may be affected by the grade of BC [122]. In view of the fact that lymphocytes constitute the largest inflammatory cell

population, lymphocytes have been suggested to be responsible for the stromal heterogeneity observed in different subtypes of BC.

1.6 Intratumour heterogeneity in IBC

BC exhibits a remarkable degree of heterogeneity in terms of morphological characteristics, molecular profiles and gene expression between different patients (intertumour heterogeneity) and within individual tumours (intratumour heterogeneity) [123, 124].

Dick *et al.* (2008) described how cancer is a heterogenous disease and BC is no exception. Histopathologists have long been aware that different regions of a tumour can exhibit morphological differences [125], which is why several samples are typically obtained for pathological examination [126].

Aleskandarany *et al.* (2018) reviewed how phenotypic and genotypic variation in BC may be under-characterised in exceptional situations, as a single biopsy sample may not be adequate for assessment of tumour type, grade, lymphovascular invasion, molecular profiles and stroma-related markers and lead to different results than surgical excision specimens, which indicates that single biopsy samples may not accurately reflect intratumour heterogeneity [127].

Though the implications of BC heterogeneity were previously poorly understood, BC heterogeneity has now been recognised as an important determinant of patient outcome [128].

The variety of cell populations within a tumour mass, along with other supportive cells such as endothelial cells, fibroblasts and tumour-infiltrating cells and the highly complex extracellular matrix (ECM) naturally confer a degree of

intratumoural heterogeneity, which is further complicated by the influence of the tumour microenvironment and clonal variation [129, 130]. Indeed, distinct intratumoral molecular and phenotypic heterogeneity may be responsible for the variations in the behaviour and features of primary tumours and the matched metastases [130-132].

Two hypotheses have been put forward and widely studied in an attempt to explain the concept of intratumoural heterogeneity: the cancer stem cell (CSC) hypothesis and the clonal evolution/selection model (Figure 1.2). Both hypotheses suggest that molecular alterations in a single cell enable cells to acquire the ability to proliferate, and also indicate that the influence of the tumour microenvironment affects the composition of BC to a degree.

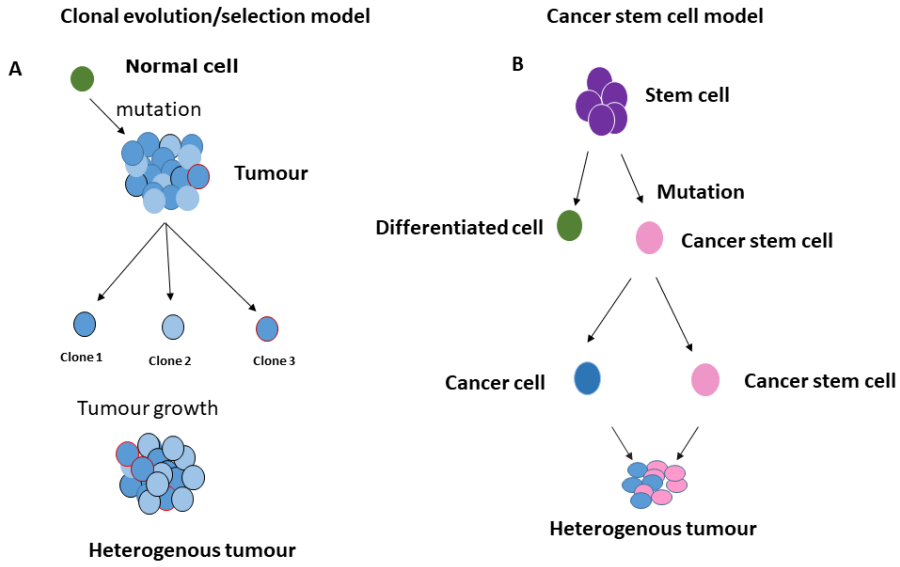


Figure 1.2: Schematic representing the two models for the concept of intratumoural heterogeneity. (A) Clonal evolution/selection model, in which random mutations occur in differentiated cells, which subsequently generate cancerous clonal populations that develop further mutations. (B) Cancer stem cell model, in which tumours are proposed to arise from stem cells that generate hierarchically organised tumours [133].

1.7 Cancer stem cell hypothesis

The cancer stem cell hypothesis proposes that tumours contain subpopulations of cells with similar properties as normal adult stem cells. These cells have the capacity to divide indefinitely and give rise to cancer stem cells (CSCs) and progeny that differentiate into the variety of cell types within a tumour mass [134, 135].

CSCs possess the ability to self-renew and to differentiate in response to certain stimuli, and exhibit high telomerase activity, activation of anti-apoptotic pathways and can metastasise to distant organs [136, 137].

The origin of CSCs is still under debate: some researchers have proposed that normal stem cells with long-lived features acquire genetic alterations and become CSCs, while other researchers state that genetic mutations may enable cells to obtain stem cell characteristics [138]. The central hypothesis of the cancer stem cell lineage hierarchy described by Dick *et al.* (2001) proposes that CSCs arise from normal stem cells and acquire mutations that make the stem cells cancerous [139].

However, this model was questioned by Gupta and colleagues (2011), who demonstrated de-differentiation of cancer cells back into CSC could occur under certain circumstances and that a dynamic phenotypic equilibrium within cancer cell populations is maintained by stochastic state transitions. Their study relied on the quantitative Markov model of cell-state interconversion, which postulates that the probability of a cell transitioning into another cell type depends on its current state, and not its prior state [140].

CSCs were first identified by Dick and colleagues in human acute myeloid leukaemia (AML), which confirmed the existence of a cellular hierarchy in AML, similar to the hierarchy observed during normal haematopoiesis [141]. Cell surface markers for hematopoietic stem and progenitor cells were discovered and used to isolate similar cell populations in AML [142]. Dick and colleagues isolated CD34+/CD38- human CSCS using FACS [143]. This CSC subpopulation had the ability to generate clones *in vitro* and formed tumours when transplanted into immunodeficient mice [144].

Moreover, CSCs can generate spontaneously from neoplastic non-stem cells as a result of epigenetic and genetic changes, which may also contribute to tumour heterogeneity [145]. CSCs express multiple genes related to the epithelial-to-mesenchymal transition (EMT); therefore CSCs have the capacity to migrate into distant sites [146].

The EMT, the transformation of epithelial cells into mesenchymal cells which can migrate through the ECM, is critically important to embryogenesis and normal developmental processes [146]. Several mechanisms are associated with both the EMT and CSCs, such as Twist, ZEB1/2, HIF, TGF-β, WNT/β-Catenin, Notch and Hedgehog signalling pathways.

1.7.1 Cancer stem cell phenotype in BC

BCSCs were first identified by Al-Hajj and colleagues utilising cell surface markers ($CD44^{-}/CD24^{+}$) and can form tumours when only a few hundred cells are injected into Nonobese diabetic/severe combined immunodeficiency (NOD/SCID) mice. In contrast, $CD44^{-}/CD24^{+}$ BC cells transplanted to NOD/SCID mice failed to form tumours [147]. Following this work, other researchers identified CSCs in various

solid cancers, including colon, pancreatic, brain, head and neck cancer [148]. Ginestier *et al.* (2004) showed that normal and malignant breast stem cells with increased aldehyde dehydrogenase activity (ALDH), as assessed by the ALDEFLUOR assay, possess stem/progenitor characteristics and tumour-initiating potential. Furthermore, ALDH+/CD44+/CD24- cells had the greatest tumour-forming potential in NOD/SCID mice, with the ability to form tumours after injection of only 20 cells [149].

BC cell lines express varied levels of BCSC markers. The MDA-MB-231 (TNBC) cell line is comprised of 85% $CD44^{-}/CD24^{+}$ cells, but only 0.9% of MDA-MB-231 cells are $ALDH1^{+}$. Only 0.2% of MCF7 (luminal A) cells express ALDH1, which further indicates the heterogeneity of BCSC marker expression in BC subtypes.

While numerous researchers have successfully isolated BCSCs from either BC patient tissues or BC cell lines using a variety of markers, no universal BCSC markers have yet been identified.

1.7.2 Signalling pathways that regulate BCSCs

Several signalling pathways are activated in BCSCs to maintain stem cell properties and promote tumour progression, including the Notch, Hedgehog, WNT/β-Catenin, JAK/STAT, and NF-κB signalling pathways [150]. Some of the fundamental pathways active in BCSCs are summarised below.

1.7.2.1 Hedgehog signalling

Hedgehog signalling plays a role in the control of embryonic development, tissue regeneration and stem cell renewal [151]. The Hedgehog family is comprised of Sonic (SHH), Indian (IHH) and Desert (DHH) Hedgehog, which bind to the Patched 1 and/or Patched 2 (PTCH1 and PTCH2) receptors and subsequently deactivate/inhibit Smoothened (SMO). Downregulation of SMO activates the glioma-associated oncogene (GLI) transcription factors (GLI1, GLI2 and GLI3), which promote transcription of downstream target genes. Moreover, GLI2 and GLI3 act as transcriptional activators to induce GLI1 and act as transcriptional activators [152].

Overexpression of SHH increased the proliferation, colony formation, and migratory and invasive ability of TNBC cell lines *in vitro*. TNBC contains a high proportion of GLI1-positive basal-like progenitors, which suggests TNBC may respond to ligand-dependent stimulation of canonical Hh pathways [153, 154]. Overexpression of SHH has been associated with a poor prognosis and shorter overall survival in BC [154-157].

Jeng *et al*. (2013) showed that the SHH pathway is abnormally activated in CD133+ CSCs [158]. Moreover, studies have indicated that Hedgehog signalling regulates the expression of CSC markers, such as ALDH1, BMI-1, Snail, Wnt2, PDGFRα, Jagged-1 and CD44 [159-163].

1.7.2.2 Wnt signalling

The highly conserved Wnt signalling pathway consists of 19 Wnt ligands and 10 frizzled (FZD) receptors. The Wnt pathway comprises two signalling pathways. The canonical Wnt pathway is activated by the transcriptional regulator β-

catenin, the downstream transcriptional co-activator of TCF/LEF target gene expression [164, 165]. The noncanonical β-catenin-independent Wnt pathway is divided into the planar cell polarity pathway and the Wnt/Ca^{2+} pathway [166, 167].

Both the canonical and noncanonical Wnt signalling pathways play important roles in embryonic development and homeostasis in several human tissues. The canonical Wnt pathway controls proliferation, survival and cell death, and the noncanonical Wnt pathway regulates asymmetrical cell division, cell polarity and cell migration.

Mingli *et al.* (2015) reported that combined expression of ALDH1A1 and β-catenin, a Wnt pathway gene, is associated with lymph node metastasis and poorer survival in BC [168]. Another study reported overexpression of TCF4, LEF1 and β-catenin in a subset of aldefluor-positive BCSCs [169]. Moreover, induction of oncogenic Wnt/β-catenin genes is associated with metastasis and a poorer prognosis in TNBC [170-172].

1.7.2.3 Notch signalling

Notch signalling is a fundamental pathway in BCSCs. Notch signalling is comprised of four heterodimeric transmembrane Notch receptors (1-4) and the transmembrane ligands Delta-like (Dll) 1, 2, 4 and Jagged (JAG) 1 and 2.

Notch signalling is a recognised mechanism of crosstalk between neighbouring cells, with one cell expressing a ligand and the other cell expressing a receptor that binds the ligand [173]. The interaction between the receptors on different cells induces a modification within the receptors, leading to cleavage of the extracellular domain from the intracellular domain, and

translocation of the intracellular domain to the nucleus to promote transcription of Notch target genes [173].

Notch 1 is thought to be involved in the self-renewal ability of BCSCs by increasing *ErbB2* transcription [174, 175]. ALDH1A1 is activated by NOTCH signalling, as SIRT2 deacetylates and activates ALDH1A1 [176]. A meta–analysis concluded that overexpression of Notch 1 is associated with the progression from ductal carcinoma *in situ* to invasive cancer and poorer recurrence-free survival [177].

1.7.3 BCSCs and metastasis

BCSCs play a role in tumour initiation, heterogeneity, progression, and metastasis and drug resistance and are thus associated with poorer patient outcomes [178-180]. Tumour cells migrate through the bloodstream or lymphatic system to sites of metastasis [181, 182]. CSCs have been suggested to be able to proliferate sustainably, possess EMT features as such as loss of polarity and cell-to-cell adhesion, and have the ability to migrate and promote cancer metastasis, thus CSCs are implicated in tumour progression [183-185].

For instance, the BCSCs marker CD44 is a transmembrane glycoprotein that plays a major role as adhesion molecule. Co-expression of CD44 and other tyrosine kinase receptors promote the proliferation, migration and invasion of CSCs [186-188]. Sheridan *et al.* (2006) reported that the $CD44^{-}/CD24^{+}$ cell subpopulation in BC express high levels of invasion-related genes [189]. Expression of markers such as $CD44^{-}$ / and or $CD24^{+}$ by tumour cells alone does not necessarily mean these cells can create metastases. Identification of CSCs within metastatic tumour or circulating tumour cell populations does not prove that cells are capable of establishing metastatic tumour [190].

Overexpression of ALDH1 is also associated with metastasis in BC [149, 191, 192]. Croker *et al.* (2017) reported that BC cell lines with high ALDH1A1 expression have increased capacity to proliferate, adhere and migrate; however both ALDH1A1 and ALDH1A3 expression correlated with metastasis *in vivo* [193]. Li *et al.* (2017) reported cells with a high CD44/CD24 ratio were mainly reasonable for self-renewal, proliferation and tumour growth, whereas $ALDH1^{+}$ cells possess invasion and metastatic ability. The CD44/CD24 ratio and ALDH1 expression are similar in primary tumours and the matched distant metastases,

as well as circulating tumour cells, indicating these stem markers are sustained during the progression and metastasis of BC [194].

One study reported that high JAG1 protein expression, a Notch ligand, was significantly associated with metastasis to bone in IBC [116].

Current conventional therapeutic strategies for BC target highly proliferating tumour cells. These strategies do not always succeed, and tumour metastasis and recurrence occur frequently. The CSC hypothesis suggests that antitumor therapy fails to target CSCs. Accumulating data indicates that BCSCs are resistant to both chemotherapy and ionizing radiation [195-198]. Lix *et al.* (2008) reported that BCSCs ($CD44^{-}/CD24^{+}$) are resistant to doxorubicin/cyclophosphamide, and the numbers of these cells significantly increased after neoadjuvant chemotherapy in BC [199]. A clinical study showed that high ALDH1 expression was associated with a poor outcome, a low pathological complete response (pCR) rate and resistance to therapy [116].

1.8 Clonal evolution theory and clinical implications

The clonal evolution model agrees that cancer cells originate from a single cell, and that additional mutations which occur during tumorigenesis cause genetic instability in cancer cells [200]. As a tumour is exposed to selective pressures, multiple subpopulations of cells arise due to new mutations that can be disadvantageous or result in proliferation and survival through clonal expansion [201]. The tumour that eventually arises is the result of all of the clonal expansions that occurred due to selective pressure. Clonal expansion may also occur due to epigenetic influences, such as histone modification, DNA methylation, nucleosome positioning and microRNA expression, that confer phenotypic advantages on a clone [202]. The mutations that allow a clone to be selected are known as driver mutations, and in many cases are accompanied by other mutations (passenger alterations) that are not associated with a known selective survival/proliferative advantage or manifestation, but may be converted to a driver aberration if the selective pressures change [203].

Hormone receptor-positive tumours (ER+ and PgR+) account for the majority of BC cases and are treated with endocrine therapy [204]. Though endocrine therapies reduce mortality in ER+ BC, resistance-driven recurrence to hormone therapy develops in about 25-30% of patients [205]. The randomised phase III Paloma-3 trial sequenced circulating tumour DNA from patients at baseline and at the end of treatment with palbociclib plus fulvestrant versus placebo plus fulvestrant. The study demonstrated that clonal evolution is a common occurrence during therapy; clonal evolution was frequently observed in the subset of patients who previously received endocrine treatment [206].

1.9 Genomic instability and heterogeneity

In an attempt to better understand and explain cancer diversity and complexity, Hanahan and Weinberg (2011) proposed eight "hallmark biological capabilities" that normal cells must acquire via a multistep process to develop into cancer. These hallmarks are the ability to resist cell death, sustain proliferative signalling, induce angiogenesis, evade growth suppressors, enable replicative immortality, activate invasion and metastasis, evade the immune system and reprogram cellular energy metabolism. Genomic instability is critical to acquisition of these features [207].

Genomic instability refers to the propensity of the genome to gain mutations, and may occur due to defects in the DNA mismatch machinery that allow point mutations to persist (microsatellite instability) or chromosomal instability that may lead to gain or loss of entire chromosomes or aberrations in the chromosomal structure, such as inversions, duplications, deletions or translocations [208]. Genomic instability plays a key role in breast carcinogenesis and promotes heterogeneity in BC [209].

TNBC is the most heterogenous subtype of BC. TNBC exhibits increased chromosome instability (CIN) compared to the ER+ subtype. Increased CIN70 expression is associated with a good prognosis in TNBC, but a poorer prognosis in ER+ BC [208].

Raynaud *et al*. (2018) collected molecular data for different tumour types from a variety of public databases, including the numbers of mutations and the frequency of copy number alterations, including copy number gains and losses. Their study showed intratumour heterogeneity was greatest in tumours with

large numbers of mutations and copy number alterations. Extremely unstable tumours with DNA repair deficiencies or chromosomal instability were not the most intrinsically heterogeneous. Conversely, intratumour heterogeneity was greatest in tumours exhibiting relatively high copy number alterations, and number of mutations [210].

1.10 Spatial and temporal heterogeneity in BC

The concept that different regions of a tumour may contain different genetic and molecular aberrations is known as spatial heterogeneity and indicates the so-called genetic variations are causal rather than coincidental [211].

In a proof-of principle study, Geyer and colleagues (2010) determined whether the divergent phenotypic features observed within a tumour may be the result of varied genetic aberrations (genetic heterogeneity). Their team demonstrated the presence of intratumour genetic heterogeneity using microarray-based comparative genomic hybridization (CGH), *TP53* sequencing, chromogenic *in situ* hybridisation (CISH) and fluorescence *in situ* hybridisation (FISH) to analyse morphologically distinct areas of metaplastic BC. The morphologically distinct areas had similar genetic profiles, which could be explained by epigenetic changes induced by interactions with the tumour microenvironment. However, two cases contained morphologically distinct components with distinct gene copy number aberrations [212].

Spatial intratumour genetic heterogeneity has also been documented in histologically similar areas of tumours. For example, Geyer *et al.* (2009) showed the lobular component of a tumour was ER+, while the squamous component of the same tumour was triple-negative with *EGFR* gene amplification. These results suggest lobular BC, but not squamous tumours, can respond to endocrine therapy [211].

Navin and colleagues (2011) conducted single-nucleus sequencing to investigate tumour population structure and evolution in two cases of BC, and classified BC as monogenomic or polygenomic based on genomic structure variations. Monogenomic tumour cell populations have identical genomic profiles

and are the result of single clonal expansion in primary and liver metastasis, whereas polygenomic tumours contain subpopulations of "monogenomic" cell clusters which could either be geographically segregated or intermixed, but are genetically related to the initial tumour [212].

At the morphological level, 20% of BC cases are heterogeneous "mixed" histotypes with characteristic special-type areas combined with areas with no special features. Tubular or invasive cribriform carcinomas as pure forms are associated with a more favourable prognosis than the other subtypes, and endocrine therapy is the primary treatment [213]. However, when tubular or invasive cribriform carcinomas also contain a component of another histotype, the prognosis may change and the therapeutic approach should be adjusted to the most advanced biological features [214].

Immunohistochemical staining confirmed that spatial heterogeneity in expression of the hormone receptors (ER and PgR) and HER2 exists in BC, and that this heterogeneity has impact on patient outcome [215]. ER+ tumours are typically defined as the presence of ER nuclear staining of any intensity in ≥ 1% of tumour cells. However, tumours with high ER expression (100% cells) should not be treated equally as tumours with low proportions of ER expressing cells (1-9%), as endocrine therapy only targets ER+ cells [216].

Ahan *et al.* (2020) also showed high intratumour HER2 heterogeneity exists in BC, especially in terms of HER2 protein expression and low-grade *HER2* gene amplification [217]. Patients with heterogeneous intratumoral HER2 gene amplification have shorter disease-free survival than patients with homogeneous HER2 gene amplification [218].

Temporal heterogeneity describes how tumour metastases differ to the primary tumour over time. Kuukasjärvi *et al.* (1997) analysed 29 primary BC tumours and the paired metastases using comparative genomic hybridization and fluorescence *in situ* hybridisation. Gain of the Xq12-q22 regions was more common in the metastases, and the genetic composition of nine metastases (31%) differed almost completely from the paired primary tumours, which indicates how temporal changes represent another type of heterogeneity. This heterogeneity is thought to be the result of continuous evolution (by acquisition of additional genetic variations) of an early stem line clone during tumour progression [219].

For example, *ESR1* mutations can occur in metastatic lesions due to acquisition of resistance to endocrine therapy and disease progression. Schiavon *et al.* (2015) detected *ESR1* mutations in the plasma of patients with ER+ metastatic BC, which suggests that the presence of *ESR1*-mutated ctDNA is a predictive marker of resistance to endocrine therapy. The authors suggested that rare *ESR1* mutant subclones exist due to genetic intratumoral heterogeneity and clonal diversity and are selected during treatment for metastatic BC [220].

HER2 is considered to be a driver genetic aberration in BC. In addition, HER2 status may vary between primary tumours, metastases and circulating tumour cells [221]. A landmark case reported by Wu *et al.* (2010) provided a good example of intratumour heterogeneity. A young female with invasive breast carcinoma underwent core biopsy, which demonstrated *HER2* amplification. The patient eventually relapsed and passed away while receiving trastuzumab as an adjuvant. Examination of the metastases revealed no *HER2* amplification, so the initial tumour sample was re-examined. Only a focal area

corresponding to the site of the core biopsy exhibited *HER2* amplification [222]. This indicates that HER2 status can vary within a primary tumour and between the primary and metastatic lesions, which could be explained by potential selection of unamplified portions of heterogeneously amplified tumours by Trastuzumab [222]. Another study suggested that metastases may arise from a single cell or a limited, small number of primary clones from the primary tumour, or even originate from clones present in other metastases, rather than from the clones present in the primary tumour [223].

Discordance in the expression of key prognostic markers has also been reported between primary and metastatic BC, with discordant expression of ER, PgR, and HER2 observed in 20%, 33%, and 8% of cases. Receptor-positive primary tumours frequently convert to receptor-negative metastases; receptor-negative primary tumours rarely convert to receptor-positive metastases [224]. Therefore, therapies targeting these receptors may fail in the late stages of BC, as resistance is acquired due to therapy-related clonal and cellular alterations [225]. Investigation of the mechanisms underlying temporal heterogeneity and clonal evolution may help to modify targeted therapies.

1.11 Hypothesis

BC is heterogenous disease in terms of morphology, molecular profile, behaviour and response to therapy. The overall hypothesis of this thesis is that the heterogeneity of ER, BCSCs, and immune infiltrates in BC are associated with differential prognosis and patient outcome. This will provide additional prognostic insight and potentially improve clinical decision making in BC, and thus enhance personalised treatment.

1.12 Aims

The aims of this thesis are to:

(i) Investigate the prognostic significance of ALDH1A1 and BMI-1 expression in various BC subtypes, and explore the correlations between ALDH1A1 and BMI-1 expression and the expression of other BCSC markers.

(ii) Identify BCSC subpopulations in BC tissue samples from a large, well-characterised BC cohort using double immunohistochemical staining and examine the relationships between BCSC phenotypes and the clinicopathological features of BC and patient outcome.

(iii) Assess low ER+ expression (1–9% positive cells) at both the protein (IHC) and mRNA (RNAscope) levels in needle core biopsies and the corresponding surgical excision specimens from patients with BC.

(iv) Assess the heterogeneity of ER expression in primary IBC and the paired lymph node metastases and explore the associations between intratumoral ER heterogeneity and tumour behaviour and patient outcomes.

(v) Investigate the prognostic and predictive value of SLC39A6 in invasive BC, particularly in ER+ tumours.

(vi) Quantify the spatial and temporal distribution of TIL infiltrates in BC and explore the associations with patient outcomes.

Chapter 2

Materials and Methods

2.1 Study cohorts

2.1.1 Nottingham patient cohorts

Four cohorts of patients with breast cancer (BC) treated at a single institute, Nottingham City Hospital, between 1988 and 2019 were assessed in this thesis. The primary cohort (*n* = 930) and secondary cohort (*n* = 875) included well-characterised, consecutive patients with early-stage operable primary invasive BC diagnosed between 1988 to 1998 and 1998 to 2006, respectively, for whom long-term clinical follow-up was available. Samples from the primary and secondary cohorts were used to construct tissue microarrays (TMAs).

The third series, the low oestrogen receptor (ER) cohort (*n* = 143), was selected from the Nottingham City Hospital patient database. Patients treated with primary invasive BC at Nottingham City Hospital who had known ER status based on a needle core biopsy (NCB) were included in the low ER series. A fourth series, the ER heterogeneity series (*n* = 502), was also identified from the Nottingham City Hospital database. Patients with tumours expressing 1-70% ER+ cells, and 100% ER+ cells and known lymph node status were included in the ER heterogeneity series.

Data on the clinicopathological features of the cohorts were extracted, including age at diagnosis, ethnicity, tumour grade, tubular formation, pleomorphism score, mitotic figure score, histological type, tumour stage,

Nottingham Prognostic Index (NPI) and lymphovascular invasion (LVI). Tumour stage was classified as stage 1, node-negative; stage 2, 1–3 positive nodes; stage 3, ≥ 4 positive nodes. NPI was categorised as good, NPI score ≤ 3.4; moderate, 3.41–5.4; poor, > 5.4) [226]. Data on oestrogen receptor (ER), progesterone receptor (PR) and HER2 (human epidermal growth factor 2) status were also collected for each patient from the original pathology reports.

The molecular subtypes of BC were classified based on immunohistochemical (IHC) analysis of ER, PgR, HER2 and Ki67 expression as Luminal A (ER+/HER2−; Ki67 < 10%), Luminal B (ER+/HER2−; Ki67 ⩾ 10%), HER2-positive (HER2+, regardless the ER status) and triple negative (TN; ER−, PgR− and HER2 −). In this study HER2+ subtype is defined as HER2 positive regardless the ER status, to investigate the prognostic impact of the studied biomarkers on the HER2+ tumour with no interference of the ER status.

Survival data for these cohorts was extracted from a routinely updated database. Survival status was recorded as alive or dead (and the cause of death was classified as BC or other causes). Survival time (in months) was defined as the time from date of diagnosis until last follow-up (dead or still alive). Breast cancer-specific survival (BCSS) was defined as the time (in months) from the date of primary surgical treatment to death due exclusively to breast cancer. The disease-free interval (DFI) is the time (in months) from the date of surgery to the date at which recurrence was first detected, including local and regional recurrence. Post-recurrence survival was defined as the time (in months) from the date of surgical removal of the recurrent tumour to the date of death due to BC. Distant metastasis-free survival (DMFS) was defined as the time (in months) from the primary surgery until distant metastasis was first detected.

All patients either underwent mastectomy or wide local excision with radiotherapy. Data on adjuvant therapy was also retrieved, including radiotherapy, hormonal therapy and chemotherapy. Treatment was uniformly selected based on tumour features, NPI and hormone receptor status. Endocrine therapy was administered to patients with ER+ tumours and high NPI scores (> 3.4); no adjuvant therapy was given to patients with excellent NPI scores (≤ 3.4). Premenopausal patients with moderate and poor NPI scores were candidates for chemotherapy, while postmenopausal patients with moderate or poor NPI scores only received hormonal therapy. None of the patients in the study cohorts received neoadjuvant therapy.

Table 2.1 summarises the patient demographics and clinicopathological features of the primary, secondary and low ER Nottingham cohorts; Table 2.2 summarises the patient demographics and clinicopathological features of the Nottingham ER heterogeneity cohort.

2.1.1.1 Confirmation of ER, PgR and HER2 status for the Nottingham cohorts

Expression of the hormone receptors ER and PgR in the Nottingham cohorts was confirmed by re-staining and reassessing the tissue sections according to the recent ASCO guidelines [2]. ER and PgR positivity were defined as ≥ 1% of tumour cells exhibiting nuclear staining.

The primary and secondary series were also re-stained for HER2. HER2 positivity was defined as ≥ 10% of the tumour cells exhibiting strong membranous staining (score of +3). A score of +2 was defined as weak to moderate intensity staining in > 10% of the tumour cells or incomplete membrane staining for >

10% of the tumour cells. Negative HER2 staining (score 0 or 1) was defined as no or incomplete membrane staining in < 10% of tumour cells [3]. Chromogenic *in situ* hybridisation (CISH) was used to assess *HER2* gene amplification status in borderline cases with an IHC score +2.

Ki67 was also stained and assessed in the primary cohort. Ki67 positivity was defined as > 10% positive tumour cell nuclei [4].

Table 2.1: Clinicopathological features of the primary, secondary and low oestrogen receptor (ER) Nottingham cohorts.

Clinicopathological parameters	**Nottingham primary series** *N* (%)	**Nottingham secondary** *N* (%)	**Low ER cohort** *N* (%)
Age at diagnosis			
≥ 50 years	714 (61.3)	602 (69)	95 (69)
< 50 years	395 (38.7)	273 (31)	43 (31)
Tumour size			
≥ 2 cm	577 (52.0)	315 (36)	77 (56)
<2 cm	532 (48.0)	559 (64)	61 (44)
Grade			
Grade I	190 (17.2)	135 (15)	6 (4)
Grade II	366 (33.1)	347 (40)	32 (23)
Grade III	550 (49.7)	392 (45)	99 (72)
Tumour type			
Ductal (Inc. mixed)	922 (83.1)	502 (90)	98 (72)
Lobular	101 (9.1)	28 (5)	8 (6)
Medullary-like	26 (2.3)	10 (2)	21 (15)
Miscellaneous	7 (0.6)	1 (0.2)	2 (1)
Special type	53 (4.9)	15 (2.8)	7 (5)
Vascular invasion status			
Definite	382 (34.6)	230 (26)	15 (14)
Negative/Probable	722 (65.4)	644 (74)	94 (86)
Lymph node stage			
Stage I	674 (70.0)	596 (68)	77 (73)
Stage II	341 (30.8)	240 (27)	22 (20)
Stage III	91 (8.2)	38 (5)	7 (7)
Follow-up status			
Alive	569 (51.3)	649 (74)	112 (81)
Died of BC	365 (33.0)	151 (17)	19 (14)
Died of other causes	175 (15.7)	75 (9)	7 (5)
Oestrogen receptor status			
Negative	270 (24.6)	175 (20)	27 (20)
Positive	827 (75.4)	694 (80)	115 (80)
Progesterone receptor status			
Negative	435 (40.8)	363 (42)	98 (77)
Positive	630 (59.2)	508 (58)	27 (23)
Human Epidermal Growth Factor Receptor 2 status			
Negative	921 (86.6)	758 (87)	72 (51)
Positive	143(13.4)	117 (13)	42 (49)

Table 2.2: Clinicopathological features of the Nottingham oestrogen receptor (ER) heterogeneity cohort.

Clinicopathological feature	**ER+ (1-70%)** ***N*** **(%)**	**ER+ (100%)** ***N*** **(%)**
Age at diagnosis		
< 50 years	32 (28.6)	28 (7)
≥ 50 years	80 (71.4)	363 (93)
Tumour size		
< 2 cm	70 (62.5)	277 (71)
≥ 2 cm	42 (37.5)	114 (29)
Tumour Grade		
Grade I	16 (14.3)	116 (30)
Grade II	42 (37.5)	185 (47)
Grade III	54 (48.2)	90 (23)
Axillary nodal stage		
Stage I	68 (60.7)	284 (73)
Stage II	26 (23.2)	87 (22)
Stage III	15 (13.4)	20 (5)
Lymphovascular invasion		
Negative	74 (66.1)	308 (79)
Positive	25 (22.3)	83 (21)
Nottingham Prognostic Index		
Poor Prognostic Group	15 (13.4)	27 (7)
Moderate Prognostic Group	59 (52.7)	153 (39)
Good Prognostic Group	33 (29.5)	210 (54)
Histological type		
Ductal (including mixed)	91 (81)	312 (80)
Special type	8 (7)	29 (7)
Lobular	12 (11)	39 (10)
Non-special type	0 (0)	1 (0.2)
Tubular carcinoma	0 (0)	23 (2.8)
Patient follow-up		
Alive	57(51)	309(79)
Died of BC	19(17)	40(10)
Died of other causes	11(10)	57(15)
Endocrine therapy		
Yes	43 (39)	215 (55)
No	38 (61)	76 (19)

2.1.1.2. Availability of data on biomarker expression for the primary Nottingham cohort

A wide range of data on several biomarkers is available for the primary Nottingham cohort used in this study, including the expression of ER-related proteins (FOXA1, GATA3, TFF1) [227], breast cancer stem cell markers (CD24, CD44, CD133, SOX10, EpCAM, SOX9) [228-230], and immune markers (CD3, CD8, FOXP3, CD20, CD68, PDL-1, PD-1) [231-234]. Table 2.3 summarises the details of the antibodies used to detect these biomarkers in the primary series of the Nottingham cohort.

Table 2.3: Antibodies used for immunohistochemical analysis of the Nottingham cohort (All the antibodies listed below stained by the previous PhD students in the group)

Antibody	Clone	Dilution	Antigen retrieval	Source
CD3	SP7	1:100	Microwave in EDTA	Vector Laboratories
CD8	1A5	1:50	Microwave in citrate	Vector Laboratories
FOXP3	23A/E7	1:100	Microwave in EDTA	Abcam
CD20	L26	1:300	Microwave in citrate	Dako
CD68	KP1	1:300	Microwave in citrate	Dako
EpCAM	VU-ID9	1:100	Microwave in citrate	Abcam
CD24	SWA11	1:500	Microwave in citrate	German Cancer Research Centre
CD44	156-3C11	1:100	Microwave in citrate	Cell Signalling
CD133	AC-15	1:75	Microwave in citrate	Sigma
SOX9	PAS-86301	1:10000	Microwave in citrate	Santa Cruz
SOX10	CL4455	1:1000	Microwave in citrate	Novus Biologicals
GATA3	HG3-31	1:80	Microwave in citrate	Santa Cruz
TIFF	Ab57752	1:2000	Microwave in citrate	Abcam
FOXA1	2F83	1:2000	Microwave in citrate	Abcam

2.1.2 Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) cohort

The Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) cohort (*n* = 1980) was established from tumour banks in the UK and Canada to assess genomic alterations and altered mRNA expression in BC [235]. We acknowledge the METABRIC data team, who made this data available to use in this study.

Briefly, matched DNA and RNA samples were isolated from freshly frozen tumour samples and analysed using Affymetrix SNP 6.0 and Illumina HT-12 v3 platforms, respectively. The DNA samples were hybridized to Affymetrix SNP 6.0 arrays per the manufacturer's instructions (Affymetrix, Santa Clara, CA) at AROS Applied Biotechnology (Aarhus, Denmark). The samples met the quality control criteria established by AROS. The RNA samples were used to prepare biotin-labelled cRNA with the Illumina Totalprep RNA amplification kit (Ambion, Warrington, UK) and hybridized to Illumina Human HT-12 v3 Expression Beadchips, as per the manufacturer's instructions. The hybridized Beadchips were scanned using an Illumina BeadArray Reader, as described by Curtis *et al.* (2012). The DNA and RNA data were normalised using the log transformation method, as previously described [235].

Clinicopathological data and follow-up data are available for this cohort including age at diagnosis, tumour grade, histological type, tumour stage, LVI, nodal stage and BCSS. The clinicopathological features of the METABRIC cohort are summarised in Table 2.4.

Patients with ER-positive and/or lymph node-negative tumours did not receive adjuvant chemotherapy; patients ER-negative and/or lymph node-positive

tumours received adjuvant chemotherapy. Additionally, patients with the HER2+ subtype did not receive trastuzumab.

Table 2.4: Clinicopathological features of the METABRIC cohort

Clinicopathological feature	**METABRIC cohort *N* (%)**
Age at diagnosis	
≥ 50 years	1426 (78.6)
< 50 years	424 (21.4)
Tumour size	
≥ 2 cm	1337 (68.2)
<2 cm	623 (31.8)
Grade	
Grade I	169 (9.0)
Grade II	770 (40.7)
Grade III	952 (50.3)
Tumour type	
Ductal (including mixed)	1545 (83.6)
Lobular	148 (8.0)
Medullary-like	32 (1.7)
Miscellaneous	12 (0.6)
Special type	113 (6.1)
Lymph node stage status	
Stage I	1035 (52.5)
Stage II	623 (31.5)
Stage III	315 (16.0)
Follow-up status	
Alive	1070 (55.7)
Died of breast cancer	505 (26.3)
Died of other causes	345 (18.0)
Oestrogen receptor status	
Negative	472 (23.8)
Positive	1508 (76.2)
Progesterone receptor status	
Negative	938 (47.4)
Positive	1042 (52.6)
Human Epidermal Growth Factor	
Negative	1734 (87.5)
Positive	246 (12.5)

2.2 Assessment of tumour burden for tissue specimens from the Nottingham cohort

Formalin-fixed paraffin-wax-embedded (FFPE) blocks were retrieved from the Nottingham Health Science Biobank and sectioned at 4 μm using a microtome (Leica RM2125; Leica Microsystems, Cambridge, United Kingdom). Haematoxylin and eosin (H&E) staining was performed for each case and the slides were reviewed using a Nikon light microscope to confirm the presence of a sufficient area of invasive tumour.

H&E-stained lymph node diagnostic slides were also reviewed by light microscopy to select positive lymph node sections with sufficient tumour area. All available lymph node positive tumour blocks were collected.

2.2.1 Haematoxylin and eosin (H&E) staining

H&E staining was performed using a Leica Autostainer XL (Leica Biosystems, Milton Keynes, UK). Prior to loading the slides, the purity of the haematoxylin solution was checked and the solution was filtered, if necessary. If present, scum was removed by drawing a piece of tissue paper across the top of the solution. Only recently prepared (within one month) eosin (#12668616; Thermo Fisher Scientific, Winsford, UK), acid alcohol, borax (#B9876; Sigma, Gillingham, UK) and isopropanol solutions (#LH-MO273E, Thermo Fisher Scientific) were used.

The FFPE tissue sections were heated at 60 °C for 10 min on a hotplate (Leica Biosystems), cooled, and loaded into the Leica XL Autostainer. The sections were deparaffinised in two 5-min changes of 100% xylene (#X/0250/PB17; Thermo Fisher Scientific), rinsed three times in 100% industrial methylated spirits

IMS (#11492874; Thermo Fisher Scientific) for 2 min to remove the xylene, and washed in tap water three times for 2 min each to remove the alcohol.

The tissue sections were stained in haematoxylin solution (#12658626; Thermo Fisher Scientific) for 7 min to stain the cell nuclei a reddish-purple colour, washed with running tap water for 5 min, excess background staining was removed by incubation in 1% acid alcohol for 30 sec, following by a wash in tap water for 20 sec, and the sections were blued with 1% Borax (#B9876; Sigma) for 40 sec to convert the haematoxylin purple shade into a dark blue colour.

Next, the sections were counterstained using an alcoholic solution of 1% eosin (#12668616; Thermo Fisher Scientific) for 30 sec to stain non-nuclear elements a pink colour. Then, the sections were rinsed in tap water for 15 sec, dehydrated by immersion in 100% propanol (#P/7507/PB17; Thermo Fisher Scientific) three times for 1 min each, incubated three times in 100% xylene for 5 min each to dehydrate and diaphanise the sections, mounted in DPX (Sigma), and dried overnight in a fume hood.

2.3 Immunohistochemical staining (IHC) of tumour sections from the Nottingham cohorts

2.3.1 Construction of tissue microarrays

Tissue microarrays (TMAs) had been previously constructed for the primary Nottingham cohort and were created for the secondary Nottingham cohort during this study. Briefly, full-face H&E-stained tumour sections were digitally scanned using the Panoramic 250 Flash II Scanner (3DHISTECH Ltd, Budapest, Hungary) and the digital images were annotated using Caseviewer software (Caseviewer v.2.3; 3DHISTECH Ltd, Budapest, Hungary). Representative areas matching the richest area of invasive tumour were selected for each TMA core.

The tumour blocks were imaged, drilled and punched using an automated GrandMaster TMA arrayer® (3D HISTECH) with a 0.6-mm core needle. A TMA Grand Master® with a punch set of 0.6 mm was used for automated TMA production. The donor blocks were inserted and loaded into the first row of the machine at a maximum of 60 blocks at a time. The digital H&E slides were overlaid and marked to guide the transfers from the donor blocks, and the marked area was punched from each donor block into the recipient block. The TMA blocks were incubated at 37° C for 20 min and 4 μm-thick sections were cut and stained with H&E for quality control purposes.

The array grids were designed to contain 150 cases/cores, subdivided into three subarrays. Additional orientation cores (normal kidney and/or normal liver) were added outside the margin of the array on the left, right, top and bottom to enable orientation. Each grid contained the unique ID number for each patient arranged in the same sequence as the corresponding position in the recipient block.

2.3.2 Validation of antibody specificity by Western blotting

Western blotting was performed to validate the specificity of the antibodies used for IHC staining.

2.3.2.1 Preparation of cell lysates for Western blotting

MCF-7, MD-MBA-231, SKbr3 and HeLa cell lines were obtained from the American Type Culture Collection (Rockville, MD, USA) and cultured according to the supplier's recommendations.

The cells were cultured to approximately 80-90% confluency by a researcher associate (Dr Chitra Joseph). The media was removed, the adherent cells were washed with 10 mL of 1x phosphate buffered saline pH 7.4 (PBS; #RNBH8631; Sigma), incubated with trypsin (#29259717; Sigma) for 5 min at 37 °C, and the trypsin in the detached cell suspension was neutralized by adding 10 mL of fresh cell culture media containing 10% foetal bovine serum (FBS; Sigma), which contains α1-antitrypsin to inhibit cell dissociation. The suspended cells were transferred to a sterile 50 mL tube, centrifuged at 1000 rpm for 5 min and the cell pellet was resuspended in 10 mL of complete media.

To count the viable cells, 50 µL of trypan blue (#72-57-1; Sigma) was mixed with 50 µL cell suspension in labelled microfuge tubes and the mixture was loaded into a Neubauer haemocytometer chamber. The cells in all of the squares in the small central grid were counted. The total number of cells per mL was calculated by determining the mean number of cells in the four central squares and multiplying by 10^4 (as the volume underneath the squares corresponds to 1.0

x 10^{-4} mL), and then multiplying by the trypan blue dilution factor (x2), using the following equation:

$$cell\ density\ (^{cells}/_{mL}) = \frac{total\ number\ of\ cells\ in\ 4\ squares}{4}\ x\ 10^4\ x\ 2$$

Cell suspensions containing 10 million cells (1.0 x 10^7)/total cell suspension were pelleted at 1000 rpm for 5 min, the supernatant was removed, the cell pellet was placed on ice, and the cells were resuspended and lysed in 1 mL RIPA buffer (#89901; Thermo Fisher Scientific) containing 1% protease/phosphatase inhibitors (Complete Mini Protease Inhibitor Cocktail tablets, #04693124001; Roche).

The lysates were incubated on ice for 15 min, gently inverting the tube every few minutes, centrifuged at 13,000 rpm for 10 min at 4 °C, and the supernatants were removed and stored at -20 °C for short-term storage or at -80 °C for long-term storage.

2.3.2.2 Electrophoresis and transfer

Protein lysates and the Prestained Protein Ladder marker (#26616; Thermo Fisher Scientific, UK) were thawed on ice. Each 30-µL sample contained 20 µL cell lysate, 7.5 µL of 4x Bolt LDS sample buffer (#NW04122BOX; Thermo Fisher Scientific) and 3 µL of 10x reducing agent (#NP0009; Thermo Fisher Scientific). The sample mixtures were vortexed and placed in a heating block for 5 min at 100 °C to denature the proteins.

The gel cassette (#NW04120BOX; Life Technologies) was placed into the running tank with the lower side of the dam facing inwards to the buffer chamber.

The tank was filled with 1x running buffer (950 mL deionised H_2O and 50 mL of 20x MOPS SDS Running Buffer, #NP000102; Invitrogen/Thermo Fisher Scientific) up to the cathode level and the comb was removed from the gel. The 30-µL protein samples and 7 µL of the protein ladder marker (#00328797; Thermo Fisher Scientific) were loaded into individual wells and 500 µL antioxidant (#1889064; Life Technologies) was added to the inner buffer chamber to maintain the proteins in a reduced state. Electrophoresis was performed at a constant voltage of 160 V for 60 min using a Bio-Rad Power supply.

After electrophoresis, the gel was removed from the cassette and sandwiched to a 0.045-µm pore size nitrocellulose membrane (#GE1060008; Sigma) between thick filter paper and sponges that had been pre-soaked in transfer buffer, as shown in Figure 2.3. The proteins were transferred to the membrane at a constant current of 10 V for 60 min.

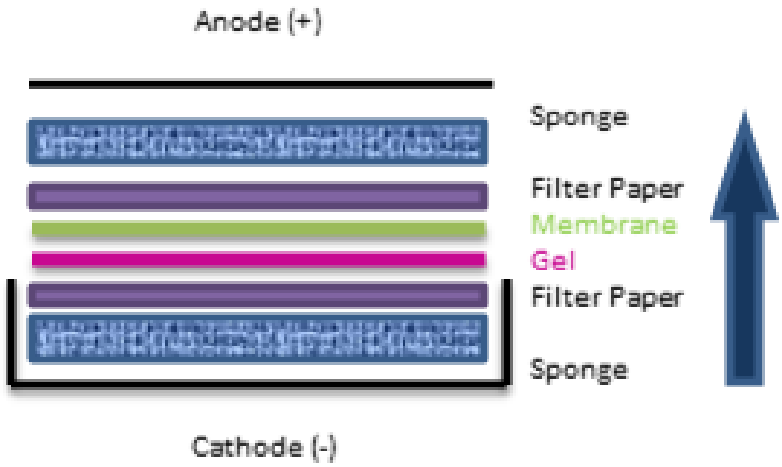


Figure 2.1: Assembly of a western blot transfer sandwich.

The membrane was placed face-up in blocking solution (2.5 g non-fat dried milk in 50 mL wash solution) and incubated for 60 min at room

temperature on an orbital shaker, followed by incubation with the primary antibody against the target protein and anti-β-actin antibody (as a loading control) at 4 °C overnight or at room temperature for 60 min. The antibodies were diluted in blocking solution to the concentrations recommended by the manufacturers and the β-actin (mouse or rabbit) antibodies were diluted at 1:5000. Table 2.5 lists the antibodies and dilution factors.

The membrane was subsequently incubated with the appropriate fluorescently labelled secondary antibodies (IRDye 800CW Donkey Anti-Rabbit or IRDye 680RD Donkey Anti-Mouse, Li-cor Biosciences, Cambridge, UK) for 60 min at room temperature on the orbital shaker while protected from light with aluminium foil. The secondary antibodies were diluted 1:15000 in 5 mL blocking solution. After 60 min incubation, the membrane washed three times for 10 min each in wash buffer (1 mL Tween-20 in 1,000 mL PBS).

Finally, the protein bands were visualised using an Odyssey Fc scanner and Image Studio 4.0 (LI-COR Biosciences) at wavelengths of 600, 700 and 800 nm.

2.3.2.3 Alternative enhanced chemiluminescence (ECL) detection method

The appropriate horseradish peroxidase (HRP)-conjugated secondary antibody (anti-rabbit or anti-mouse monoclonal horseradish peroxidase conjugate; Dako, Carpinteria, CA, USA), was diluted 1:1000 in blocking buffer and incubated with the membrane for 60 min at room temperature on an orbital shaker. Following three washes in wash buffer, ECL detection reagent (GE Healthcare, Buckinghamshire, UK) was added to the membrane for one minute and then

removed, and the bands were visualised using a LI-COR Odyssey Fc and the chemiluminescence detection module of Image Studio 4.0 (LI-COR Biosciences).

Table 2.5. Details of the antibodies used in this thesis.

Antibody	Catalogue number	Dilution WB	Dilution IHC	Incubation time* for IHC	Antigen retrieval	Supplier	Positive control
ALDH1A1	HPA002123	1:2500	1:100	1 h/overnight	Microwave in citrate	Sigma	Liver tissues
BMI-1	EPR3745	1:5000	1:100	1 h	Microwave in citrate	Abcam	Kidney tissues
EpCAM	MAB960	1:250	1:100	1 h	Microwave in citrate	R&D Systems	Colon tissue
SOX9	NBP2-24659	1:1000	1:100	1 h	Microwave in citrate	Novus Biologicals	Kidney tissues
SLC39A6	AA170320	1:1500	1:200	1 h	Microwave in citrate	LSBio	Kidney tissue

*The incubation time for all antibodies for western blotting was overnight.

2.3.3. Automated IHC using the Ventana Autostainer

IHC staining for ER and PgR were carried out on a Ventana Benchmark® ULTRA system (Roche) using the rabbit monoclonal primary anti-ER (SP1) antibody (#790-4325; Roche) and rabbit monoclonal primary anti-PgR (1E2) antibody (#790-4296; Roche) following the manufacturer's recommended protocol. This work has been performed by NHS technicians in histopathology laboratory in Queen medical centre, Nottingham.

Briefly, the tissue sections were heated at 72 °C for 10 min on a hot plate, Cell Conditioner 1 (CC1) (#952 124; Roche) was applied for 64 min for antigen retrieval, then a casein block (#06440002001; Roche) was applied at 100 µL/slide for 8 min. The diluted primary antibodies (100 µL/slide) were applied for 16 min at 37 °C, followed by the Ultraview detection kit (UV) for 10 min, then the OptiView HRP Multimer for 8 min, and the sections were counterstained with modified Mayer's haematoxylin (haematoxylin II) for 12 min.

The slides were removed from the machine, washed in Ezprep detergent (Roche) for 1 min, then washed in four changes of water for 10 sec each, followed by three changes of ~100% alcohol for 15, 20 and 10 seconds, respectively. Finally, the sections were incubated in xylene for 15 sec and mounted using Clearvue mountant XYL (Thermo Fisher Scientific).

The optimal concentrations for the primary antibodies had been previously determined by the supplier (Roche) and were validated by the IHC diagnostic laboratory team (Histopathology department, Queen's Medical Centre, Nottingham, UK). Internal and external controls were included in each run; specific details are provided in Chapter 6.

2.3.4 Manual immunohistochemistry

The NovolinkTM Max Polymer Detection System (#RE7280-K; Leica Biosystems) used in this study was previously optimised in our laboratory by Mr Christopher Nolan. The kit includes a peroxidase block, protein block, post-primary antibody, detection polymer, DAB (diaminobenzidine tetrahydrochloride) chromogen substrate, and haematoxylin counterstain.

The tissue sections were heated on hot plate at 60 °C for 10 min to dry and fix the sections and melt the paraffin, cooled for 5 min at room temperature, and the deparaffinization and rehydration steps were performed using the Leica Autostainer XL, as described in section 2.2.1.

2.3.4.1 Antigen retrieval

The slides were placed into a plastic container containing citrate buffer (1 M sodium citrate, pH 6.0), covered with a lid and microwaved (Whirlpool JT359 Jet Chef 1000 W) for 20 min at 1000 W, then cooled in running tap water for 10 min.

In accordance with the manufacturers' instructions, this antigen retrieval method was employed during the IHC staining procedures for all antibodies in this study.

2.3.4.2 Manual IHC process

A hydrophobic barrier was applied around the sections using a hydrophobic marker pen (Vector Laboratories) and the slides were placed in a humidified chamber. The peroxidase block from the NovolinkTM Max Polymer Detection System was applied for 5 min to inhibit endogenous peroxidase activity, the slides were washed three times with TBST for 5 min each, the protein block from the kit was applied for 5 min to block non-specific antibody binding and reduce background staining, then the slides were washed three times in TBST for 5 min.

The primary antibodies for each target protein were diluted to the optimal concentrations in Leica antibody diluent (#SLBL4196V; Leica Biosystems) and applied to the sections for 60 min at room temperature, followed by three washes with TBST for 5 min each. Next, 300 µL of the rabbit anti-mouse secondary antibody was applied for 30 min, the sections were washed three times for 5 min each, then the horseradish peroxidase (HRP) detection polymer was applied for 30 min at room temperature, and the sections were washed three times for 5 min each.

DAB working solution was prepared immediately before application according to the manufacturer's instructions by diluting DAB substrate 1:20 and protected from light until use. The sections were incubated with 300 µL DAB substrate for 5 min. DAB chromogen staining was observed as a permanent brown colour that developed slowly and was preserved after three washes with TBST.

Finally, the slides were counterstained with haematoxylin for 6 min, transferred to a rack and loaded into the Leica Autostainer, rinsed in tap water for 5 min, rinsed three times for 2 min in 100% IMS, immersed twice in xylene for 5 min, cleared in xylene, mounted with DPX (#3808600E, Leica Biosystems) and

allowed to dry overnight. Figure 2.2 illustrates the single IHC staining assay within the NovolinkTM Max Polymer Detection System.

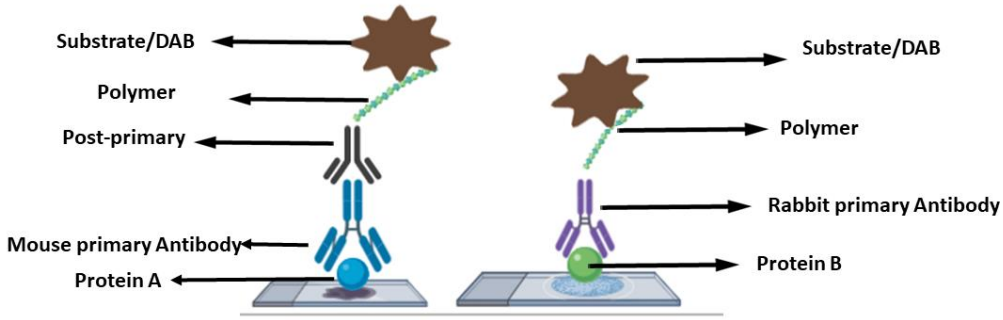


Figure 2.2: Schematic illustration of the single IHC assay in the NovolinkTM Max Polymer Detection System. Proteins A and B are detected on separate slides using primary mouse IgG or IgM or rabbit IgG antibodies. The post-primary rabbit anti-mouse immunoglobulin binds to any mouse primary antibody. The Novolink™ Polymer recognizes rabbit immunoglobulins, thus detects the post-primary antibody and rabbit primary antibodies. The Poly-HRP-IgG Polymer is applied to detect the binding sites using the chromogen DAB.

2.3.5 Optimisation of the antibodies for IHC

The antibodies were optimised by staining recently cut sections from different molecular subtypes of BC. Initially, the antibody manufacturers' recommendations in terms of antibody concentration range, incubation time and antigen retrieval method were tested. The positive control for each antibody was selected according to the supplier's datasheet or the Human Protein Atlas, as shown in Table 2.5. For the negative controls, the primary antibody was replaced with only the antibody diluent.

The stained sections were assessed on the following day by a pathologist and trained observer. The optimal antibody concentration was then selected or the optimisation steps were repeated and the variables in the optimisation protocol were adjusted until the desired result was obtained. The optimal antibody concentration was selected on the basis of no/low background staining and a high degree of heterogeneous staining between TMA cores.

Prior to immunostaining the TMAs, full-face tissue sections from 20 randomly selected FFPE BC specimens were subjected to IHC for ALDH1A1, BMI-1 and SLC39A6 and the staining distributions were assessed to confirm the appropriateness of using TMAs.

2.3.6 Assessment of IHC staining using the H-score

Stained TMA slides were scanned as high-resolution digital images using a NanoZoomer (Hamamatsu Photonics, Welwyn Garden City, UK) at x20 magnification and scored visually on high-resolution monitors using xplore Viewer software (Philips, Belfast, UK).

Prior to scoring the slides, all TMA slides were evaluated to estimate the threshold intensity level for each biomarker. Only tumour cores containing at least 15% invasive cancer cells within the tissue cores were considered.

Expression of the biomarkers was visually evaluated in a semi-quantitative manner using the modified Histochemical-score (H-score), as previously described [236]. Both the percentage and intensity of stained cells were assessed. Intensity was determined as 0, 1, 2 or 3 for negative, weak, moderate and strong staining, respectively. The H-score was estimated by multiplying the intensity score by the percentage of cells stained positive for each core. Thus, the H-score ranges from 0 to 300. Scoring was conducted in a blinded manner, without access to the clinicopathological or outcome data for the patients.

2.3.7 Double IHC staining

Double IHC was used to detect two distinct proteins/antigens in the same BC tissue specimens. The Double IHC Kit from A. Menarini Diagnostics (Winnersh, UK) was used for double staining.

2.3.7.1 Experimental steps

The deparaffinisation, rehydration, antigen retrieval, and blocking steps were performed as described in sections 2.2.1, 2.3.4.1 and 2.3.4.2

The first primary antibody (anti-ALDH1A1) was diluted to the optimal concentration (1:100) in Leica antibody diluent and 300 µL was applied to the section overnight at 4 °C. The next day, the slides were washed three times in TBS-T for 5 min each, incubated with the second primary antibody (anti-SOX9 or

anti-EpCAM) diluted 1:100 for 1 h at room temperature, and washed three times for 5 min each.

The detection kit (#MP-XLCT53-K25; A. Menarini Diagnostics) contains goat anti-mouse alkaline phosphatase (AP)-polymer conjugated and goat anti-rabbit HRP-polymer conjugated secondary antibodies, which react with the heavy and light chains of mouse and rabbit antibodies, respectively. The secondary antibodies were applied for 30 min at room temperature, then the sections were washed three times for 5 min each.

Fast-red chromogen (#MP-FR805-25; A. Menarini Diagnostics), which binds to the alkaline phosphatase (AP) enzyme and produces a bright fuchsin-red precipitate, was applied to the tissue sections for 10 min at room temperature, then the slides were washed three times for 5 min each. Green chromogen (#MP-807-K25, A. Menarini Diagnostics), which binds to HRP, was applied to the tissue sections for 10 min, then the slides were washed 3 times for 5 min with deionised water. Excess water was removed from the slides using filter paper and the slides were placed in an oven 37 °C to dry for 30 min and mounted with VectaMount (#H-5000, Vector Laboratories, UK). Figure 2.3 illustrates the A. Menarini Diagnostics Double IHC Detection System.

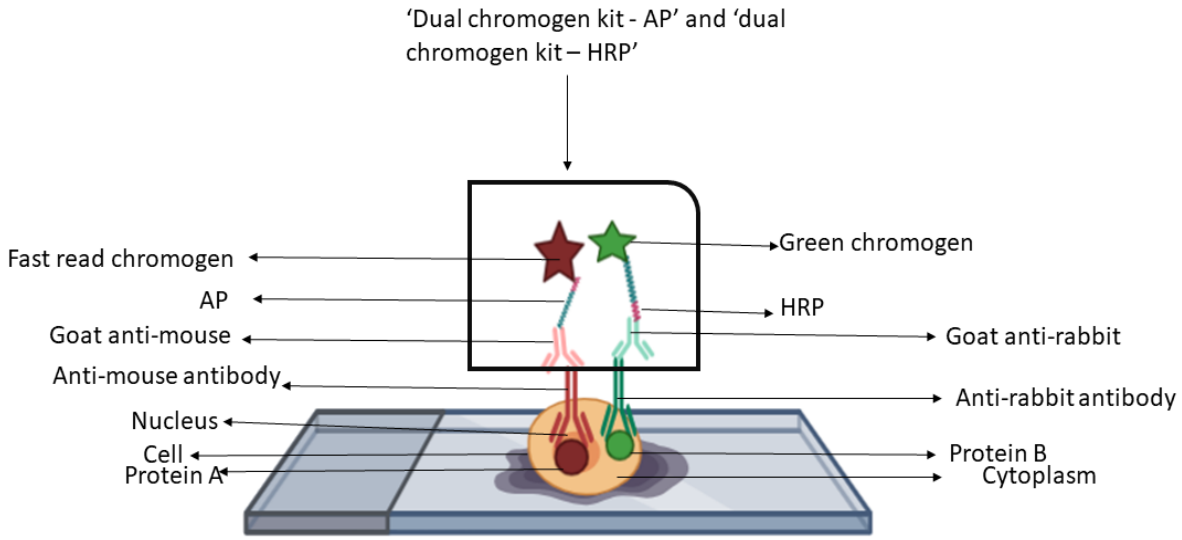


Figure 2.3: Schematic illustration of the double IHC staining assay. Primary anti-rabbit and primary anti-mouse antibodies are used to individually detect two specific protein targets in the same tissue section. The dual detection kit includes two secondary antibodies (goat anti rabbit conjugated to HRP and goat anti mouse conjugated to PA). The staining is detected by metabolism of green chromogen and red fast by HRP and AP, respectively.

2.3.7.2 Optimisation of antibodies for double IHC

The concentration and incubation time of the antibodies were initially optimised individually to produce the desired results, to achieve little or no background and varied degrees of heterogenous expression between the TMA cores. The optimal concentrations were 1:50, 1:100, and 1:200 for the ALDH1A1, SOX9, and EpCAM antibodies, respectively.

When the antibodies were combined for the double IHC staining (ALDH1A1+SOX9, and ALDH1A1+EpCAM), the concentrations and incubation times were adjusted slightly to provide the optimal results. The optimal conditions selected for double IHC were 1:100 with overnight incubation for the ALDH1A1

antibody and 1:100 with 1 hour incubation for both the SOX9 and EpCAM antibodies.

2.4 Chromogenic *in situ* hybridization assay

A chromogenic *in situ* hybridization (CISH) assay was performed on FFPE BC tissue sections using specific probes to detect the expression of individual *ESR* and *PgR* mRNA molecules using a hybridization-based signal amplification method [237, 238] with a commercial RNAscope assay kit (#322350; Advanced Cell Diagnostics, Newark, CA, USA).

To confirm the reliability of the RNA protocol and tissue quality, CISH was performed in parallel to evaluate the positive control housekeeping gene peptidyl prolyl isomerase B (*PPIB*; #310043; Advanced Cell Diagnostics) and the negative control bacterial gene *BA-DapB* (#701021; Advanced Cell Diagnostics).

Briefly, tissue sections were deparaffinised by heating at 60 °C for 60 min, incubated in xylene twice for 5 min at room temperature (RT) with occasional vertical shaking, immediately rehydrated by incubation with IMS (Thermo Fisher Scientific) twice for 1 min, and dried on a hot plate for 5 min. Hydrogen peroxide was applied for 10 min, followed two 5-min washes in deionised water. The slides were immediately placed in target retrieval reagent, microwaved at 750 W for 10 min and 450 W for 10 min, washed in deionised water for 1 min, placed in IMS for 3 min, dried on a hot plate at 60 °C for 15 min, and a hydrophobic barrier pen (#H-4000; Vector Laboratories) was used to mark the sections. The protease was applied to the slides and the sections were placed in a humidified chamber for 30 min at 40 °C in a HybEZ hybridisation oven (Thermo Fisher Scientific), then the sections were washed three times with deionised water for 2 min each.

The target probes, positive control PPIB-probe and negative control probe were applied and the slides were incubated at 40 °C for 2 h in the HybEZ hybridisation oven. Next, the sections were washed twice with washing buffer (60 mL washing buffer reagent added to 2.94 L deionised water) for 2 min.

The preamplifier was applied for 30 min at 40 °C. Amplifier 1, 2, 3, and 4 were applied in order at 40 °C for 30 min each, with a 2-min wash between each step. Amplifier 5 and 6 were applied for 30 min at RT, then fast-red chromogen was added (120 µL per slide) for 10 min, the sections were washed twice in tap water, the counterstain was applied for 1 min, and the sections were washed in running tap water, dried for 15 min and mounted with VectaMount (Vector Laboratories).

The tissue sections were visualised under a standard bright-field microscope at 20–40X magnification. The positive control showed strongly visible punctate dots. The negative control showed a clear background with only up to 1 dot per 10 cells, which was considered acceptable.

2.5 Statistical analysis

Statistical analysis was conducted using SPSS statistical software (IBM SPSS 25.0 (Chicago, IL, USA).

2.5.1 Dichotomisation of data

Prior the analysis, the cut-offs for high/low expression of each marker (ALDH1A1, BMI-1, and SLC29A6) in IHC-stained were determined using the non-commercial research tool X-tile software (University of Yale, Yale, USA; Camp et al., 2004). The data was dichotomised using cut-off points resulting from the prediction of BCSS [239].

2.5.2 Analysis of the associations between clinicopathological parameters and biomarker expression

The associations between expression of the biomarkers as categories (i.e., high/low expression) and clinicopathological parameters were assessed using the Chi-square test.

Continuous variables with a non-Gaussian distribution were analysed using the Mann-Whitney test and Kruskal-Wallis test for two and three or more categorical groups (clinicopathological parameters), respectively.

The correlations between normally and non-normally distributed continuous variables were examined using the Spearman's rank and Pearson tests, respectively. *P*-values less than 0.05 (two- tailed) were considered significant for all statistical tests.

2.5.3 Univariate analysis and multivariate analysis

Survival rates were determined by univariate analysis using Kaplan-Meier survival curves and the log-rank test. Patients were censored if they were lost to follow up

or if they died of causes other than BC. The end-points assessed were BCSS, DMFS and DFI. Two-tailed *P*-values less than 0.05 were considered significant.

Cox proportional hazards regression models were constructed to perform multivariate survival analyses in order to estimate hazard ratios (HRs) and investigate the independent prognostic significance of differential gene or protein expression when adjusted for other standard prognostic factors, including tumour grade, tumour size and histological stage [240]. *P*-values < 0.05 were considered significant.

Chapter 3

Prognostic significance of ALDH1A1 expression in invasive breast cancer

3.1 Introduction

The aldehyde dehydrogenase (ALDH) superfamily is comprised of 19 genes in 11 families and four subfamilies with distinct chromosomal locations [241, 242]. The ALDH enzymes can be found in the cytosol, nucleus, mitochondria and endoplasmic reticulum of the cell [243, 244]. The enzymes of the ALDH1 family are involved in the metabolism of several aliphatic and aromatic aldehydes [245] that play important roles in physiological processes and have tumorigenic and cytotoxic effects [246].

Several ALDH1 isoforms, including ALDH1A1, ALDH1A2 and ALDH1A3, share the same subcellular cytoplasmic and mitochondrial localisation and tissue distribution. However, these enzymes have different metabolites [247, 248]. ALDH1A1 is one of three highly conserved cytosolic isozymes (along with ALDH1A2 and ALDH1A3) involved in the retinoic acid (RA) pathway [249]. All members of the ALDH1 family convert retinol to retinaldehyde, and ALDH1A1 specifically oxidises retinaldehyde to RA [250, 251]. RA binds to nuclear retinoid receptors (RARs), leading to transactivation and/or transrepression of downstream target genes, such as retinoic acid receptor beta (*RARβ*). RA binds to retinoid X receptors (RXRs) and dimerises to peroxisome proliferator activated receptor β/δ (PPARβ/δ) to induce the expression of downstream target genes. Moreover, cooperative

interaction between oestrogen receptor alpha (ERα) and retinoic acid receptor alpha (RARα) indicates the RARα is involved in ERα-dependent gene transcription and cell proliferation [252, 253]. RA also activates the PI3K/Akt pathway through binding to the RARα in the cytoplasm [254].

ALDH1A1 is also regulated by the Wnt pathway via β-catenin/TCF-dependent transcription [255]. Deactivation of β-catenin/TCF transcription using a siRNA reduced ALDH1A1 metabolic activity [256]. One study suggested a MUC1-C→ERK→CEBPβ→ ALDH1A1 pathway exists in BC [16]. Notch also promotes ALDH activity through inducing the deacetylase SIRT2, which deacetylates and activates ALDH1A1, while the acetyl transferase PCAF induces acetylation of ALDH1A1 and inhibits ALDH activity. Figure 4.1 illustrates the various signalling pathways involved in regulation of ALDH1A1.

Much emphasis has been placed on ALDH1A1 as a CSC marker. ALDH1A1 is the most widely used BCSC marker and has been studied in a variety of BC models [257-259]. Cells with high expression of ALDH1A1 display stem cell-like features, such as self-renewal, differentiation and tumour-initiating ability [249]. High ALDH1A1 expressing cell populations from breast tumour have been reported to give rise to myoepithelial (CD10+/ESA-(Epithelial Specific Antigen)), luminal epithelial (CD10-/ESA+) and mixed colonies (CD10+/ESA+), while the low ALDH1-expressing populations only generate luminal epithelial colonies (CD10-/ESA+)[260].

Additionally, high ALDH1A1 expression is a prognostic marker associated with varied outcomes in several tumour types [261-263]. Croker *et al.* (2017) reported that knockdown of *ALDH1A1*, but not *ALDH1A3*, reduced metastatic behaviour and resistance to both chemotherapy and radiation therapy in BC cells

[264]. ALDH1A1 may also play a role in angiogenesis, as human umbilical vein endothelial cells (HUVEC) co-cultured with MCF-7 breast cancer cells express varied levels of ALDH1A1 [265].

Although several studies have demonstrated a role for ALDH1A1 in BC, the clinical significance of ALDH1A1 protein expression (based on IHC) and mRNA expression in different subtypes of BC are poorly characterised. Moreover, the association between the expression of ALDH1A1 and expression of other BCSC markers using BC tissue sample remain to be defined.

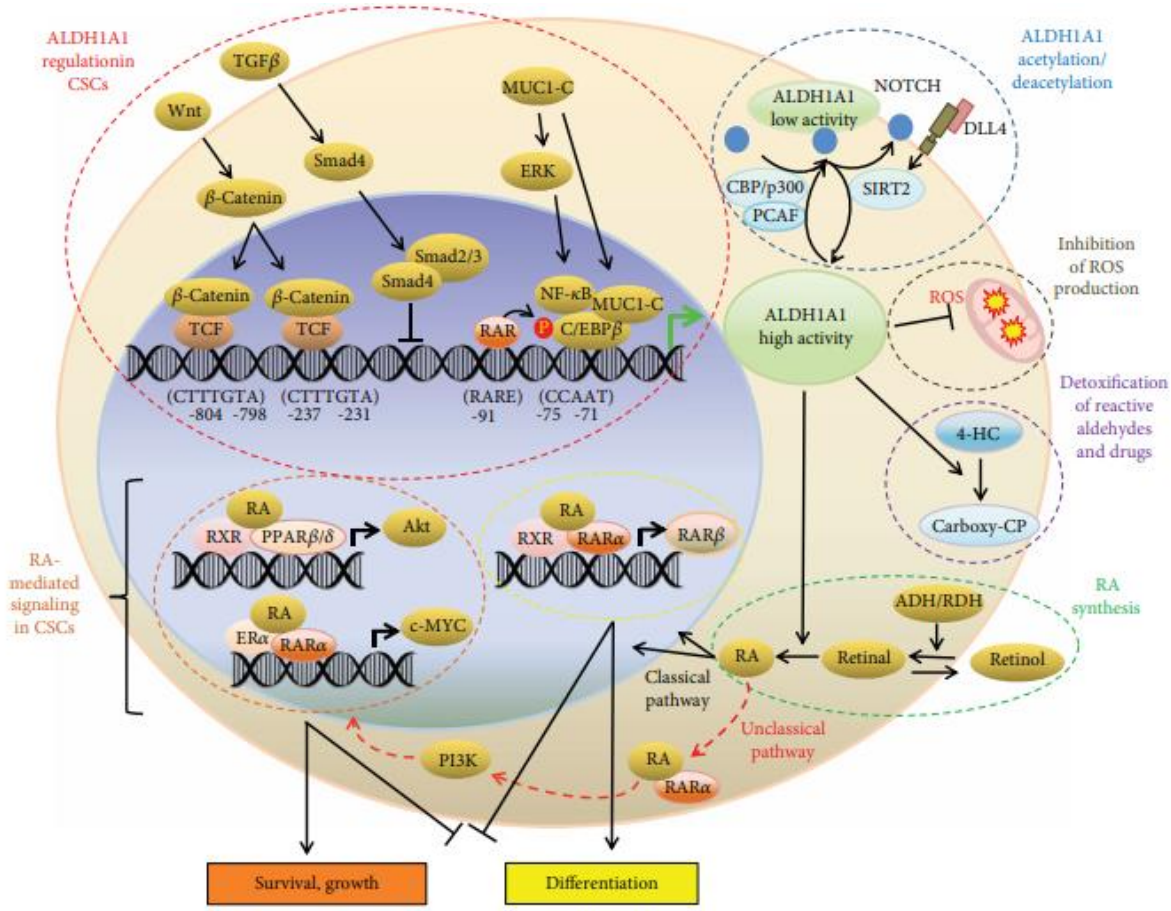


Figure 3.1: Several pathways regulate the ability of ALDH1A1 to affect cell growth, differentiation and survival. The classic retinoid singling pathway promotes expression of RARβ (yellow dotted line). The nonclassical RA pathway (orange dotted line) promotes expression of downstream RA genes, including P13K. The Wnt pathway regulates ALDH1A1 via β-catenin/TCF transcription. MUC1-C and ERK complex signalling also occur upstream of ALDH1 transcription. SMAD4 is regulated by β-catenin/TCF, and SMAD4 downregulates ALDH1A1. Notch promotes/inhibits ALDH1 activity by promoting acetylation or deacetylation of ALDH1A1 by inducing PCAF or SIRT2, respectively. Image adapted from Vassalli (2019) [255].

3.1.1 Hypotheses

The hypotheses tested in this chapter are:

(i) ALDH1A1 is a prognostic marker with distinct associations to tumour behaviour and patient outcomes in different subtypes of IBC.

(ii) ALDH1A1 expression is associated with the BCSC phenotype.

3.1.2 Aims of the study

The aims of this chapter are to:

(i) Investigate the prognostic significance of ALDH1A1 expression in various subtypes of BC.

(ii) Explore the correlations between ALDH1A1 expression and the expression of other BCSC markers.

3.1.3 Objectives of the study

The objectives of this chapter are to:

(i) Assess ALDH1A1 protein and mRNA expression in different subtypes of IBC.

(ii) Investigate the associations between ALDH1A1 mRNA and protein expression and the clinicopathological features and patient outcomes in various subtypes of IBC.

(iii) Assess the correlations between ALDH1A1 mRNA and protein expression and the expression of other BCSC markers.

3.2 Materials and Methods

3.2.3 Characteristics of the study cohorts

The study cohort used to evaluate ALDH1A1 protein expression by IHC was comprised of 930 patients with IBC derived from a retrospective series of patients with primary breast carcinoma who presented at Nottingham City Hospital between 1988 and 1998 (see Chapter 2 for more details). Table 3.1 summarises the demographic features of the Nottingham cohort; 64% of the patients were younger than 50 years-old at diagnosis, 47% had grade 3 BC, 19% had triple negative BC and 30% died from BC during the 30-year follow-up.

ALDH1A1 mRNA expression and gene copy number (CN) aberrations were examined using the METABRIC cohort of 1,980 BC patients. DNA and RNA were isolated from freshly frozen samples, as described in Chapter 2, we acknowledged the METABRIC group, who made this data available for us. The clinicopathological features of the METABRIC cohort are summarised in Table 3.1; 21% of patients were younger than 50 years-old at diagnosis, 50% had grade 3 tumours, 23% had ER- tumours and 26% died of BC over the 30-year follow-up.

3.2.4 Availability of biomarker data

Data on IHC analysis of the expression of BCSC markers, including CD133, CD24, CD44, EpCAM, SOX9 and SOX10, was available for the Nottingham study cohort. Details of these biomarkers are shown in Table 2.2 (Chapter 2).

Table 3.1: Summary of the patient demographics of the cohorts used to analyse ALDH1A1 protein and mRNA expression.

Clinicopathological parameter	Nottingham series *N* (%)	METABRIC series *N* (%)
Age, years		
≥ 50	267 (36)	1426 (79)
< 50	472 (64)	424 (21)
Tumour size		
> 2 cm	380 (52)	1337 (68)
≤ 2 cm	355 (48)	623 (32)
Tumour grade		
Grade 1	148 (19)	169 (9)
Grade 2	263 (34)	770 (41)
Grade 3	373 (47)	952 (50)
Tumour type		
Ductal (including mixed)	596 (81.6)	1545 (84)
Lobular	79 (10.8)	148 (8)
Medullary-like	13 (1.7)	32 (2)
Miscellaneous	2 (0.2)	12 (0.6)
Special type	40 (5.4)	113 (6)
Vascular invasion		Not available
Definite	244 (33)	
Negative/probable	490 (67)	
Lymph node stage		
Stage 1	527 (65)	1035 (52)
Stage 2	213 (27)	623 (31)
Stage 3	66 (8)	315 (16)
Follow-up status		
Alive	474 (56)	1070 (56)
Died of breast cancer	249 (30)	505 (26)
Died of other causes	118 (14)	345 (18)
Oestrogen receptor status		
Negative	197 (25)	472 (23)
Positive	578 (75)	1508 (76)
Progesterone receptor status		
Negative	281 (40)	938 (47)
Positive	426 (60)	1042 (53)
HER2 status		
Negative	621 (88)	1734 (88)
Positive	88 (12)	246 (12)

3.2.1 ALDH1A1 antibody validation

The specificity of the rabbit monoclonal anti-ALDH1A1 antibody (HPA002123; Sigma-Aldrich) was validated by western blotting of MCF7 and HeLa human cell lysates, as previously described (Chapter 2). The optimal dilution of the ALDH1A1 primary antibody for western blotting was 1:2500.

3.2.2 Tissue microarrays and immunohistochemical analysis

FFPE BC tissue samples were arrayed and immunohistochemical (IHC) staining was performed on 4 µm-thick TMA sections using a Novolink polymer detection system (Leica, Newcastle, UK), as described in Chapter 2.

Prior to immunostaining the TMAs, full-face tissue sections from 20 randomly selected FFPE BC specimens were subjected to IHC and the distribution of ALDH1A1 staining was assessed to confirm the appropriateness of using TMAs.

The optimal dilution of the ALDH1A1 antibody and incubation time for IHC staining were 1:100 and 1 h at room temperature. Normal liver tissues were used as a positive control, while the negative control was obtained by omitting the primary antibody.

ALDH1A1 staining was semi-quantitatively assessed using the modified histochemical score (H-score; see Chapter 2 for details). Inter-observer scoring reproducibility was examined by randomly selecting 10% of the cases; these cases were independently rescored by another assessor (M.Toss). Both assessors were blinded to the clinical data and patient outcomes.

3.2.5 Statistical analysis

IBM SPSS 24.0 software (Chicago, IL, USA) was used for statistical analysis, as described in Chapter 2. Inter-observer agreement was determined using kappa statistics. ALDH1A1 expression was categorised at the protein and mRNA levels using X-tile (X-tile Bioinformatics Software, Yale University, version 3.6.1) based on predictions of patient survival. The associations between ALDH1A1 categorical groups and clinicopathological parameters were tested using the Chi-squared test. The associations between ALDH1A1 expression and patient outcomes were assessed using Kaplan–Meier curves and the log-rank test. Cox proportional hazard regression models were constructed for multivariate survival analyses to estimate the hazard ratios (HRs) for the associations between ALDH1A1 expression and patient outcome adjusted for other prognostic factors.

3.3 Results

3.3.1 ALDH1A1 expression in BC

The specificity of the anti-ALDH1A1 antibody was validated by western blotting. A single band for ALDH1A1 was observed at the predicted size of 56 kDa, which confirmed the specificity of the antibody (Figure 3.2).

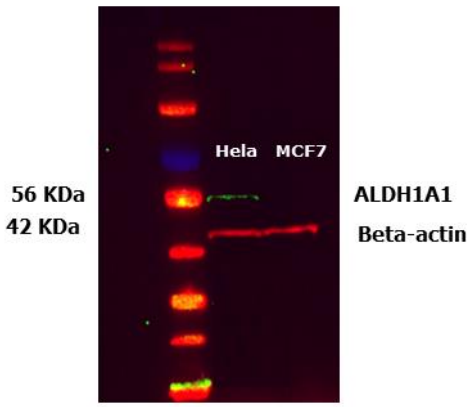


Figure 3.2: Western blot to confirm the specificity of the rabbit monoclonal anti-ALDH1A1 antibody (HPA002123; Sigma-Aldrich, UK). A single specific band (green) was observed at the expected molecular weight of 56 kDa in HeLa cell lysates. The red bands represent the β-actin positive control at 42 kDa. The protein ladder ranges from 10 to 180 KDa.

3.3.2 ALDH1A1 protein expression in BC

ALDH1A1 immunostaining was homogenously distributed through the full-face tissue sections, which validated the use of TMAs.

In analysis of the TMAs from the primary of Nottingham cohort, a high level of reproducibility in the scoring of ALDH1A1 expression between the two scorers was confirmed by the kappa statistic of 0.865. Inter-observer scoring reproducibility was examined by randomly selecting 10% of the cases; these cases were independently rescored by another assessor (M.Toss), then the kappa test was generated for the categorical scoring.

ALDH1A1 expression was mainly localised in the cytoplasm of the tumour cells, with the staining intensities varying from negative to high (Figures 3.3A and B). In addition, positive ALDH1A1 immunostaining was observed in stromal cells, including immune infiltrates and stromal fibroblasts (Figures 3.4A and B).

Interestingly, strong ALDH1A1 IHC staining intensities were observed in the minority of tumour cells, while the majority of cells showed weak to moderate staining intensities. Overall, 32 (3.4%) of the BC cases exhibited strong ALDH1A1 staining intensities (in 5% to 100% of tumour cells). Based on the X-tile-generated cut-off point of H-score of > 5, high ALDH1A1 expression was noted in 663/930 (71%) cases.

ALDH1A1 protein expression was also examined in the molecular subtypes of BC. ALDH1A1 was highly expressed in highly proliferative tumours, including the luminal B subtype (P = 0.028; Table 3.2).

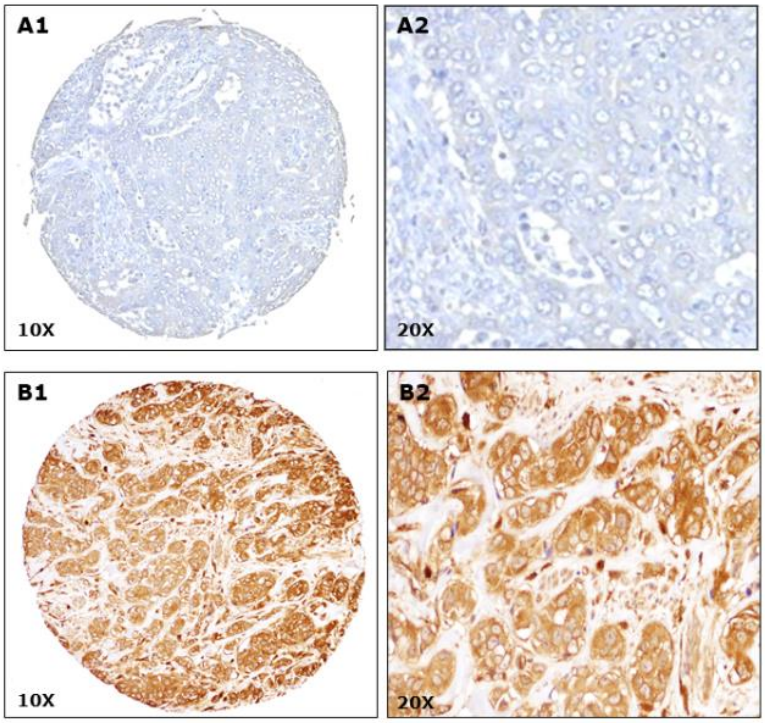


Figure 3.3: Immunohistochemical (IHC) staining for ALDH1A1 in invasive BC. (A) Negative ALDH1A1 staining; (B) positive ALDH1A1 staining.

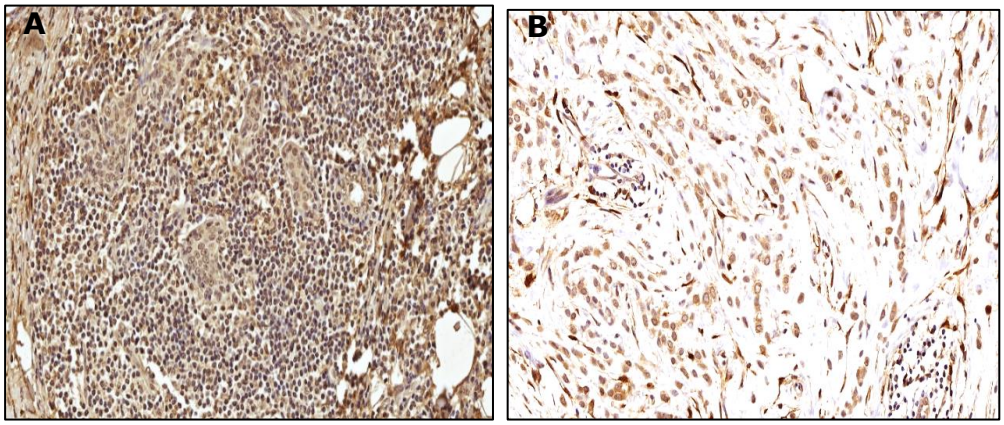


Figure 3.4: IHC analysis of ALDH1A1 expression in stromal cells (20×). (A) ALDH1A1-positive immune cells; (B) ALDH1A1-positive fibroblasts.

3.3.3 ALDH1A1 mRNA expression and CN changes

ALDH1A1 mRNA was highly expressed in 1,405/1,980 (71%) cases in the METABRIC cohort based on the X-tile cut-off point (6.9 log fold-change). *ALDH1A1* CN gain was detected in 35/1,980 (1.7%) cases and 32/1,980 (1.6%) cases showed CN loss. There were no significant associations between CN gain or loss and *ALDH1A1* mRNA expression ($P = 0.61$).

Of the cases present in both the Nottingham and METABRIC cohorts, *ALDH1A1* mRNA expression was not significantly associated with ALDH1A1 protein expression ($n = 184$; $P > 0.05$).

3.3.4 Associations between ALDH1A1 expression and clinicopathological features

High ALDH1A1 protein expression was associated with aggressive prognostic features, including high grade ($P = 0.007$), high mitotic count ($P = 0.004$), increased nuclear pleomorphism ($P = 0.001$), poor NPI ($P = 0.008$), advanced nodal stage ($P = 0.002$), lympho-vascular invasion ($P = 0.043$) and high Ki67 expression ($P = 0.045$). There were no significant associations between high ALDH1A1 protein expression and patient age ($P = 0.15$), tubule formation ($P = 0.56$) or tumour size ($P = 0.90$; Table 3.2).

High *ALDH1A1* mRNA expression was positively associated with positive axillary node involvement ($P = 0.002$), poor NPI ($P = 0.029$), small tumour size ($P < 0.001$) and the normal-like subtype ($P < 0.0001$; Table 3.2).

Table 3.2: Associations between ALDH1A1 protein and mRNA expression and clinicopathological features in breast cancer*

Clinicopathological feature	ALDH1A1 protein expression			ALDH1A1 mRNA expression		
	Low No. (%)	**High No. (%)**	**x^2 *P*-value**	**Low No. (%)**	**High No. (%)**	**x^2 *P*-value**
Patient age						
< 50 y	178 (30)	416 (70)	1.99	92 (25)	271 (75)	2.12
≥ 50 y	85 (26)	247 (74)	0.158	435 (29)	1055 (71)	0.145
Tumour size						
< 2 cm	134 (30)	320 (70)	0.730	124 (21)	467 (79)	220.80
≥ 2 cm	126 (27)	341 (73)	0.90	404 (32)	871 (68)	**<0.0001**
Tumour grade						
Grade 1	50 (38)	82 (62)	9.91	39 (24)	123 (76)	1.94
Grade 2	87 (30)	202 (70)	**0.007**	203 (28)	534 (72)	0.379
Grade 3	123 (25)	378 (75)		265 (29)	643 (71)	
Tubule formation						
1	12 (35)	22 (65)	1.14	N/A		
2	83 (27)	228 (73)	0.564			
3	151 (27)	402 (73)				
Mitotic count						
1	91 (35)	173 (65)	11.29	N/A		
2	53 (28)	135 (72)	**0.004**			
3	102 (23)	344 (77)				
Nuclear pleomorphism				N/A		
1	9 (50)	9 (50)	13.76			
2	103 (33)	208 (67)	**0.001**			
3	134 (24)	433 (76)				
Axillary nodal stage	173 (30)	399 (70)	12.81	308 (31)	681 (69)	12.17
Stage 1	57 (21)	216 (79)	**0.002**	162 (28)	428 (73)	**0.002**
Stage 2	30 (39)	47 (61)		64 (21)	241 (79)	
Stage 3						
Nottingham Prognostic Index						
Good	90 (36)	116 (64)	9.58	192 (30)	457 (70)	7.05
Moderate	26 (26)	382 (74)	**0.008**	305 (29)	744 (71)	**0.029**
Poor	36 (24)	163 (76)		39 (20)	154 (80)	
Vascular invasion						
Negative	183 (30)	418 (70)	4.094	NA		
Positive	77 (24)	242 (77)	**0.043**			
Molecular Subtype						
Luminal A	67 (34)	131 (66)	9.11	N/A		
Luminal B	66 (22)	235 (78)	**0.028**			
HER2	32 (25)	95 (75)				
Triple negative	44 (25)	134 (75)				
PAM50 Classification	N/A			184 (27)	505 (73)	51.35
Luminal A				154 (34)	305 (66)	**<**
Luminal B				107 (35)	201 (65)	**0.0001**
HER2				72 (31)	161 (69)	
Basal				17 (9)	179 (91)	
Normal like						

* Chi square test.

3.3.6 Correlations between ALDH1A1 expression and the expression of other BCSC markers

The correlations between ALDH1A1 mRNA and protein expression and the expression of other BCSC markers were assessed. The biomarkers selected for this analysis were based on previous investigations of putative BCSC markers and the availability of expression data for BCSC markers in the study cohorts *[266-271]*.

ALDH1A1 protein expression correlated positively with EpCAM (P = 0.007), SOX9 (P = 0.028) and SOX10 (P = 0.028) protein expression, but was not significantly correlated with CD24 or CD44 protein expression (P > 0.05; Table 3.3). Furthermore, when CD44 and CD24 expression were classified as CD44+/CD24+, CD44-/CD24-, CD44+/CD24- or CD44-/CD24+, no significant correlation was detected between ALDH1A1 protein expression and any category of CD44/CD24 expression (P = 0.481).

ALDH1A1 mRNA levels correlated with the mRNA levels of *CD44* (P = 0.014), *CD133* (P = 0.003), *EpCAM* (P = 0.006) and *SOX10* (P < 0.0001; Table 3.3).

In the limited number of cases with CN gain or loss, there was no significant correlation between *ALDH1A1* CN gain/loss and the expression of any of the tested BCSC markers (P > 0.05).

Table 3.3: Spearman's rank analysis of the correlations between ALDH1A1 protein and mRNA expression and the expression of other breast cancer stem cell (BCSC) markers.

BCSC marker	ALDH1A1 protein expression Correlation coefficient (*P*-value)	*ALDH1A1* mRNA expression Correlation coefficient (*P*-value)
CD133	0.082 (0.074)	0.067 (0.003)
CD44	0.000 (0.999)	0.055 (0.014)
CD24	0.001 (0.982)	0.044 (0.055)
EPCAM	0.089 (0.007)	0.062 (0.006)
SOX10	-0.088 (0.028)	0.117 (<0.0001)
SOX9	0.402 (0.028)	0.041 (0.071)

3.3.7 Associations between ALDH1A1 expression and patient outcomes

Survival analysis showed that high ALDH1A1 protein expression was associated with poor BCSS ($P < 0.0001$; Figure 3.5A). In contrast, high *ALDH1A1* mRNA expression was associated with better BCSS ($P < 0.0001$; Figure 3.5B).

When the molecular subtypes were assessed individually, high ALDH1A1 protein expression was associated with poor BCSS in TNBC ($P = 0.003$) and the luminal B subtype ($P = 0.042$), but not in luminal A and HER2+ tumours ($P > 0.05$; Figures 3.6A–D). However, no significant associations were identified between *ALDH1A1* mRNA expression and BCSS in any molecular subtype of BC (all $P \geq 0.05$; Figure 3.7).

ALDH1A1 protein and mRNA expression did not exhibit any significant associations with DMFS or DFI ($P > 0.05$; data not shown).

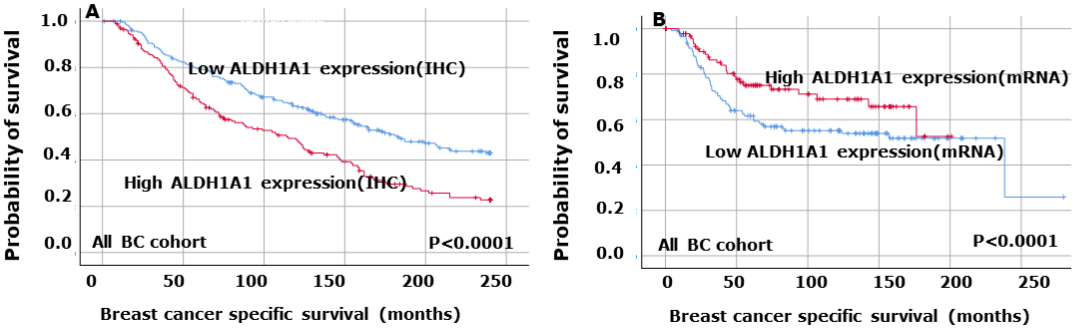


Fig. 3.5: Kaplan–Meier survival curves of breast-cancer specific survival for the entire Nottingham cohort stratified by ALDH1A1 protein expression (A) and the entire METABRIC cohort stratified by ALDH1A1 mRNA expression (B).

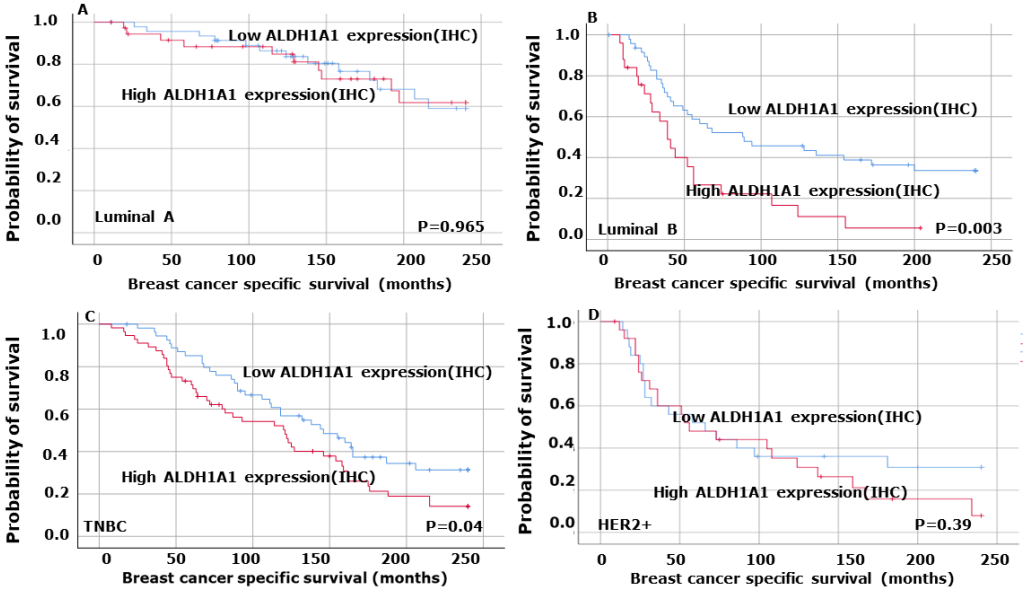


Fig. 3.6: Kaplan–Meier survival curves of breast-cancer specific survival for patients with different molecular subtypes of BC stratified by ALDH1A1 protein expression.

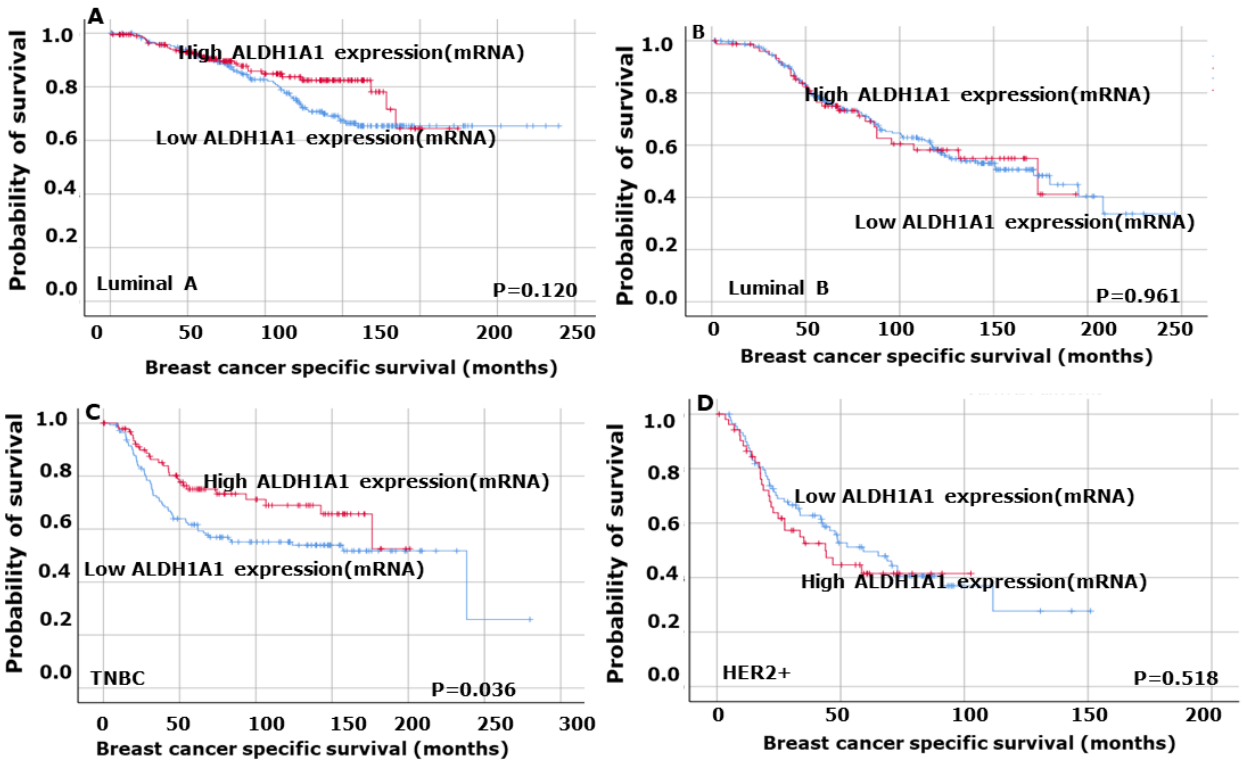


Fig. 3.7: Kaplan–Meier survival curves of breast-cancer specific survival for patients with different molecular subtypes of BC stratified by *ALDH1A1* mRNA expression.

3.3.8 Multivariate analysis

Cox-proportional hazards models incorporating a variety of other prognostic parameters (tumour size, tumour grade and nodal stage) revealed that high ALDH1A1 protein expression was an independent predictor of poor BCSS (P = 0.001, HR = 1.52, 95% CI: 1.180–1.968) and high *ALDH1A1* mRNA expression was an independent predictor of good BCSS (P < 0.0001, HR = 0.640, 95% CI: 0.527–0.777; Table 5). However, the multivariate model showed that ALDH1A1 protein expression was not an independent predictor for poor BCSS in the luminal B and TNBC subtypes (Table 6).

Table 3.4: Multivariate Cox regression hazard model of the associations between ALDH1A1 protein and mRNA levels and other prognostic clinicopathological features and breast cancer-specific survival in the entire breast cancer cohorts.

	ALDH1A1 protein expression			***ALDH1A1* mRNA expression**		
Parameter	*P*-value	Hazard ratio	95% CI	*P*-value	Hazard ratio	95% CI
ALDH1A1	**0.001**	1.524	1.180–1.968	< 0.0001	0.640	0.527–0.777
Tumour size	0.008	1.395	1.089–1.785	< 0.0001	1.516	1.203–1.910
Tumour grade	< 0.0001	1.846	1.529–2.273	< 0.0001	1.364	1.163–1.599
Axillary nodal stage	< 0.0001	1.790	1.531–2.132	< 0.0001	1.92	1.703–2.174

Table 3.5: Multivariate Cox regression hazard model of the associations between ALDH1A1 protein levels and other prognostic clinicopathological features and breast cancer-specific survival in the luminal B and triple negative (TN) breast cancer subtypes.

	Luminal B			**TNBC**		
Parameter	***P*-value**	**Hazard ratio**	**95% CI**	***P*-value**	**Hazard ratio**	**95% CI**
ALDH1A1	0.075	1.513	0.959–2.387	0.133	1.677	0.855–3.289
Tumour size	0.435	1.210	0.750–1.950	0.732	0.901	0.498–1.633
Tumour grade	0.022	1.609	1.070–2.420	0.398	1.513	0.579–3.951
Axillary nodal stage	0.114	1.295	0.940–1.785	0.063	1.470	0.980–2.204

3.4 Discussion

ALDH1A1 has been reported as a specific marker for identification of the BCSC phenotype and as a marker for CSCs in other solid cancers, including lung and colorectal cancer [260] [272] [273]. Several studies reported that ALDH1A1 plays a major role in the progression of several solid tumours and is associated with poor outcomes [274-277]. However, Chang *et al.* (2009) found that ALDH1A1 was associated with improved patient outcome in ovarian cancer, which indicates tumours is occurring in different locations might present different molecular expression patterns, and ALDH1A1 may play a different role in ovarian cancer than it does in other cancers [278, 279].

In the current study, ALDH1A1 expression was assessed at the genomic, transcriptomic and proteomic levels in large cohorts of IBC patients to explore the prognostic significance of ALDH1A1 and assess the correlation between ALDH1A1 expression and the expression of other BCSC markers.

High ALDH1A1 protein (cytoplasmic expression) and mRNA expression were observed in 71% of the IBC tumours overall. Other studies reported much lower frequencies (~50%) of high ALDH1A1 expression [280-282]. However, the results of this study are in agreement with Yan *et al.* (2018), who detected high ALDH1A1 in 60% of their BC cohort based on H-scores.

Such variability between studies may be the result of different cut-offs and definitions of ALDH1A1 positivity. This study investigated the prognostic significance of ALDH1A1 in two large cohorts of patients with BC and derived cut-off values based on the prediction of patient survival, which may explain some of the differences to other studies. Additionally, the lack of concordance with other studies could partly be explained by the use of different scoring methods. H-

scoring was used in this study, which is a widely accepted system in both clinical and research settings, whereas previous studies used other scoring systems [283-285].

CSCs represent a small percentage (0.05–1%) of the tumour cell population [286, 287]. However, diffuse ALDH1A1 immunostaining may not accurately predict the BCSC population when employed as an individual tissue marker. However, IHC analysis using a combination of more than one BCSC marker may more precisely identify CSCs. For example, Neumeister *et al.* (2010) performed multiplex immunofluorescent staining of BC tissues and reported that combined analysis of ALDH1A1 and CD44 could predict patient survival more accurately than either marker alone [288]. In this study, strong ALDH1A1 immunoreactivity was observed in a minority of the tumour cells, which may indicate that the ALDH1A1-positive subpopulation represents BCSCs. A previous study also showed low levels of strong ALDH1A1 expression with focal positivity in BC, which is consistent with the concept that CSCs represent a small subgroup of tumour cells [289]. In their review, Giuseppe (2019) stated that ALDH1A1 can be expressed by true stem cells, transit amplifying progenitor cells, differentiating cells, progenitor cells and even mature cells [255].

Identification of CSCs is commonly accomplished using combinations of several biomarkers, as there is no universal marker or consensus CSC marker for all cancer types, and some molecules are frequently shared across CSCs in various tumour types. Al-Hajj *et al.* (2003) identified that BCSCs expressing the cell surface markers CD44+/CD24- drive tumour growth [290]. Furthermore, Ahmed *et al.* [291] showed that expression of CD44-/CD24+ in BC tissue samples was significantly associated with more aggressive behaviour and poorer

patient outcomes. This study indicates significant correlations exist between *ALDH1A1* mRNA expression and *CD44+*, *EPCAM, SOX10* and *CD133* mRNA expression. Furthermore, ALDH1A1 protein expression correlated positively with the protein expression of other BCSC markers, including SOX9 and EpCAM.

CD133/prominin 1 is a promising BCSC candidate marker associated with the characteristics of self-renewal, high proliferation and drug resistance [292]. Recently, an assessment of CD133 expression in BC concluded that overexpression of CD133 is likely to be a stemness feature and prognostic marker in IBC [229]. EpCAM is well-established as a BCSC and epithelial–mesenchymal transition (EMT) marker that can maintain self-renewal, the pluripotent phenotype and the EMT in cancer progression [293, 294]. Domenic *et al.* [295] reported that SOX9 expression may maintain human breast luminal progenitor cells and be associated with more aggressive features in BC [296]. Thus, this study indicates that ALDH1A1 may have potential as both a CSC marker and prognostic marker in BC.

High ALDH1A1 protein expression was more frequent in highly proliferative (luminal B and TNBC) tumours, which suggests that ALDH1A1 may contribute to cell proliferation and tumorigenesis in more aggressive subtypes of BC [297]. Furthermore, in agreement with another study [298], ALDH1A1 was positively associated with the proliferation marker Ki67 in this study.

ALDH1A1 mainly catalyses the conversion of retinaldehyde to retinoic acid, which subsequently binds and activates the retinoic acid receptor or retinoid X receptor in the nucleus to promote target gene expression. The downstream genes of the retinoic acid pathway are also involved in tumour cell differentiation and proliferation [247]. This study demonstrates that high

expression of ALDH1A1 at both the protein and mRNA levels are positively associated with aggressive tumour features in IBC. Fei *et al.* (2017) concluded that ALDH1A1 was associated with larger tumour size and advanced stage in TNBC [289]. Another study reported that high expression of ALDH1A1 was associated with larger tumour size, high grade and advanced tumour stage in BC, in agreement with the current study [299]. Importantly, high ALDH1A1 protein expression was an independent prognostic marker for poor BCSS in the entire BC cohort, thus ALDH1A1 may have clinical relevance to improve survival rate prediction.

Furthermore, high ALDH1A1 was significantly associated with poorer BCSS in the luminal B and TNBC subtypes, which further indicates the heterogeneity of BC with respect to stemness markers. TNBC has been reported to contain a higher percentage of cells that possess a stem-like phenotype [300]. However, ALDH1A1 was not an independent prognostic factor for BCSS in the multivariate analysis, which indicates ALDH1A1 expression may depend on other factors. For example, Mieog *et al*. (2012) reported that the prognostic value of ALDH1A1 expression is age-dependent and was only observed in patients aged < 65-years-old [299, 301].

Contradictory associations between ALDH1A1 protein and mRNA expression and patient outcomes were observed in this study, similarly to other studies [302]. In this study, high *ALDH1A1* mRNA expression was associated with improved BCSS. Another study indicated that high ALDH1A1 mRNA expression was associated with good outcomes in TNBC [302]. Sjöström *et al.* (2015) reported that high stromal ALDH1A1 expression was associated with better distant disease-free survival in BC [303]. One possible underlying cause

of the discrepancies between the survival outcomes and ALDH1A1 mRNA and protein expression is that the stromal components are not removed during tissue processing for RNA extraction. ALDH1A1 may also be expressed in cytotoxic T cells, B cells or dendritic cells, which have been shown to suppress tumour progression [304]. However, the IHC analysis in this study only considered ALDH1A1 protein expression in invasive tumour cells.

Moreover, there was no significant correlation between ALDH1A1 protein and mRNA expression in the subset of 184 cases present in both cohorts, which could also be attributed to interference by stromal components.

Study suggested that ALDH1 expressed in different subtypes of cells can involve in varied molecular pathways, and has been speculated that ALDH1 positive stromal cells might produce retinoic acid into the tumour microenvironment. In turn, stromal cell-derived retinoic acid might induce tumour cell differentiation, inhibit proliferation and migratory abilities of tumour cells directly or via regulation of other immune cells, which, in consequence, decreases the aggressiveness of a tumour and reduces its progression [304]. Future experimental studies are required to identify precise mechanisms behind this putative crosstalk between tumour and stroma cells in BC. Future identification of ALDH1A1 in stromal cells in the full face section, as stromal expression is not suitable for assessment in TMA, is recommended.

3.5 Conclusion

This chapter demonstrates and confirms that high ALDH1A1 protein expression is associated with advanced prognostic features and poorer BCSS in BC, particularly in the luminal B and TNBC subtypes. High expression of ALDH1A1 was also associated with high expression of some of BCSC markers. Therefore, combined analysis of the expression of ALDH1A1 and other BCSC markers may provide a method to more accurately identify the BCSC subpopulation in BC tissue samples.

Chapter 4

The prognostic significance of BMI-1 expression varies in different subtypes of invasive breast cancer

4.1 Introduction

B-lymphoma Moloney murine leukaemia virus insertion region-1 *(*BMI-1) is a member of the polycomb complex protein family of transcriptional repressors [305, 306]. The *BMI-1* gene is located on the short arm of chromosome 10 (10p11.23) and extends over 10 kb. The highly-conserved protein is composed of 326 amino acids [307] and contains a RING finger at the N-terminus, a central helix-turn-helix domain and a carboxyl-terminal PEST-like domain. BMI-1 is involved in the DNA damage response. The RING domain of BMI-1 locates DNA strand breaks. The central helix-turn-helix domain and RING domain are associated with the life span of the cell. The PEST domain is required for protein degradation of BMI-1 [308].

High BMI-1 expression has been reported in several types of cancer, including gastric, head and neck, pancreatic, ovarian, breast and lung tumours, hepatocellular carcinoma and endometrial carcinoma [309-316]. High levels of BMI-1 expression have been observed in human cancer tissues compared with the matched normal tissues [317-321]. BMI-1 has been reported to contribute to axillary lymph node metastases, highly aggressive tumour behaviour and advanced-stage relapse in BC [322, 323]. Overexpression of BMI-1 leads to increased cell motility and invasion in *in vitro* models of BC, [312, 324].

BMI-1 plays important roles in BC and BCSCs through activation of numerous signalling pathways. BMI-1 upregulates the canonical Wnt pathway and overexpression of BMI-1 leads to upregulation of cyclin D1 and c-Myc [325, 326]. Moreover, overexpression of BMI-1 appears to downregulate Wnt inhibitors, such as the DKK family [327]. SALL4, a transcriptional regulator of the stem cell phenotype, has been reported to induce histone methylation (H3-K4 and H3-K79) on the *BMI-1* gene promoter [328, 329]. Zhu *et al* (2013) observed knockdown of nuclear factor erythroid-2-related factor 2 (Nrf2a) reduced the expression of BMI-1 and inhibited the proliferation of glioma stem cells [330]. Furthermore, BMI-1 has been shown to regulate cancer cell invasion, metastasis and chemosensitivity through the VEGF, PTEN and PI3K/AKT signalling pathways [315, 331-333]. Hence, BMI-1 is a key regulator of the self-renewal, differentiation and tumour initiation potential of BCSCs [334].

However, Arnes *et al.* (2010) reported that low expression of BMI-1 was associated with high expression of ALDH1, a functional marker of the BCSC phenotype, in African patients with BC [335]. Interestingly, expression of BMI-1 is upregulated by oestrogen in the tumours of ER-positive BC patients [336]. More recently, BMI-1 has been suggested as a prognostic and predictive biomarker of poorer outcomes in several types of cancer [305, 337, 338]. However, conflicting data were reported in the same study, as high expression of BMI-1 at the mRNA and protein levels were associated with high expression of the ER and positive axillary lymph node metastasis [339].

Although BMI-1 has been established to play a role in maintenance of the properties of CSCs in various types of cancer, there is still little evidence on the

prognostic value of BMI-1 in different subtypes of IBC and the associations of BMI-1 with BCSCs need to be defined in different subtypes of BC.

4.1.1 Hypotheses

The hypotheses tested in this chapter are:

(i) BMI-1 has varied prognostic value in different subtypes of BC.

(ii) BMI-1 has varied correlations with BCSC markers in different subtypes of BC.

4.1.2 Aims

The aims of this study are to:

(i) Investigate the prognostic value of BMI-1 expression in a large cohort of BC patients with long-term follow-up.

(ii) Explore the correlations between BMI-1 expression and the expression of BCSC markers in different BC subtypes.

4.1.3 Objectives

The objectives of this chapter are to:

(i) Assess BMI-1 mRNA and protein expression in different subtypes of BC.

(ii) Examine the correlations between BMI-1 expression and the expression of other BCSC markers at the mRNA and protein levels in BC.

(iii) Investigate the associations between BMI-1 expression and the clinicopathological features and patient outcomes in different subtypes of BC.

4.2 Materials and Methods

4.2.1 Study cohort for analysis of BMI-1 expression

The study cohort for evaluation of BMI-1 protein expression contained 870 patients with invasive BC derived from a retrospective series of patients with primary breast carcinoma who presented at Nottingham City Hospital between 1988 and 1998 (see Chapter 2).

In this cohort, 65% of the patients were less than 50-years-old, 15% and 52% had grade 1 and grade 3 tumours, and 28% and 72% had ER- and ER+ tumours, respectively.

BMI-1 mRNA expression and gene copy number (CN) aberrations were assessed in the METABRIC cohort of 1,980 BC patients. The clinicopathological parameters of both the Nottingham and METABRIC study cohorts are summarised in Tables 4.6 and 4.7.

4.2.2 Availability of data on biomarker expression

Immunohistochemical detection of a panel of BCSC markers, including ALDH1A1, CD133, CD24, CD44, EPCAM, SOX9 and SOX10, had been performed previously for the Nottingham cohort, Details of these biomarkers are shown in Chapter 2.

4.2.3 Antibody validation

The specificity of the anti-BMI-1 antibody was validated by western blotting of MCF7, MDA-MB-231, SKBR3 and MDA-MB-468 BC cell lysates and HeLa cell lysates using a 1:5000 dilution of the primary antibody (EPR3745, Abcam, UK).

4.2.4 IHC of BMI-1 protein expression

IHC staining was performed on 4 µm TMA sections using a Novolink polymer detection system (Leica, Newcastle, UK), as described in Chapter 2. The optimal dilution of the BMI-1 antibody for IHC analysis was 1:100 with 1 h incubation at room temperature. Normal kidney tissues were used as a positive control, while the negative control was obtained by omitting the primary antibody.

4.2.5 Assessment of BMI-1 protein expression

BMI-1 IHC staining was semi-quantitatively evaluated using the modified histochemical score (H-score). Inter-observer reproducibility was tested by randomly selecting 10% of the cases for double scoring.

4.2.6 Statistical analysis

BMI-1 protein and mRNA expression were categorised using X-tile bioinformatics software version 3.6.1 (School of Medicine, Yale University, New Haven, USA) based on prediction of BCSS.

Data were analysed as described in section in Chapter 2 using IBM SPSS 24.0 software (Chicago, IL, USA).

4.3 Results

4.3.1 Validation of the BMI-1 antibody

The specificity of the anti-BMI-1 antibody was validated using western blotting (Figure 1). A single band was observed at the predicted size of 40 kDa for BMI-1, which confirmed the specificity of the antibody.

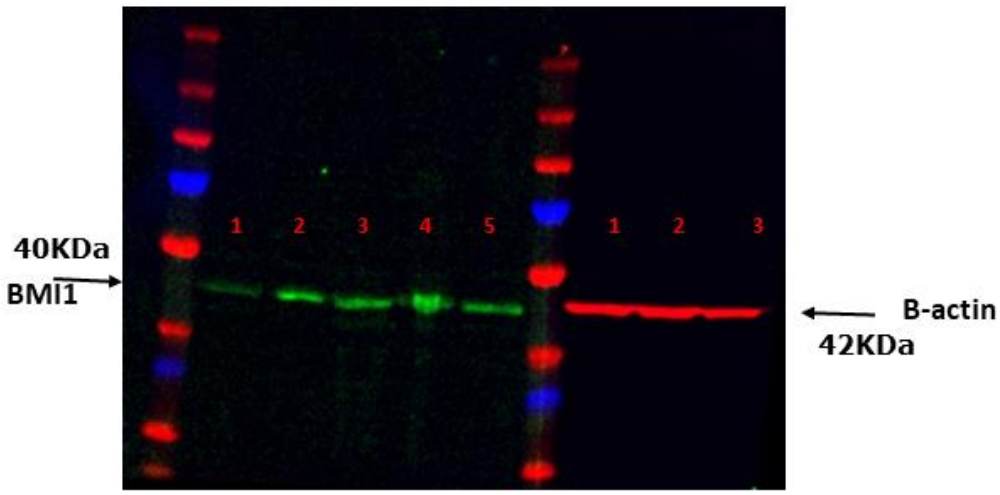


Figure 4.1: Western blot to confirm the specificity of the rabbit monoclonal anti-BMI-1 antibody. A single specific band (green) was observed at the expected molecular weight for BMI-1 (40 kDa) in 1) MCF7; 2) MDA-MB-231; 3) SKBR3; 4) MDA-MB-468; and 5) HeLa cell lysates. The red bands are β-actin (positive control) at 42 kDa.

4.3.2 BMI-1 protein expression

IHC revealed a homogenous pattern of staining, with BMI-1 protein expression localised in the nuclei of the invasive tumour cells (Figure 2B). The staining intensities varied from negative (no stain) to strong (Figures 2A and 2B). The interclass correlation co-efficient was 0.931, indicating excellent inter-observer agreement between the two scorers.

In the entire Nottingham cohort, the H-score for BMI-1 protein expression ranged from 0 to 270. A high/positive BMI-1 protein level (H-score > 130) was observed in 176/870 (20%) cases, whereas 694/870 (80%) cases were classified as low/negative BMI-1 protein expression.

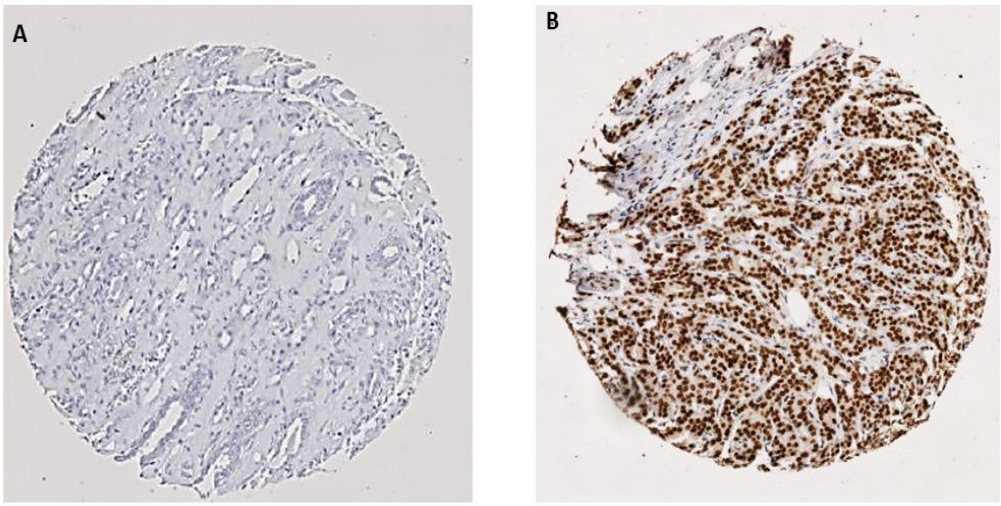


Fig 4.2: Representative photomicrographs of BMI-1 protein expression in invasive BC. (A) Negative immunohistochemical (IHC) staining. (B) Positive IHC staining for BMI-1.

4.3.3 BMI-1 mRNA expression

In the METABRIC cohort, high *BMI-1* mRNA expression (9.2 log fold-change) was observed in 1288/1980 (65%) cases.

BMI-1 CN gains were observed in 113/1980 (6%) of cases and 25/1980 (1.3%) of cases exhibited CN loss.

Among the cases included in both the Nottingham cohort and METABRIC cohort (n = 310), a strong correlation was observed between *BMI-1* mRNA expression and protein expression (P = 0.001).

4.3.4. BMI-1 expression in BC molecular subtypes

ER+ tumours expressed high levels of BMI-1 (mean = 94 H score) compared to ER- tumours (mean = 48 H score). Table 4.3 summaries BMI-1 protein and mRNA expression in various BC subtypes.

BMI-1 protein expression was higher in the luminal ER+ subtype than the ER- subtypes (P < 0.0001). High *BMI-1* mRNA expression was observed in 73% of tumours of the ER+ subtype and 60% of tumours of the ER- subtype.

Table 4.3: BMI-1 protein expression H-scores and mRNA log fold changes in different molecular subtypes of breast cancer.

	Protein (H-score)			mRNA (log fold-change)		
	Entire cohort	**Luminal ER+ BC subtype**	**TNBC subtype**	**Entire cohort**	**Luminal ER+ BC subtype**	**Basal ER- BC subtype**
Mean	81.73	93.91	47.80	9.5	9.6	9.1
Median	80.00	90.00	30.00	9.4	9.5	9.0
Range	0–270	0–270	0–240	7.6–12	7.8–12	7.6–11.5

4.3.5 Correlations between expression of BMI-1 and expression other biomarkers at the protein and mRNA levels

The correlations between BMI-1 protein and mRNA expression and the expression of other BCSC biomarkers were investigated in the entire Nottingham and METABRIC cohorts and in varied subtypes of BC.

In the entire Nottingham cohort, significant negative correlations were observed between BMI-1 protein expression and the protein levels of the BCSC markers ALDH1A1 (P = 0.017), CD133 (P = 0.006), CD24 (P = 0.023) and SOX9 (P = 0.004). There were no significant correlations between ALDH1A1 protein expression and EpCAM or CD44 protein expression ($P \geq 0.05$; Table 4.4).

In ER+ tumours, BMI-1 protein expression correlated negatively with ALDH1A1 (P = 0.005) and CD133 (p = 0.001) protein expression. Conversely, in ER- tumours, BMI-1 protein expression correlated positively with CD133 ($p < 0.0001$) and SOX10 protein expression (P = 0.01; Table 4.4).

In the entire METABRIC BC cohort, *BMI-1* mRNA expression correlated negatively with the mRNA expression levels of the BCSC markers *ALDH1A3* ($P <$

0.0001) *CD133* (*P* = 0.0003), *CD24* (*P* = 0.002) and *CD44* (*P* = 0.0007; Table 4.5).

When the analysis was restricted to ER+ tumours, high *BMI-1* mRNA expression correlated negatively with the mRNA expression of *ALDH1A1* (*P* = 0.025), *ALDH1A3* (*P* = 0.026), *CD133* (*P* = 0.038), *CD44* (*P* = 0.0001), *CD24* (*P* = 0.004) and *SOX10* (*P* = 0.0006; Table 4.5). However, in the TNBC subtype, positive correlations were observed between *BMI-1* mRNA expression and *EPCAM* and *SOX10* mRNA expression (both *P* < 0.001).

Table 4.4: Correlations between BMI-1 protein expression and the protein expression of other breast cancer stem cell markers in the entire Nottingham cohort and luminal ER+ subtype.

BCSC Marker	BMI-1 Protein (Entire cohort) Correlation coefficient (*P*-value)	BMI-1 Protein (Luminal ER+ BC) Correlation coefficient (*P*-value)
ALDH1A1	-0.101 (0.017)	-0.150 (0.005)
CD133	-0.123 (0.006)	-0.171 (0.001)
CD44	-0.046 (0.321)	-0.004 (0.766)
CD24	-0.090 (0.023)	-0.076 (0.057)
EPCAM	-0.051 (0.133)	0.052 (0.191)
SOX10	0.235 (< 0.0001)	0.155 (< 0.0001)
SOX9	-0.098 (0.004)	-0.012 (0.766)

Table 4.5: Correlations between *BMI-1* mRNA expression and the mRNA expression of other breast cancer stem cell markers in the entire METABRIC cohort and luminal ER+ BC.

BCSC marker	***BMI-1* mRNA (Entire cohort) Correlation coefficient (*P*-value)**	***BMI-1* mRNA (Luminal ER+ BC) Correlation coefficient (*P*-value)**
ALDH1A1	-0.033 (0.138)	-0.058 (0.025)
ALDH1A3	-0.181 (<0.0001)	-0.058 (0.026)
CD133	-0.205 (<0.0003)	-0.054 (0.038)
CD44	-0.103 (<0.0007)	-0.115 (0.0001)
CD24	-0.232 (0.002)	-0.195 (0.0004)
EPCAM	0.073 (0.001)	0.123 (0.0008)
SOX10	-0.257 (0.0001)	-0.119 (0.0006)
SOX9	-0.011 (0.616)	-0.090 (0.0007)

4.3.6 Associations between BMI-1 expression and the clinicopathological features of BC

High BMI-1 protein expression was associated with various clinicopathological features characteristic of a good prognosis, such as lower histological grade ($P < 0.0001$), increased tubule formation ($P = 0.004$), lower mitotic count ($P < 0.0001$), lower nuclear pleomorphism ($P < 0.0001$), low NPI ($P = 0.024$), ER+ status ($P < 0.0001$) and HER2-negative status ($P = 0.002$). There were no significant associations between BMI-1 protein levels and other factors such as tumour stage, LVI or tumour size ($P \geq 0.05$; Table 4.6).

Similar findings were observed in the METABRIC cohort. High *BMI-1* mRNA expression was positively associated with older age ($P < 0.001$), postmenopausal status ($P < 0.001$), lower grade tumours ($P < 0.0001$), good NPI score ($P < 0.0001$) and invasive tubular carcinoma ($P < 0.0001$; Table 4.7).

Table 4.6: Associations between BMI-1 protein expression and the clinicopathological features of breast cancer.

	Protein expression (*n* = 870)			
Clinicopathological parameters	**No. (%)**	**Low BMI-1 No. (%)**	**High BMI-1 No. (%)**	**x^2 *P*-value**
Patient age				
< 50 y	562 (65)	451 (80)	111 (20)	0.22
≥ 50 y	308 (35)	243 (79)	65 (21)	0.635
Tumour size				
< 2 cm	452 (52)	369 (82)	83 (18)	1.77
≥ 2 cm	409 (48)	319 (78)	90 (22)	0.106
Tumour grade				
Grade 1	130 (15)	90 (69)	40 (31)	22.33
Grade 2	280 (33)	211 (75)	69 (25)	**< 0.0001**
Grade 3	449 (52)	385 (86)	64 (14)	
Tubule formation				
1	39 (5)	26 (67)	13 (33)	11.02
2	289 (34)	219 (76)	70 (24)	**0.004**
3	504 (59)	420 (83)	84 (17)	
Mitotic count				
1	255 (31)	187 (73)	68 (27)	14.99
2	170 (20)	131 (77)	39 (23)	**< 0.0001**
3	407 (49)	347 (85)	60 (15)	
Nuclear pleomorphism	14 (2)	11 (78.6)	3 (21)	18.76
1	309 (37)	223 (72.2)	86 (28)	**< 0.0001**
2	509 (61)	431 (84.7)	78 (15)	
3				
Lymph node stage				
Stage 1	529 (62)	426 (80)	103 (20)	0.65
Stage 2	271 (32)	214 (79)	57 (21)	0.721
Stage 3	60 (6)	46 (77)	14 (23)	
Nottingham Prognostic Index				
Poor	142 (16)	119 (83.8)	23 (16.2)	7.46
Moderate	476 (55)	389 (81.7)	87 (18.3)	**0.024**
Good	243 (29)	180 (74.1)	63 (25.9)	
Oestrogen receptor status	233 (27)	213 (91)	20 (9)	26.12
Negative	626 (73)	474 (76)	152 (24)	**< 0.0001**
Positive				
Progesterone receptor status				14.48
Negative	357 (43)	309 (87)	48 (13)	**< 0.0001**
Positive	475 (57)	361 (76)	114 (24)	
HER2 status				
Negative	708 (86)	557 (79)	151 (21)	8.46
Positive	114 (14)	103 (90)	11 (10)	**0.002**

Table 4.7: Associations between *BMI-1* mRNA expression and the clinicopathological features of breast cancer.

	mRNA Expression (*n* = 1980)			
Clinicopathological parameters	**No. (%)**	**Low *BMI-1* No. (%)**	**High *BMI-1* No. (%)**	**x^2 *P*-value**
Patient age				
< 50 years	396 (21)	147 (37)	249 (63)	11.809
≥ 50 years	1458 (79)	411 (28)	1047 (72)	**0.001**
Tumour size				
< 2 cm	622 (32)	193 (31)	429 (69)	4.218
≥ 2 cm	1331 (68)	476 (36)	855 (64)	0.040
Tumour grade				
Grade 1	163 (9)	20 (12)	143 (88)	91.336
Grade 2	729 (41)	162 (22)	567 (78)	**< 0.0001**
Grade 3	880 (50)	357 (41)	523 (59)	
Axillary nodal stage				
Stage I				
Stage II	964 (49)	281 (29)	683 (71)	3.418
Stage III	443 (23)	29 (29)	414 (71)	0.181
	544 (28)	345 (34)	199 (66)	
Nottingham Prognostic Index				
Poor	643 (35)	135 (21)	508 (79)	42.356
Moderate	1022 (55)	346 (34)	676 (66)	**< 0.0001**
Good	189 (10)	77 (41)	112 (59)	
Oestrogen receptor status				
Negative	401 (22)	224 (56)	177 (44)	165.226
Positive	1414 (78)	319 (23)	1095 (77)	**< 0.0001**
Progesterone receptor status				
Negative	870 (47)	368 (42)	502 (58)	116.00
Positive	984 (53)	190 (19)	794 (81)	**< 0.0001**
HER2 status				
Negative	1624 (88)	467 (29)	1157 (71)	10.616
Positive	229 (12)	90 (39)	139 (6)	**0.001**

4.3.7 Associations between BMI-1 protein expression and patient outcomes

Univariate survival analysis showed that high BMI-1 protein expression was significantly associated with longer BCSS ($P = 0.02$) and distant metastasis-free survival (DMFS; $P = 0.04$; Figure 4.3). In ER+ tumours, high BMI-1 protein expression was significantly associated with longer BCSS ($P = 0.04$), but not with DMFS (Figure 4.4). However, high BMI-1 protein expression was significantly associated with shorter BCSS in TNBC ($P = 0.04$; Figure 4.5). No significant associations were observed between BMI-1 and DFI in the entire Nottingham cohort or any subtype of BC ($P > 0.05$).

High *BMI-1* mRNA expression was significantly associated with longer BCSS in the entire METABRIC cohort ($P < 0.0001$) and in patients with luminal ER+ tumours ($P < 0.0001$), but not in patients with ER- tumours ($P > 0.05$; Figure 4.6). Moreover, *BMI-1* CN gain was associated with longer BCSS in the entire METABRIC cohort ($P = 0.031$; Figure 4.7), while no significant associated of CN loss with BCSS ($P>0.05$; Figure 4.7).

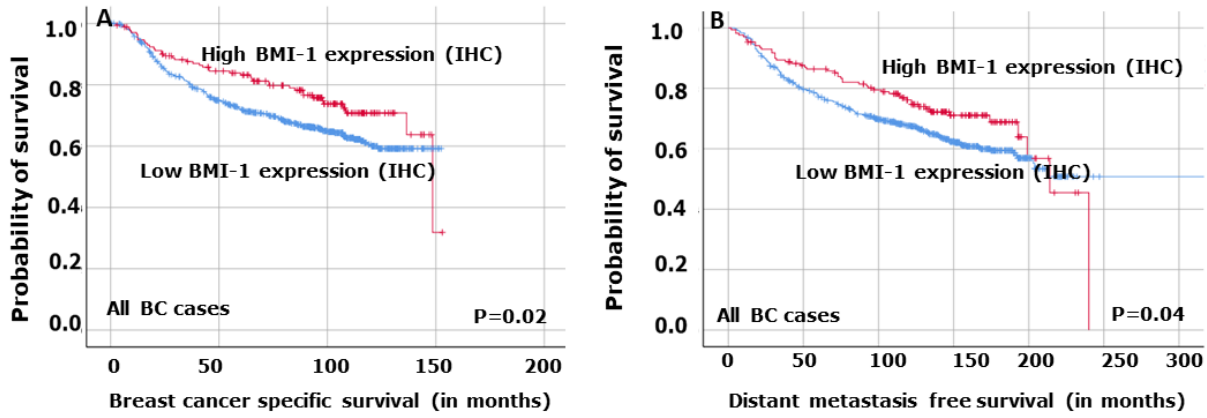


Figure 4.3: Kaplan-Meier plots of (A) breast cancer-specific survival (BCSS) and (B) distant metastasis-free survival (DMFS) for the entire Nottingham cohort of patients with breast cancer (BC) stratified by BMI-1 protein expression.

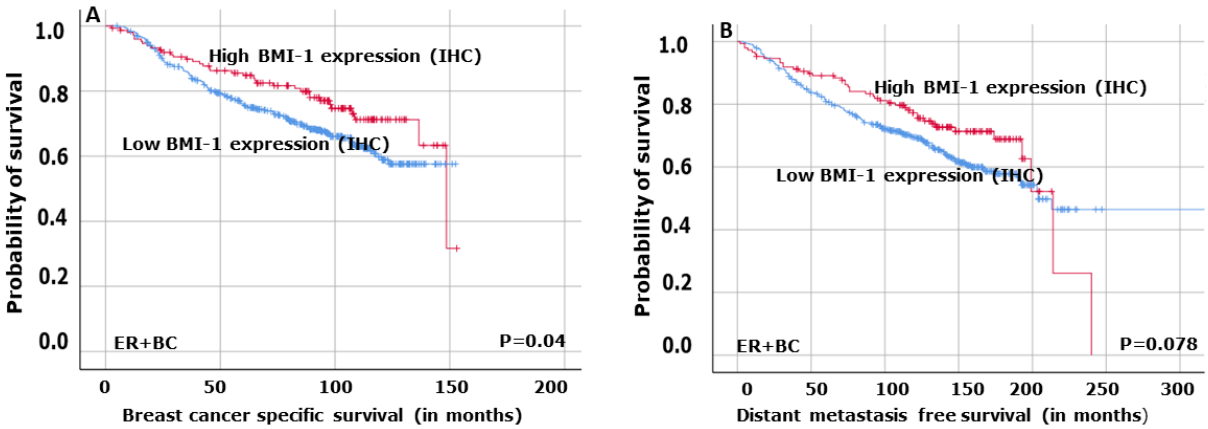


Figure 4.4: Kaplan-Meier plots of (A) breast cancer-specific survival (BCSS) and (B) distant metastasis-free survival (DMFS) for the subgroup of patients with oestrogen receptor positive (ER+) breast cancer (BC) stratified by BMI-1 protein expression.

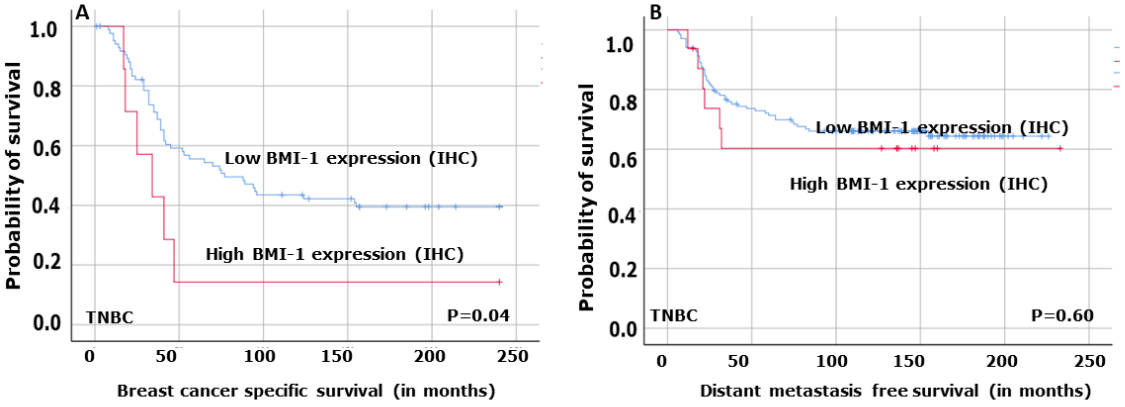


Figure 4.5: Kaplan-Meier plots of (A) breast cancer-specific survival (BCSS) and (B) distant metastasis-free survival (DMFS) for the subgroup of patients with TNBC stratified by BMI-1 protein expression.

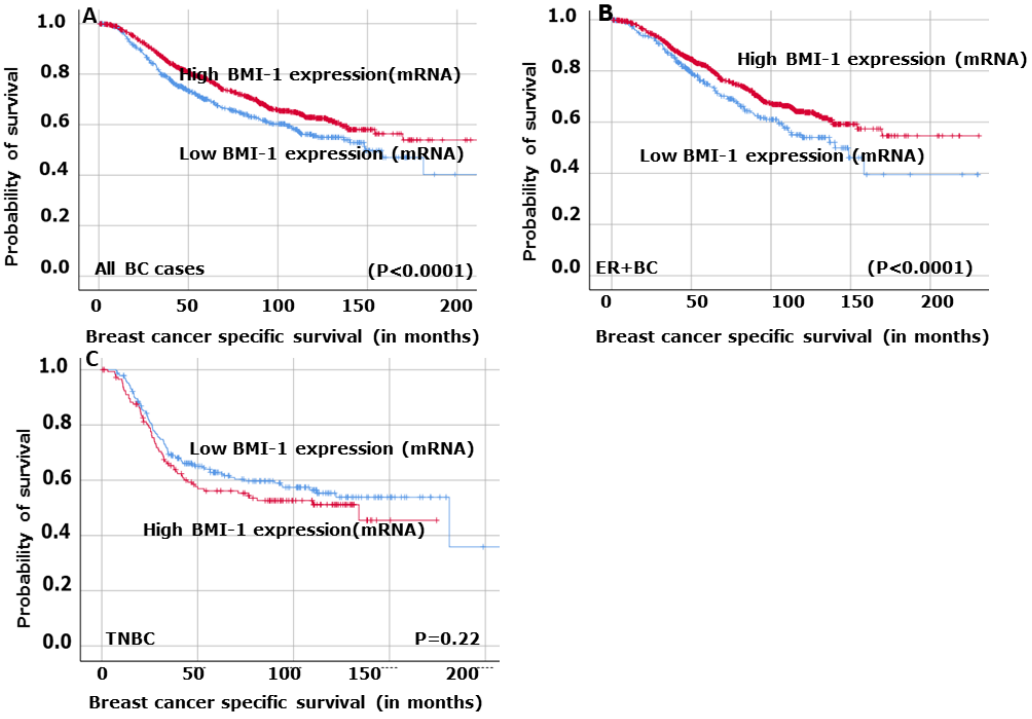


Figure 4.6: Kaplan-Meier plots of breast cancer-specific survival (BCSS) for (A) the entire METABRIC breast cancer (BC) cohort and subgroups with (B) oestrogen receptor positive (ER+) tumours and (C) triple negative breast cancer (TNBC) stratified by *BMI-1* mRNA expression.

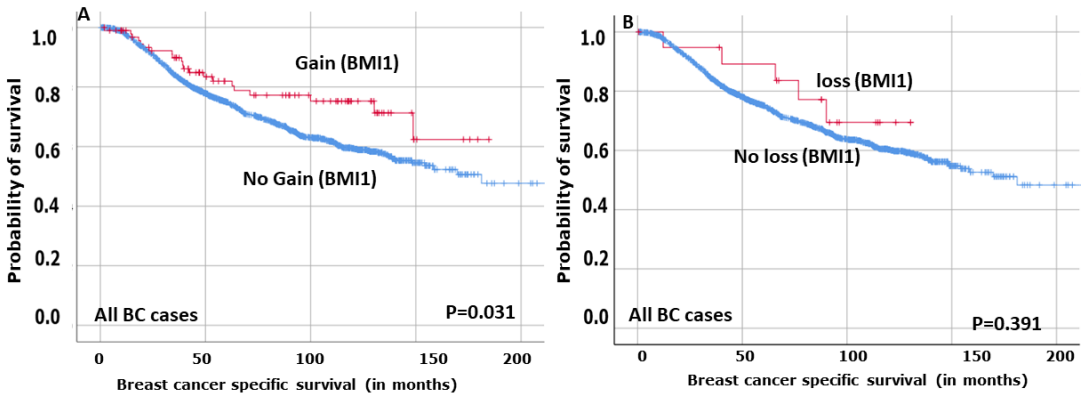


Figure 4.7: Kaplan-Meier plots of breast cancer-specific survival (BCSS) for the entire METABRIC breast cancer (BC) cohort stratified by (A) BMI-1 copy number gain and (B) BMI-1 copy number loss.

4.3.8 Multivariate analysis

The Cox regression model including tumour size, tumour grade and nodal stage revealed that BMI-1 was an independent predictor of better BCSS for the entire Nottingham BC cohort (*P* = 0.04, HR = 0.72, 95% CI: 0.51–0.99) and in the subgroup of patients with luminal ER+ tumours (*P* = 0.04, HR = 0.69, 95% CI: 0.48–0.99; Table 4.8), but not in TNBC (Table 4.9).

Multivariate analysis including age at diagnosis, tumour size, tumour grade and nodal stage demonstrated that *BMI-1* mRNA expression was an independent prognostic factor, but only in ER+ tumours (*P* = 0.045; HR = 0.852; 95% CI: 0.724–1.003; Table 4.10).

Table 4.8: Multivariate Cox regression hazard model of prognostic clinicopathological parameters for the entire Nottingham BC cohort and the subgroup of patients with ER+ tumours

	Entire BC cohort				**ER+ tumours**			
Parameters	**Hazard ratio**	**95% CI**		***P*-value**	**Hazard ratio**	**95% CI**		***P*-value**
		Lower	**Upper**			**Lower**	**Upper**	
Tumour Grade	1.61	1.32	1.96	< 0.0001	1.70	1.36	2.13	< 0.0001
Nodal Stage	1.99	1.67	2.38	< 0.0001	2.03	1.62	2.56	< 0.0001
Tumour Size	0.67	0.52	0.87	0.002	0.63	0.46	0.86	0.004
BMI-1 Protein Expression	0.72	0.51	0.99	**0.048**	0.69	0.48	0.99	**0.047**

Table 4.9: Multivariate Cox regression hazard model of prognostic clinicopathological features in TNBC.

Parameters	Hazard ratio	95% CI		*P*-value
		Lower	Upper	
Tumour grade	2.621	0.734	9.353	0.138
Node stage	1.585	1.094	2.297	0.015
Tumour size	1.161	0.891	4.960	0.633
BMI-1 protein expression	2.102	0.629	2.141	0.090

Table 4.10: Multivariate Cox regression hazard model including other prognostic clinicopathological parameters shows that high BMI1 (mRNA) provided a dependent prognostic factor in all BC and ER+ tumour.

Parameters	entire BC				ER+ tumour			
	Hazard ratio	95% Confidence Interval (CI)		p- value	Hazard ratio	95% Confidence Interval (CI)		p- value
		Lower	Upper			Lower	Upper	
Patient Age	1.113	0.904	1.372	0.312	1.624	1.149	2.295	0.006
Tumour Grade	1.385	1.180	1.625	<0.0001	1.364	1.138	1.635	0.001
Node Stage	1.858	1.649	2.093	<0.0001	1.740	1.508	2.022	<0.0001
Tumour Size	1.011	1.007	1.015	<0.0001	1.018	1.013	1.023	<0.0001
BMI-1	0.911	0.800	1.038	0.162	0.852	0.724	1.003	**0.045**

4.4 Discussion

BC is a heterogeneous disease. Understanding the causes and clinical implications of heterogeneity may help to improve patient outcomes and treatment regimens in BC [340]. Numerous studies of prognostic factors or biomarkers have illustrated the complex clinical and biological differences between different subtypes of BC [341-343]. In the current study, BMI-1 expression was evaluated at the proteomic, transcriptomic and genomic levels using tissue samples from well-characterised cohorts of patients with various molecular subtypes of BC.

BMI-1 has been linked to a poor prognosis and poorer survival in several types of cancer [344, 345]. Wang *et al.* (2012) reported that BMI-1 was associated with advanced features and poorer prognosis—including grade 3 tumours, the basal-like phenotype and poorer five-year overall survival—in 171 cases of BC [346]. However, Choi *et al.* (2009) reported that BMI-1 was an excellent prognostic biomarker in 960 cases of BC, of which 58% were ER+ [347]. Kim *et al.* (2004) reported conflicting results, as they found that BMI-1 was associated with ER+ tumours, which are less aggressive than tumours of other subtypes, and with lymph node metastasis [339]. The present study detected high BMI-1 protein and mRNA expression in 20% and 70% of BC cases, respectively. Two previous studies showed that high BMI-1 protein expression was present in 62% and 53% of BC cases [339, 347]. The slight differences between our results and previous studies could be explained by used of varied scoring methods, definitions of positive staining and cut-off values. This study used the H-scoring method, which

is widely-accepted in clinical and research settings, while previous studies used other scoring systems and different antibodies.

This work showed that the ER+ luminal subtype frequently expresses high levels of BMI-1 protein compared to the ER- subtype. Engelsen *et al.* (2008) reported low BMI-1 protein expression was associated with loss of the ER and PgR in endometrial carcinoma, suggesting a potential link between BMI-1 and hormone receptor status [316]. Significant associations between high BMI-1 expression and less advanced tumour features were observed in this study. Similarly, Choi *et al.* (2009) reported that overexpression of BMI-1 protein was associated with ER+ status and other favourable clinicopathological parameters, including smaller tumour size, negative lymph node metastasis, early tumour stage and intermediate nuclear grade in 960 cases of BC [348].

The present study also demonstrated that BMI-1 was an independent prognostic biomarker in the ER+ luminal subtype, regardless of other clinicopathological features. Although there was a limited number of cases of TNBC in the Nottingham cohort (n = 56), high BMI-1 protein expression was associated with poorer BCSS in that small subgroup. This agrees with a previous report that overexpression of BMI-1 was associated with significantly poorer BCSS in TNBC [349]. Overall, these results suggest BMI-1 plays diverse roles in various subtypes of BC and suggest other factors may regulate the BMI-1 signalling pathway in BC subtypes. To speculate, our findings might be related to ER-mediated BMI-1 role, and in TNBC where the ER absent, the BMI-1 regulated by others signalling pathways.

BMI-1 plays a vital function in regulation of stem cell transcriptional signalling pathways and is also implicated in the self-renewal, proliferation and cell cycle in CSCs [350, 351]. To the best of our knowledge, this is the first study to analyse the correlations between BMI-1 expression and the expression of BCSC-like biomarkers at both the transcriptomic and proteomic levels. BMI-1 expression correlated negatively with the expression of some BCSC markers at both the mRNA and protein in the luminal ER+ subtype. However, no significant correlations were observed between BMI-1 expression and the expression of most BCSC markers (only CD133 and SOX10) in TNBC, which further indicates BMI-1 exerts distinct roles in ER+ tumours and TNBC. A previous study reported high expression of BMI-1 in CD133+ cells in human glioblastoma multiforme [352]. The current study suggests that BMI-1 may only represent a BCSC-like marker in TNBC. The explanation for this inverse relationship is not known yet, however the possible explanation is ALDH1A1 is a BCSC markers and highly expressed in more aggressive tumour such as TNBC, where BMI-1 is mediated by ER expression, which suggest ER is inhibiting pro-tumorigenic role of BMI-1.

4.5 Conclusion

This study indicates that BMI-1 plays varied roles in different subtypes of BC. High BMI-1 was associated with significantly longer BCSS in ER+ BC but shorter BCSS in TNBC. These results indicate that ER status may modulate the biological function of BMI-1 in BC. Further functional studies are required to clarify the roles and prognostic value of BMI-1 in various subtypes of BC.

Chapter 5

Identification of a breast cancer stem cell phenotype: an immunohistochemical study

5.1 Introduction

Breast cancer (BC) is comprised of heterogeneous populations of cells with diverse biological properties, including a subset of cells termed breast cancer stem cells (BCSCs). BCSCs are characterised by the ability to self-renew, differentiate and evade the immune system, and are associated with aggressive tumour behaviour, resistance to therapy and metastatic potential [353-355]. Moreover, differentiation of BCSCs leads to intratumour and intertumour heterogeneity, and this heterogeneity poses a major challenge to the design of effective therapies [356]. Thus, due to their significant role in tumour progression, it is important to characterise BCSCs in order to identify develop potential anti-BCSC therapies [291].

Identification of cells with BCSC phenotypes represents an essential step in the study of BCSCs. Various molecules have been reported to selectively mark BCSC populations [357, 358]. Indeed, the existence of multiple stemness-related markers indicates the complexity of the mechanisms that regulate BCSCs, which involve a variety of signalling pathways [359, 360].

Multiple studies have reported that a number of transcription factors, such as SOX2, SOX9, OCT4, KLF4, Nanog and SALL4, are overexpressed by cancer stem cells [361-364]. Furthermore, cell surface proteins, including CD24, CD44,

CD133 and EpCAM, have been used to isolate BCSCs [365-368]. Other BCSC markers are associated with metabolism in tumour cells, for example, the members of the ALDH1 family [193, 251, 369, 370].

Although several markers are expressed by CSCs, no single marker can be used to detect all CSCs. A study in acute myeloid leukaemia pioneered the concept of CSCs by combining the markers $CD34^{+}$ and $CD38^{-}$. In a subsequent study, Al-Hajj *et al.* (2003) combined CD44+ and CD24- to identify the BCSC phenotype [147, 371]. However, other studies have demonstrated high expression of these markers throughout BC tissue samples [372-374], which contradicts the suggestion that BCSCs only represent a subpopulation of cells equivalent to 1–10% of a tumour mass [375].

Although previous investigations have provided some understanding of BCSCs [376-379], accurate identification of BCSCs is not yet fully understood. BCSC populations have been reported to express different markers. Moreover, most BCSC markers are not exclusive to BCSCs and are also expressed in other normal or tumour cell types in the breast [148]. Therefore, the isolation of BCSCs remains challenging and there are no standardised criteria for identification of BCSCs in BC tissue samples.

5.1.1 Hypothesis

Assessment of the co-localisation of key BCSC markers in BC tissue samples using immunohistochemistry will enable identification of sub-clones of tumour cells that are likely to represent BCSCs.

5.1.2 Study aims

The aims of this chapter are to:

(i) Identify BCSC subpopulations in BC tissue samples from a large, well-characterised BC cohort using double immunohistochemical staining.

(ii) Examine the relationships between the identified BCSC phenotype and the clinicopathological features of BC and patient outcomes.

5.1.3 Objectives

The objectives of this chapter are to:

(i) Compare the expression of the immunohistochemical BCSC markers ALDH1A1, SOX9, EpCAM, BMI-1, SOX10, CD133, CD24 and CD44 in different subtypes of IBC.

(ii) Analyse the correlations between expression of these BCSC markers in IBC.

(iii) Analyse the prognostic significance of ALDH1A1, SOX9 and EpCAM co-expression in IBC.

(iv) Perform double IHC staining for selected biomarkers (ALDH1A1 and SOX9) and (ALDH1A1 and EpCAM) in IBC.

(v) Identify the specific population of tumour cells characterised by co-localisation of ALDH1A1, SOX9 and EpCAM in BC tissue samples.

(vi) Determine the prognostic value of the ALDH1A1+SOX9+EpCAM+ subpopulation in different subtypes of IBC.

5.2 Materials and Methods

5.2.1 Patient cohort

A large, well-characterised cohort (*n* = 1519) of patients with BC enrolled at Nottingham City Hospital between 1989 and 2006 was utilised in this study; the cohort is described in more detail in Chapter 2. Overall, 48% of the patients presented with stage 2 or 3 tumours and 33% presented with grade 3 tumours. Metastatic BC and recurrent BC accounted for 27% and 33% of the study cohort, respectively. The demographic features of the patients are summarised in Table 5.1.

5.2.2 Availability of biomarker data for the cohort

Data on immunohistochemical detection of a panel of BCSC markers, including CD133, CD24, CD44, EpCAM, SOX9 and SOX10, was available for the study cohort [229, 291, 380]. Moreover, the expression of ALDH1A1 and BMI-1 were assessed in Chapters 3 and 4. The details of each biomarker are shown in chapter 2.2.

Table 5.1: Demographic features of the Nottingham patient cohort

Clinicopathological feature	***N* (%)**
Age	
≥ 50 years	840 (56)
< 50 years	671 (44)
Tumour size	
≥ 2 cm	600 (40)
< 2 cm	904 (60)
Grade	
Grade 1	232 (16)
Grade 2	546 (36)
Grade 3	725 (48)
Nottingham prognostic index (NPI)	
Good	487 (32)
Moderate	821 (55)
Poor	196 (13)
Tumour type	
Ductal (including mixed)	1021 (87)
Lobular	72 (6)
Medullary-like	30 (2.5)
Miscellaneous	2 (0.5)
Special type	36 (3)
Vascular invasion	
Definite	431 (29)
Negative/probable	1069 (71)
Lymph node stage	
Stage 1	1006 (68)
Stage 2	422 (28)
Stage 3	76 (5)
Follow-up status (>20 years)	
Alive	958 (60)
Died from breast cancer	359 (23)
Died from other causes	279 (17)
Oestrogen receptor status	
Negative	356 (24)
Positive	1145 (76)
Progesterone receptor status	
Negative	627 (42)
Positive	849 (58)
HER2 status	
Negative	1274 (86)
Positive	209 (14)

5.2.3 Validation of antibody specificity by western blotting

Prior to performing IHC staining, the specificity of the selected SOX9 (1:1000) and EpCAM (1:250) antibodies were validated using western blotting, as described in chapter 2.

The resulting bands corresponded to the expected molecular weight of the proteins (SOX9: 72 KDa and EpCAM: 54 KDa) and confirmed the specificity of the antibodies (Figure 5.1). Table 2.4 provides the details of the antibodies used, including the suppliers, antibody dilutions, incubation times, cell lysates and molecular weights of the bands detected.

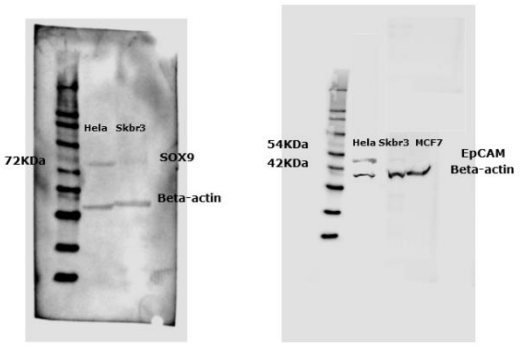


Figure 5.1: Western blots to validate the rabbit monoclonal anti-SOX9 antibody and EpCAM antibody. Single, specific bands were detected at the expected molecular weights for SOX9 and EpCAM (72 and 54 kDa, respectively) in HeLa cell lysates. The positive control β-actin was detected at 42 kDa. The bands of the Magic ladder range from 20 to 220 KDa.

5.2.4 Double IHC staining

Prior to the double IHC staining, the single staining protocols for ALDH1A1, SOX9 and EpCAM were optimised, as described in detail in Chapter 2.

The Multiplex IHC Detection Kit for staining protocols involving a combination of rabbit and mouse antibodies was used for the double IHC staining, as described in Chapter 2. In brief, antigen retrieval was performed for pairs of sequential TMA sections in a microwave (Whirlpool JT359 Jet Chef 1000 W) in citrate buffer (pH 6.0) for 20 min, and the staining was performed following the A. Menarini diagnostics protocol [381], as described in chapter 2.

The optimal dilution of the ALDH1A1 antibody and SOX9 antibody for double IHC was 1:100 for both antibodies, with an incubation time of overnight and 1 h, respectively. Double IHC for ALDH1A1 + EpCAM was performed using both antibodies at 1:100 dilution; the ALDH1A1 antibody was incubated overnight and the EpCAM antibody was incubated for 1 h.

The Multiplex IHC Detection Kit contains goat anti-mouse polymer secondary antibodies conjugated to alkaline phosphatase (AP) and goat anti-rabbit polymer secondary antibodies conjugated to horseradish peroxidase (HRP), which react with the heavy and light chains of mouse and rabbit primary antibodies, respectively. The sections were developed by applying the fast-red and green chromogens in the kit (described in greater detail in Chapter 2).

5.2.5 Double IHC scoring

The TMA-stained slides were scanned at 20× magnification into high-resolution digital images using a NanoZoomer scanner (NanoZoomer; Hamamatsu

Photonics, Welwyn Garden City, UK). Assessment of the expression of the combinations of the biomarkers ALDH1A1+SOX9 and ALDH1A1+EpCAM were performed using CaseViewer software (https://www.3dhistech.com/caseviewer).

The paired sequential TMA sections for both combinations were assessed in parallel by determining the percentages of cells expressing cytoplasmic ALDH1A1, nuclear SOX9, and cytoplasmic and membrane EpCAM+, and the percentages of ALDH1A1+SOX9+ cell, ALDH1A1+EpCAM+ cells, and ALDH1A1+ SOX9+EpCAM+ cells.

The intensity cut-off values were determined according to the expression of ALDH1A1, SOX9 and EpCAM in normal breast ducts, as shown in Figure 5.2. Positive staining (expression) was determined as higher expression than in normal ducts. The cut-off for positive staining was defined according to previous studies, in which the expression of stem cell markers was higher in tumour cells than normal cells [149, 295, 382].

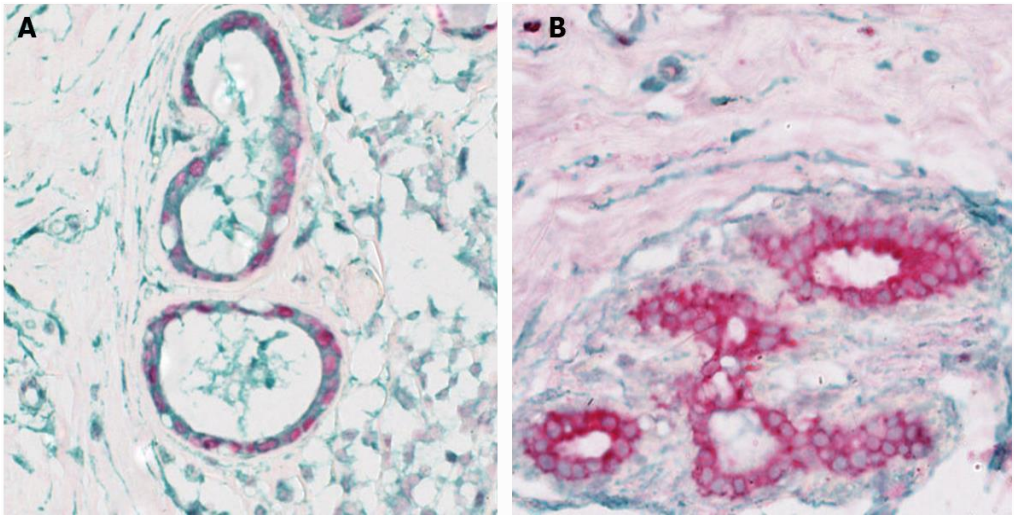


Figure 5.2: Immunohistochemical double staining of a normal human breast duct for ALDH1A1, SOX9 and EpCAM. A) The green colour in the cytoplasm indicates positive expression of ALDH1A1, and the red colour in the nucleus shows positive expression of SOX9. B) The red colour in the cytoplasm and membrane indicate positive expression of EpCAM.

5.2.6 Statistical analyses

All statistical analyses were performed using IBM SPSS 24.0 software, as described in Chapter 2. For all statistical tests, a two-sided *P*-value < 0.05 was considered significant.

5.3 Results

5.3.1 BCSC marker expression in BC tissues

Assessment of the results of single immunohistochemical staining for the study cohort revealed expression of the tested BCSC markers varied significantly in the BC tissue samples.

ALDH1A1 immunoexpression was mainly localised in the cytoplasm; EpCAM, CD44 and CD24 were mainly detected in cell membranes; SOX9, SOX10 and BMI-1 were mainly expressed in nuclei; and CD133 was detected in the membrane and cytoplasm of tumour cells.

High ALDH1A1, EpCAM, SOX9 and CD133 immunohistochemical expression (based on the mean percentage of positive cells) were observed in ER- tumours, and high SOX10 and BMI-1 expression were observed in ER+ tumours. CD24 exhibited similar expression in ER+ and ER- tumours, while HER2+ tumours exhibited the highest expression of CD44. BMI-1 was expressed in the highest percentage of cells (mean = 80%) and EpCAM was detected in the lowest percentage of cells (mean = 12%). Table 5.2 summarises the mean and median percentages of positive cells for all BCSC markers in each subtype of IBC.

Table 5.2: Mean and median percentages of cells expressing the tested breast cancer stem cell (BCSC) markers in the entire breast cancer (BC) cohort and oestrogen receptor negative (ER-), oestrogen receptor positive (ER+) and human epidermal growth factor receptor 2 positive (HER2+).

Marker	Entire cohort		ER- tumours		ER+ tumours		HER2+ tumours	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
ALDH1A1	53	65	56	70	52	60	55	72
BMI-1	82	80	48	30	93	90	59	40
CD133-C*	31	0	64	100	19	0	36	0
CD133-M*	16	0	43	0	5	0	17	0
CD24 M*	67	0	68	0	68	0	66	0
CD44 M*	25	0	27	0	25	0	30	0
EpCAM M*	12	0	25	0	11	0	24	0
SOX10	52	40	39	20	57	50	48	35
SOX9	25	0	38	0	27	0	30	9

M*, Membrane expression
C*, Cytoplasmic expression

5.3.2 Correlations between expression of the BCSC markers and selected the biomarkers for double IHC

The Spearman's rank test was performed to assess the correlations between the tested BCSC markers using the percentages of cells expressing each marker as continuous variables (Figure 5.3).

The Spearman's rho correlation values indicated that SOX10 correlated negatively with SOX9, CD133, ALDH1A1, CD24 and EpCAM. However, it is correlated positively with BMI-1 and CD44.

Furthermore, ALDH1A1, SOX9 and EpCAM all correlated positively with each other as continuous variables in the Spearman's rank analysis (all $P < 0.05$).

When assessed as categorical data using the Chi squared test, ALDH1A1, SOX9 and EpCAM all correlated positively with each other (Table 5.3).

Based on their correlations and associations as continuous or categorical data, ALDH1A1, SOX9 and EpCAM were selected as markers to identify the BCSC population by detecting co-localisation of ALDH1A1, SOX9 and EpCAM expression using double IHC. The associations between these biomarkers and the clinicopathological features and outcomes of BC were also examined.

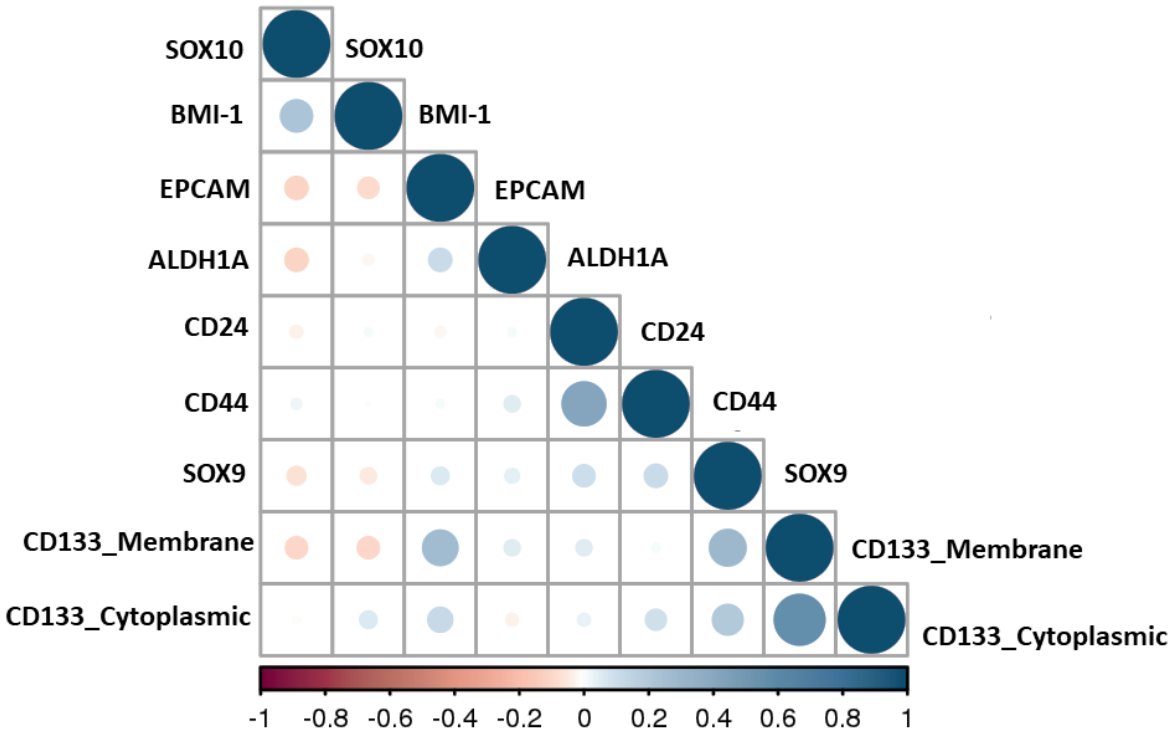


Figure 5.3: Correlation matrix summarising the correlations between the expressions of all tested breast cancer stem cell (BCSC) markers in the entire study cohort. Positive correlations are indicated in blue and negative correlations in red. The intensity of colour is proportional to the correlation coefficient.

Table 5.3. Associations between ALDH1A1, SOX9 and EpCAM expression in breast cancer (BC) (chi squared test)

	ALDH1A1 status (IHC)		
Biomarker	**Low** **No. (%)**	**High** **No. (%)**	**X^2** ***P*-value**
SOX9 status Low High	 121 (32) 142 (26)	 256 (68) 410 (74)	4.48 **0.034**
EPCAM status Low High	 209 (33) 55 (19)	 431 (67) 235 (81)	18.40 **< 0.0001**
	SOX9 status (IHC)		
Biomarker	Low No. (%)	High No. (%)	X^2 *P*-value
EPCAM status Low High	 212 (34) 18 (22)	 270 (66) 70 (78)	 4.430 **0.04**

5.3.3 Associations between EpCAM and SOX9 expression and the clinicopathological parameters of BC

Furthermore, the prognostic significance of EpCAM and SOX9 were examined. The Chi square test showed that high expression of EpCAM or SOX9 alone were associated with aggressive tumour features, including high-grade tumours, poor NPI, LVI, and stage 3 tumours (all $P < 0.05$; see Table 5.4).

Table 5.4: Associations between SOX9 and EpCAM expression and the clinicopathological features of the entire breast cancer cohort*.

Parameter	SOX9			EpCAM		
	High No (%)	Low No (%)	x^2 *P*	High No (%)	Low (%)	x^2 *P-value*
Patient age (years)						
< 50	638 (50)	646 (50)	3.64	973 (76)	312 (24)	**3.98**
≥ 50	302 (45)	367 (55)	0.056	478 (72)	190 (28)	**0.046**
Tumour size						
≤ 2 cm	405 (44)	525 (56)	14.44	641 (69)	290 (20)	**29.31**
> 2 cm	527 (52)	483 (48)	**< 0.0001**	803 (80)	206 (26)	**< 0.0001**
Tumour grade						
Grade 1	281 (64)	160 (36)	**26.55**	381 (87)	59 (13)	**81.84**
Grade 2	424 (56)	333 (44)	**< 0.0001**	636 (84)	121 (16)	**< 0.0001**
Grade 3	503 (50)	513 (50)		701 (69)	315 (31)	
Tubule formation						
1						
2	80 (56)	63 (44)	**7.48**	128 (90)	15 (11)	**25.97**
3	376 (56)	300 (44)	**0.024**	536 (80)	138 (20)	**< 0.0001**
	607 (50)	619 (50)		893 (73)	334 (27)	
Mitotic count scores						
1						
2	467 (58)	342 (42)	**19.13**	693 (86)	115 (14)	**86.91**
3	191 (51)	185 (49)	**<0.0001**	293 (78)	83 (22)	**< 0.0001**
	405 (47)	455 (53)		571 (78)	289 (34)	
Nuclear pleomorphism						
1	40 (58)	29 (42)	**17.17**	65 (94)	4 (5.8)	**50.66**
2	479 (57)	361 (43)	**< 0.0001**	692 (82)	147 (18)	**< 0.0001**
3	541 (48)	589 (52)		797 (71)	333 (29)	
Axillary nodal stage						
Stage 1						
Stage 2	880 (59)	622 (41)	**14.56**	1205 (80)	295 (20)	**12.15**
Stage 3	298 (50)	297 (50)	**0.001**	441 (74)	154 (27)	**0.002**
	89 (51)	87 (49)		130 (73)	47 (27)	
Nottingham prognostic index						
Good						
Moderate	346 (54)	299 (46)	**14.20**	537 (83)	107 (17)	**49.06**
Poor	462 (46)	535 (54)	**0.001**	717 (72)	280 (28)	**< 0.0001**
	124 (42)	174 (58)		190 (64)	109 (36)	
Vascular invasion						
Negative	678 (51)	649 (49)	**16.55**	1028 (23)	298 (22)	**22.09**
Positive	247 (41)	58.9 (59)	**< 0.0001**	406 (67)	196 (32)	**<0.0001**
Oestrogen receptor status						
Positive	809 (52)	738 (48)	0.76	1236 (80)	310 (36)	**53.76**
negative	264 (50)	263 (50)	0.38	338 (64)	189 (20)	**<0.0001**
Progesterone status						
Positive	541 (49)	570 (51)	0.81	884 (80)	226 (20)	**40.33**
Negative	354 (47)	406 (53)	0.36	506 (67)	254 (33)	**<0.0001**
HER2 status						
Positive	104 (45)	126 (55)	1.012	145 (63)	86 (37)	**18.85**
Negative	806 (49)	847 (51)	0.314	1256 (76)	395 (24)	**<0.0001**

*Chi square test; significant *P*-values are highlighted in **bold.**

5.3.4 Survival analysis for SOX9 and EpCAM expression

5.3.4.1 Breast cancer-specific survival

High expression of EpCAM and SOX9 were significantly associated with shorter breast cancer-specific survival (BCSS) in the entire BC cohort ($P = 0.008$ and $P < 0.0001$, respectively; Figure 5.4), but not in the ER- subgroup. High expression of EpCAM was also significantly associated with shorter BCSS in patients with ER+ ($P = 0.004$) and HER2+ tumours ($P = 0.019$; Figure 5.5). However, there was no significant relationship between SOX9 and BCSS in any subtype of BC ($P \geq 0.05$; Figures 5.4 and 5.5).

5.3.4.2 Disease-free interval

High SOX9 and high EpCAM expression were significantly associated with a shorter disease-free interval (DFI) in the whole BC cohort ($P = 0.001$ and $P < 0.0001$, respectively; Figure 5.6). For patients with HER2+ tumours, SOX9 expression was significantly associated with shorter DFI ($P = 0.042$; Figure 5.7), while there was no significant association between SOX9 and EpCAM expression and DFI in the ER+ and ER- subtypes ($P \geq 0.05$; Figures 5.6 and 5.7).

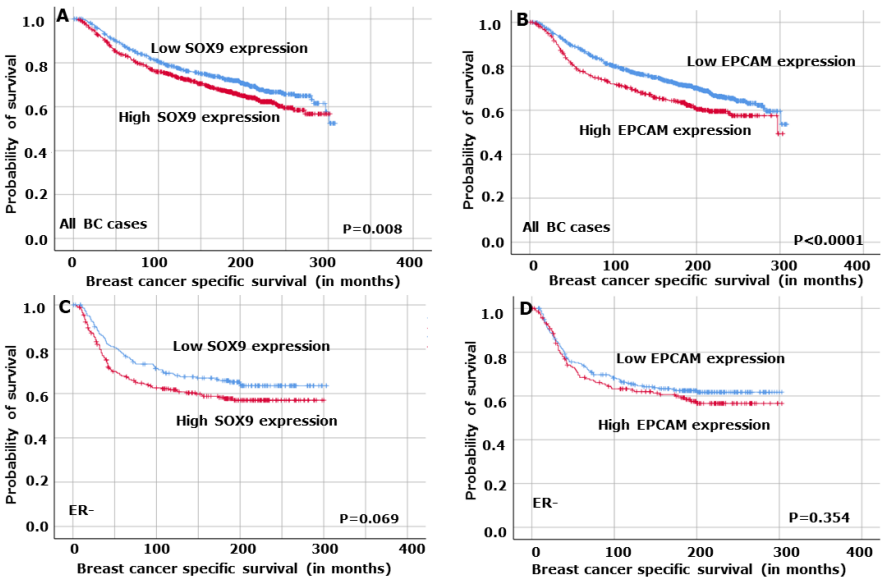


Figure 5.4: Associations between SOX9 and EpCAM expression and breast cancer-specific survival (BCSS) in the entire cohort and the Oestrogen receptor negative (ER-) subgroup. (A, and B) Associations between SOX9 and EpCAM, respectively, and BCSS in the entire breast cancer (BC) cohort. (C, D) Associations between SOX9 and EpCAM, respectively, and BCSS in the ER- subgroup.

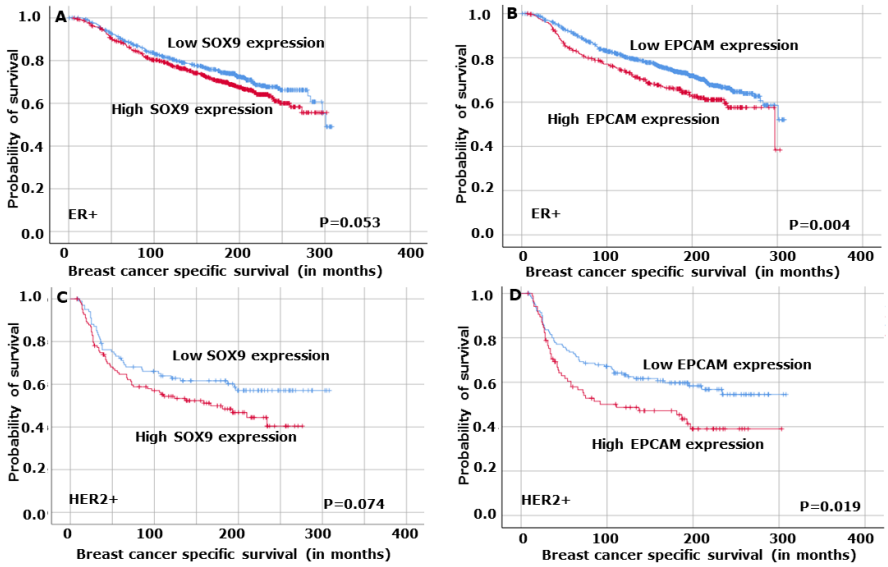


Figure 5.5: Associations between SOX9 and EpCAM expression and breast cancer-specific survival (BCSS) in the Oestrogen receptor positive (ER+) and human epidermal growth factor receptor 2 positive (HER+) subgroups. (A and B) Associations between SOX9 and EpCAM expression, respectively, and BCSS in the ER+ subgroup. (C, D) Associations between SOX9 and EpCAM expression, respectively, and BCSS in the HER2+ subgroup.

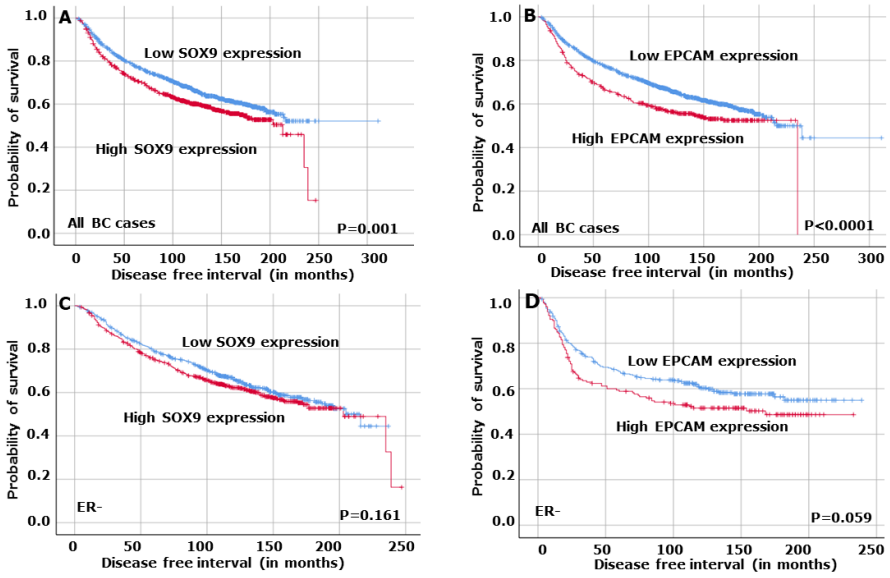


Figure 5.6: Associations between SOX9 and EpCAM expression and disease-free interval (DFI) in the entire cohort and Oestrogen receptor negative (ER-) subgroup. (A and B) Associations between SOX9 and EpCAM expression, respectively, and DFI in the entire breast cancer (BC) cohort. (D and E) Associations between SOX9 and EpCAM, respectively, and DFI in the ER- subgroup.

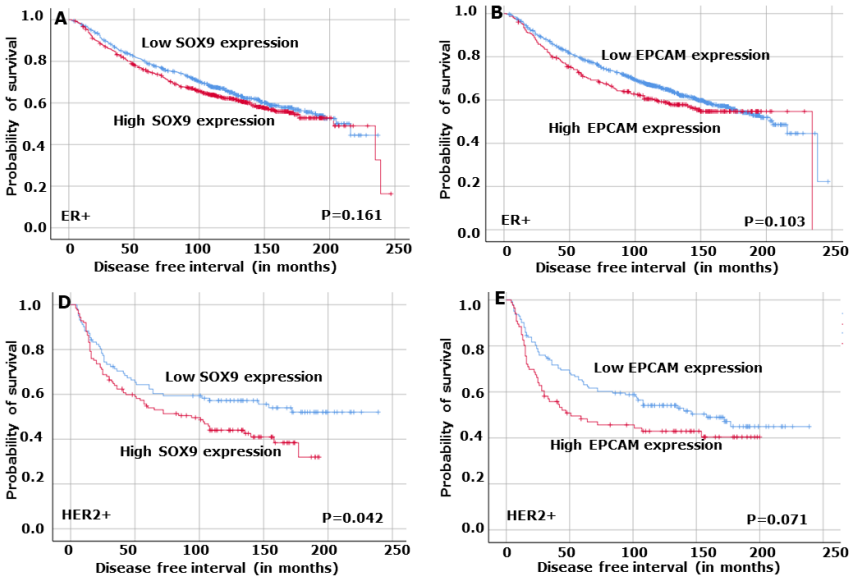


Figure 5.7: Associations between SOX9 and EpCAM expression and disease-free interval (DFI) in the oestrogen receptor positive (ER+) and human epidermal growth factor receptor 2 positive (HER2+) subgroups. (A, B) Associations between SOX9 and EpCAM expression, respectively, and DFI in the ER+ subgroup. (C and D) Associations between SOX9 and EpCAM expression, respectively, and DFI in the HER2+ subgroup.

5.3.4.3 Distant metastasis-free survival

There were significant associations between high SOX9 and high EpCAM expression and shorter distant metastasis-free survival (DMFS) in the whole BC cohort ($P < 0.0001$; Figure 5.8). High expression of SOX9 was also significantly associated with shorter DMFS in the ER+ and ER- subgroups ($P = 0.017$ and $P = 0.032$, respectively; Figures 5.8 and 5.9). High expression of EpCAM was significantly associated with shorter DMFS in the ER+ and HER2+ subgroups ($P = 0.010$ and $P = 0.0023$, respectively; Figure 5.9).

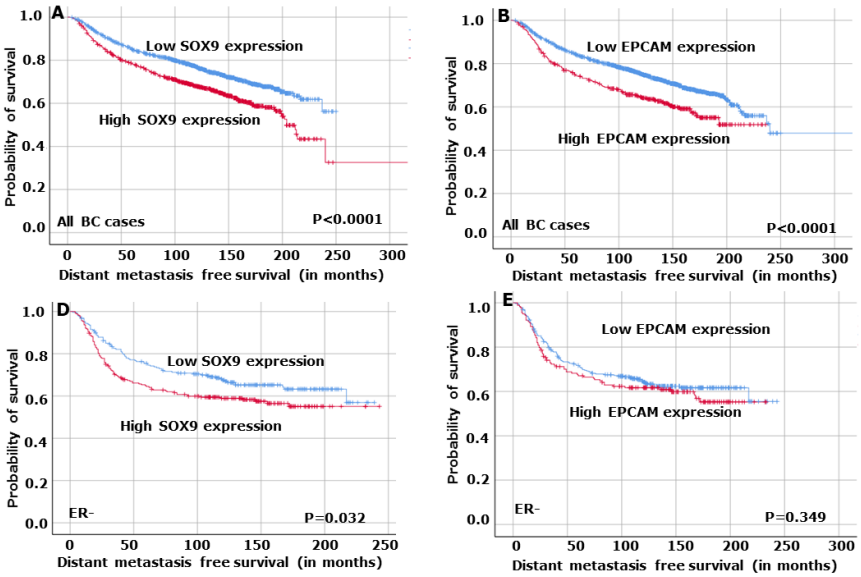


Figure 5.8: Associations between SOX9 and EpCAM expression and distant metastasis-free survival (DMFS) in the entire cohort and Oestrogen receptor negative (ER-) subgroup. (A, B) Associations between SOX9 and EpCAM expression, respectively, and DMFS in the whole breast cancer (BC) cohort. (C and D) Associations between SOX9 and EpCAM expression, respectively, and DMFS in the ER- subgroup.

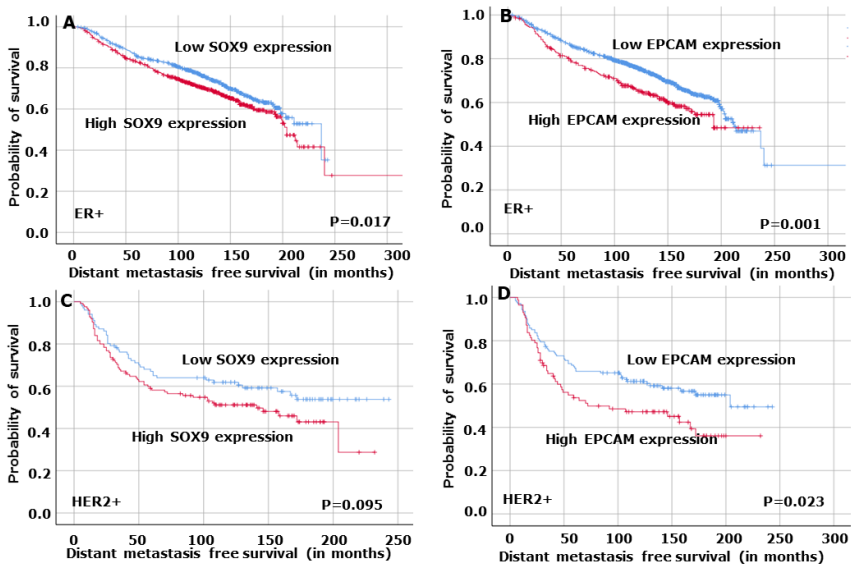


Figure 5.9: Associations between SOX9 and EpCAM expression and distant metastasis-free survival (DMFS) in the ER+ and HER2+ subgroups. (A, and B) Associations between SOX9 and EpCAM expression, respectively, and DMFS in the ER+ subgroup. (C and D) Associations between SOX9 and EpCAM expression, respectively, and DMFS in the HER2+ subgroup.

5.3.5 Co-expression of the biomarkers ALDH1A1, SOX9 and EpCAM in BC

A *k*-means test was performed to classify the cases of BC according to the expression of ALDH1A1, SOX9 and EpCAM. The classifications included ALDH1A1+SOX9+EpCAM+, ALDH1A1+SOX9-, EpCAM-, ALDH1A1+SOX9+EpCAM-, ALDH1A1+SOX9-EpCAM+, ALDH1A1-SOX9+EpCAM+, and ALHD1A1-SOX9-EpCAM-.

Kaplan-Meir survival curves and the log rank test indicated that co-expression of ALHD1A1+SOX9+EpCAM+ was associated with shorter BCSS, DFI and DMFS compared with the other classifications in the entire BC cohort (all $P < 0.05$; Figures 9A–C). Co-expression of ALHD1A1+SOX9+EpCAM+ was also associated with shorter BCSS and DMFS in ER+ tumours (all $P < 0.05$; Figures 5.10), but not in the ER- or HER2+ subtypes ($P \geq 0.05$).

To further confirm these observations, the associations between the clinicopathological features of BC and co-expression of ALDH1A1+SOX9+EpCAM+ versus ALHD1A1-SOX9-EpCAM- was tested. ALDH1A1+SOX9+EpCAM+ was associated with advanced grade, advanced stage, poor NPI and LVI compared to ALDH1A1-SOX9-EpCAM- tumours (all $P < 0.005$; Table 5.5).

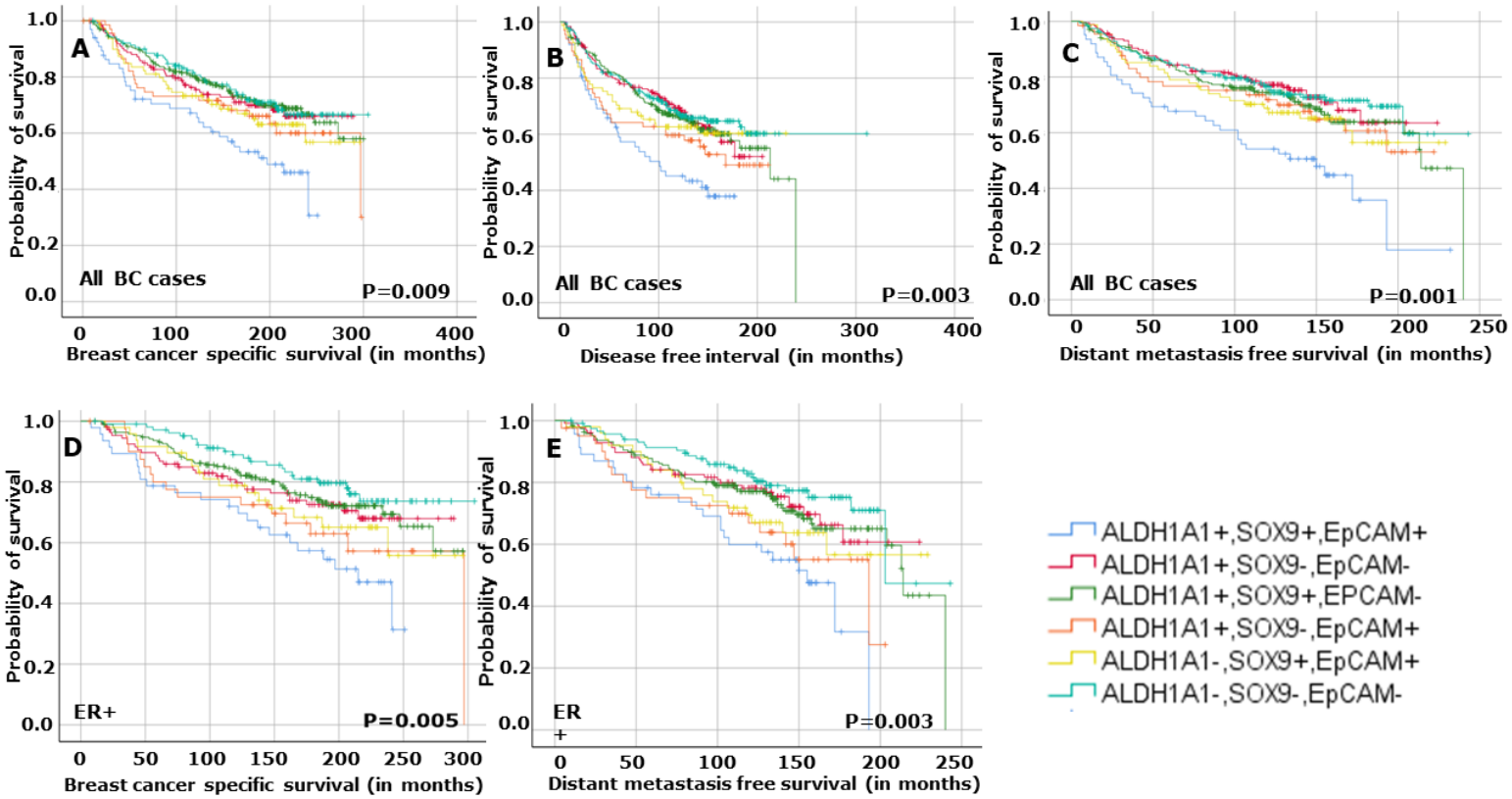


Figure 5.10: (A, B and C) Associations between co-expression of ALDH1A1, SOX9 and EpCAM and breast cancer-specific survival (BCSS), disease-free interval (DFI), and distant metastasis-free survival (DMF), respectively, in the entire cohort. (D and F) Associations between co-expression of ALDH1A1, SOX9 and EpCAM and BCSS and DMF, respectively, in the oestrogen receptor positive (ER+) subgroup.

Table 5.5: Associations between ALDH1A1+SOX9+EpCAM+ co-expression and the clinicopathological features of breast cancer (BC) in the entire cohort.

Clinicopathological Parameter	ALDH1A1+SOX9+EpCAM co-expression		
	Negative No (%)	Positive No (%)	x^2 *P*-value
Patient age (years)			
< 50	208 (85)	36 (15)	3.7
≥ 50	102 (77)	30 (23)	0.052
Tumour size			
≤ 2 cm	140 (76)	45 (24)	**11.23**
> 2 cm	168 (89)	21 (11)	**0.001**
Tumour grade			
Grade 1	76 (95)	4 (5)	**17.45**
Grade 2	120 (87)	18 (13)	**< 0.0001**
Grade 3	134 (75)	44 (25)	
Tubule formation			
1	23 (92)	2 (8)	**10.69**
2	122 (90)	14 (10)	**0.005**
3	170 (77)	50 (23)	
Mitotic count score			
1	140 (94)	9 (6)	**22.29**
2	54 (78)	15 (22)	**< 0.0001**
3	121 (74)	42 (26)	
Nuclear pleomorphism			
1	11 (92)	1 (8)	**6.98**
2	142 (88)	19 (12)	**0.030**
3	162 (78)	45 (22)	
Axillary nodal stage			
Stage 1	224 (88)	31 (12)	**9.46**
Stage 2	97 (78)	27 (22)	**0.009**
Stage 3	19 (70)	8 (30)	
Nottingham prognostic index			
Good			
Moderate	118 (83)	9 (7)	**21.00**
Poor	151 (80)	37 (20)	**< 0.0001**
	39 (66)	20 (34)	
Vascular invasion status			
Negative	220 (88)	29 (12)	**18.45**
Positive	88 (70)	37 (30)	**< 0.0001**
Oestrogen receptor status			
Positive	252 (84)	47 (16)	1.22
Negative	73 (79)	19 (21)	0.26
Progesterone receptor status			
Positive	177 (84)	33 (16)	1.77
Negative	115 (79)	31 (21)	0.18
HER2 status			
Positive	32 (80)	8 (20)	0.17
Negative	263 (83)	55 (17)	0.67

5.3.6 Co-localisation of ALDH1A1+, SOX9+ and EpCAM+ expression in BC tissue samples using double IHC

Positive co-localisation of ALDH1A1+SOX9+ was observed in 22% (308/1380 cases) of the entire study cohort, while single expression of ALDH1A1+ or SOX9+ was observed in 76% and 52% of the cohort, respectively. In addition, the percentage of ALDH1A1+SOX9+ positive tumour cells ranged from 1% to 100%.

Co-localisation of ALDH1A1+EpCAM+ was observed in 40% (562/1380) of the cases in the entire cohort, while single expression of ALDH1A1 or EpCAM were observed in 76% and 68% of the entire BC cohort, respectively. Figure 5.11 illustrates the expression of ALDH1A1, SOX9 and EpCAM.

Table 5.6 shows the mean, median and range percentages of cells co-expressing ALDH1A1+SOX9+ and ALDH1A1+EpCAM+ in the entire BC cohort and each subtype.

Co-localisation of ALDH1A1+SOX9+EpCAM+ was also assessed. Only 12% (179/1519) of the cases in the entire cohort contained cells that exhibited co-localisation of ALDH1A1+SOX9+EpCAM+. The percentage of ALDH1A1+SOX9+EpCAM+ tumour cells ranged from < 1% to 10%.

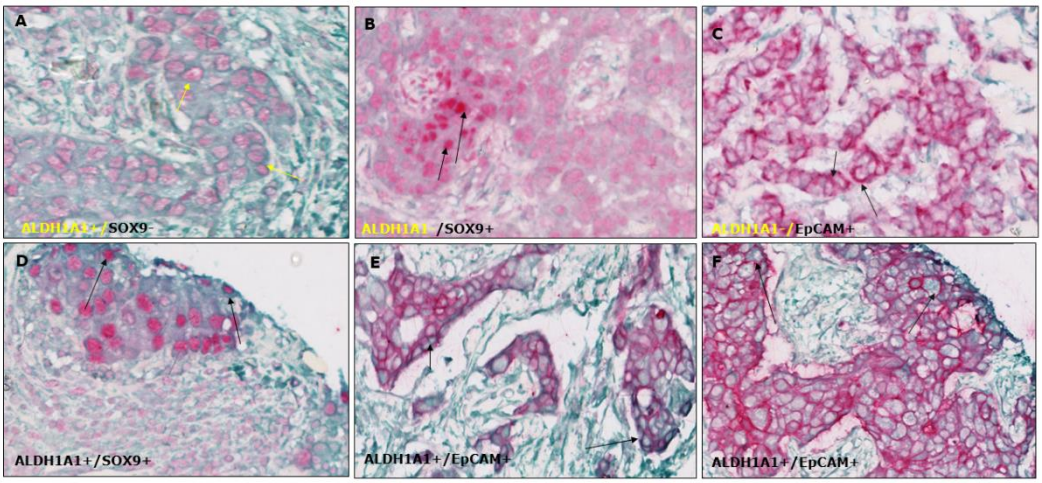


Figure. 5.11. Double immunohistochemical staining of human breast tumours for ALDH1A1, SOX9 and EpCAM. SOX9 and EpCAM were stained with fast red chromogen, and ALDH1A1 with a green chromogen. (A) ALDH1A1+/SOX9-, (B) SOX9+/ALDH1A1- tumour cells (arrow). (C) EpCAM+/ALDH1A1- tumour cells (arrow). (D) A single ALDH1A1+/SOX9+ tumour cell (arrow). (E and F) Single ALDH1A1+/EpCAM+ tumour cells (arrows).

Table 5.6: Percentages of cells co-expressing ALDH1A1+SOX9+ and ALDH1A1+EpCAM+ in the entire breast cancer (BC) cohort and ER-, ER+ and HER2+ subtypes.

Biomarker	Entire cohort			ER- subtype			ER+ subtype			HER2+ subtype		
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
ALDH1A1+SOX9+	1	0	100	2	0	80	2	0	100	2	0	100
ALDH1A1+EpCAM+	6	0	100	3	0	100	1	0	100	7	0	100

5.3.7 Associations between the ALDH1A1+SOX9+EpCAM+ phenotype and the clinicopathological features of BC

The ALDH1A1+SOX9+EpCAM+ phenotype positively correlated with advanced nodal stage in the entire study cohort and HER2+ subgroup ($P = 0.012$ and $P = 0.002$, respectively). However, no significant relationships between ALDH1A1+SOX9+EpCAM+ cells and other advanced features of BC, including high grade, NPI, advanced histological type and lymphovascular invasion (all $P \geq 0.05$), were observed in the entire cohort or HER2+ subgroup (Table 5.7).

In ER- tumours, the presence of cells with the ALDH1A1+SOX9+EpCAM+ phenotype was positively associated with poor NPI ($P = 0.031$), advanced nodal stage ($P = 0.002$) and lymphovascular invasion ($P = 0.029$). However, no significant relationships between the presence of cells with the ALDH1A1+SOX9+EpCAM+ phenotype and any clinicopathological feature were observed in ER+ tumours (Table 5.8).

Table 5.7: Associations between the ALDH1A1+SOX9+EpCAM+ breast cancer stem cell (BCSC) phenotype and the clinicopathological features of breast cancer (BC) in the entire cohort and human epidermal growth factor receptor 2 positive (HER+) subgroup.

Clinicopathological Parameters	ALDH1A1+SOX9+EpCAM+ cells in entire cohort			ALDH1A1+SOX9+EpCAM+ cells in HER2+		
	Negative No (%)	Positive No (%)	x^2 *P*-value	Negative No (%)	Positive No (%)	x^2 *P*-value
Patient age (years)						
< 50	601 (90)	70 (10)	1.91	88 (83)	18 (17)	0.08
≥ 50	733 (87)	107 (13)	0.166	87 (85)	16 (16)	0.77
Tumour size						
≤ 2 cm	800 (89)	104 (11)	0.15	113 (86)	18 (14)	1.75
> 2 cm	527 (88)	73 (12)	0.69	61 (79)	16 (21)	0.12
Tumour grade						
Grade 1	201 (87)	31 (13)	4.28	2 (50)	2 (50)	3.71
Grade 2	473 (87)	73 (13)	0.11	35 (81)	8 (18.6)	0.15
Grade 3	653 (90)	72 (10)		137 (85)	24 (15)	
Tubule formation						
1						
2	81 (87)	12 (13)	0.49	2 (68)	1 (33)	1.18
3	427 (88)	60 (12)	0.78	41 (87)	6 (13)	0.55
	790 (89)	100 (11)		128 (83)	27 (17)	
Mitotic count scores						
1	512 (88)	73 (13)	2.85	19 (76)	6 (24)	1.52
2	260 (87)	40 (13)	0.24	45 (82)	10 (18)	0.466
3	525 (90)	58 (10)		107 (86)	18 (14)	
Nuclear pleomorphism						
1	20 (95)	1 (5)	2.19	0 (0)	0 (0)	0.86
2	395 (87)	60 (13)	0.33	16 (76)	5 (24)	0.25
3	881 (89)	111 (11)		154(84)	29 (16)	
Axillary nodal stage						
Stage 1	894 (90)	112 (11)	**8.81**	96 (87)	15 (13)	12.01
Stage 2	375 (89)	47 (11)	**0.012**	66 (87)	10 (13)	0.002
Stage 3	59 (78)	17 (22)		12 (57)	9 (43)	
Nottingham prognostic index						
Good	429 (88)	58 (12)	1.020	17 (77)	5 (23)	3.42
Moderate	729 (89)	92 (11)	0.600	116 (87)	17 (13)	0.18
Poor	169 (86)	27 (14)		41 (77)	12 (23)	
Vascular invasion						
Negative	945 (88)	124 (12)	0.14	112 (86)	19 (15)	0.80
Positive	378 (88)	53 (12)	0.70	63 (81)	15 (19)	0.37
Oestrogen receptor status						
Positive	1006 (88)	139(12)	0.56	103 (86)	17 (14)	0.913
Negative	318 (89)	38 (11)	0.45	72 (81)	17 (19)	0.22
Progesterone receptor status						
Positive	748 (88)	101 (12)	0.022	110 (81)	26 (19)	2.09
negative	554 (88)	73 (12)	0.881	63 (89)	8 (11)	0.104
HER2 status						
Positive	1135 (89)	139 (11)	**5.00**	N/A	N/A	N/A
Negative	175 (84)	34 (16)	**0.025**			

Table 5.8: Associations between the ALDH1A1+SOX9+EpCAM+ breast cancer stem cell (BCSC) phenotype and the clinicopathological features of oestrogen receptor negative (ER-) and oestrogen receptor positive (ER+) tumours.

Clinicopathological Parameters	ALDH1A1+SOX9+EpCAM+ cells in ER- tumours			ALDH1A1+SOX9+EpCAM+ cells in ER+ tumours		
	Negative No (%)	Positive No (%)	x^2 *P*-value	Negative No (%)	Positive No (%)	x^2 *P*-value
Patient age (yr)						
< 50	167 (91)	16 (9)	1.47	426 (89)	54 (11)	0.613
≥ 50	151 (87)	22 (13)	0.22	580 (87)	85 (13)	0.43
Tumour size						
≤ 2 cm	190 (90)	22 (10)	0.059	606 (88)	82 (12)	0.13
> 2 cm	127 (89)	16 (11)	0.80	394 (87)	57 (13)	0.71
Tumour grade						
Grade 1	5 (100)	0 (0)	1.76	194 (86)	31 (14)	3.38
Grade 2	25 (83)	5 (17)	0.41	444 (87)	68 (13)	0.18
Grade 3	287 (90)	33 (10)		362 (90)	39 (10)	
Tubule formation						
1	3 (100)	0 (0)	0.37	77 (87)	12 (13)	0.44
2	47 (89)	6 (11)	0.28	379 (88)	54 (12)	0.799
3	262 (89)	32 (11)		526 (89)	68 (11)	
Mitotic count scores						
1	17 (94)	1 (6)	**10.6**	493 (87)	72 (12)	0.71
2	38 (76)	12 (24)	**0.005**	221 (89)	28 (11)	0.70
3	257 (91)	25 (9)		267 (89)	33 (11)	
Nuclear pleomorphism						
1	2 (100)	0 (0)	0.28	17 (94)	1 (6)	1.67
2	10 (91)	1 (9)	0.86	385 (87)	59 (13)	0.43
3	299 (89)	37 (11)		579 (89)	74 (11)	
Axillary nodal stage*						
Stage 1	226 (91)	21 (9)	**12.56**	659 (88)	91 (12)	2.21
Stage 2	77 (88)	10 (12)	**0.002**	296 (90)	37 (11)	0.33
Stage 3	14 (67)	7 (33)		45 (82)	10 (18)	
Nottingham prognostic index						
Good	18 (86)	3 (14)	**6.96**	408 (88)	55 (12)	0.40
Moderate	244 (92)	22 (8)	**0.031**	478 (87)	70 (13)	0.81
Poor	55 (81)	13 (19)		114 (89)	14 (11)	
Vascular invasion						
Negative	242 (91)	23 (9)	**4.64**	698 (87)	101 (13)	0.45
Positive	74 (83)	15 (17)	**0.029**	301 (89)	38 (11)	0.50
Progesterone receptor status						
Positive						
Negative	1 (100)	0 (0)	0.12	746 (88)	101 (12)	0.12
	304 (89)	37 (11)	0.72	247 (87)	36 (13)	0.72
HER2 status						
Positive	72 (81)	17 (19)	**8.15**	103 (86)	17 (14)	0.61
Negative	237 (92)	21 (8)	**0.004**	890 (88)	118 (12)	0.43

5.3.8 Associations between the ALDH1A1+SOX9+EpCAM+ BCSC phenotype and patient outcomes in BC

The presence of cells with the ALDH1A1+SOX9+EpCAM+ phenotype was positively associated with poorer BCSS and shorter DMFS in the entire BC cohort (P = 0.021 and P = 0.012, respectively). The ALDH1A1+SOX9+EpCAM+ phenotype was also positively associated with poorer BCSS and shorter DMFS in ER- tumours (P = 0.008 and P = 0.014, respectively) and HER2+ tumours (P = 0.014 and P = 0.044, respectively), but not in ER+ tumours ($P \geq 0.05$). Moreover, the presence of cells with the ALDH1A1+SOX9+EpCAM+ BCSC phenotype was positively associated with shorter DFI in the whole cohort (P = 0.022; Figures 5.12–5.14).

Furthermore, Cox-proportional hazards models including other known prognostic parameters (tumour size, tumour grade and patient age) showed that the ALDH1A1+SOX9+EpCAM+ BCSC phenotype was an independent predictor of poor BCSS in the entire BC cohort (P = 0.005, HR: 1.53, 95% CI: 1.14–2.20), ER- tumours (P = 0.01, HR: 1.96, 95% CI: 1.16–3.30) and HER2+ tumours (P = 0.026, HR: 1.83, 95% CI: 1.07–3.12; Table 5.9). However, when the Cox-proportional hazards models included nodal stage in addition to other prognostic parameters, the associations between the ALDH1A1+SOX9+EpCAM+ phenotype and all outcomes became insignificant ($P \geq 0.05$).

As the ALDH1A1+SOX9+EpCAM+ BCSC phenotype was associated with chemotherapy resistance, the effect of chemotherapy versus no chemotherapy in terms of BCSS, RFS and DMFS was assessed for patients whose tumours contained cells with the ALDH1A1+SOX9+EpCAM+ phenotype. BCSS, RFS or DMFS were not significantly different between the chemotherapy subgroup or

chemotherapy-naive subgroup among the patients with tumours containing ALDH1A1+SOX9+EpCAM+ cells ($P > 0.05$; data not shown).

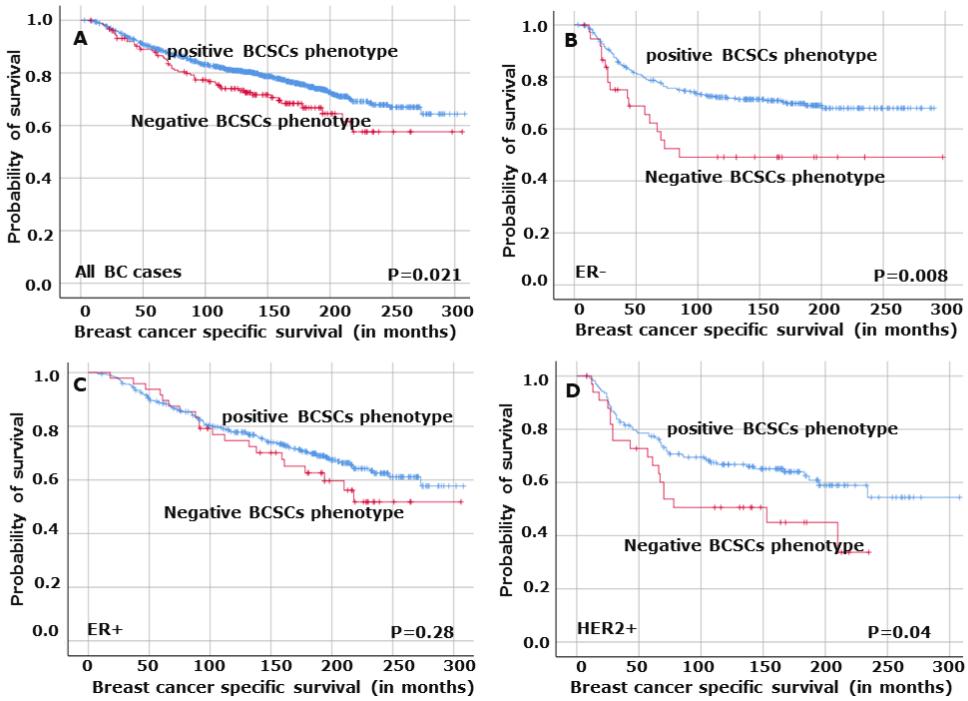


Figure 5.12: Associations between the ALDH1A1+SOX9+EpCAM cell phenotype and disease-free interval in (A) the entire cohort, (B) oestrogen receptor negative (ER-) tumours, (C) oestrogen receptor positive (ER+) tumours and (D) human epidermal growth factor receptor 2 positive (HER2+) tumours.

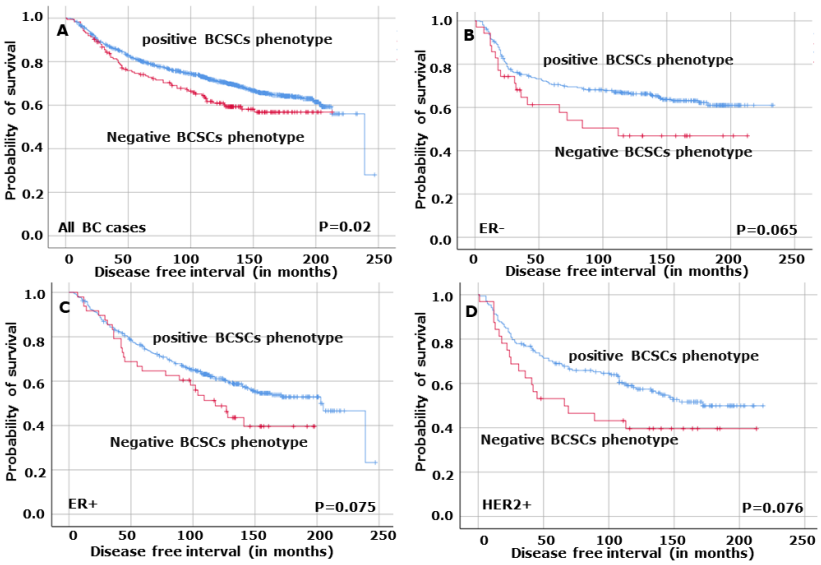


Figure 5.13: Associations between the ALDH1A1+SOX9+EpCAM cell phenotype and disease-free interval in (A) the entire cohort, (B) oestrogen receptor negative (ER-) tumours, (C) oestrogen receptor positive (ER+) tumours and (D) human epidermal growth factor receptor 2 positive (HER2+) tumours.

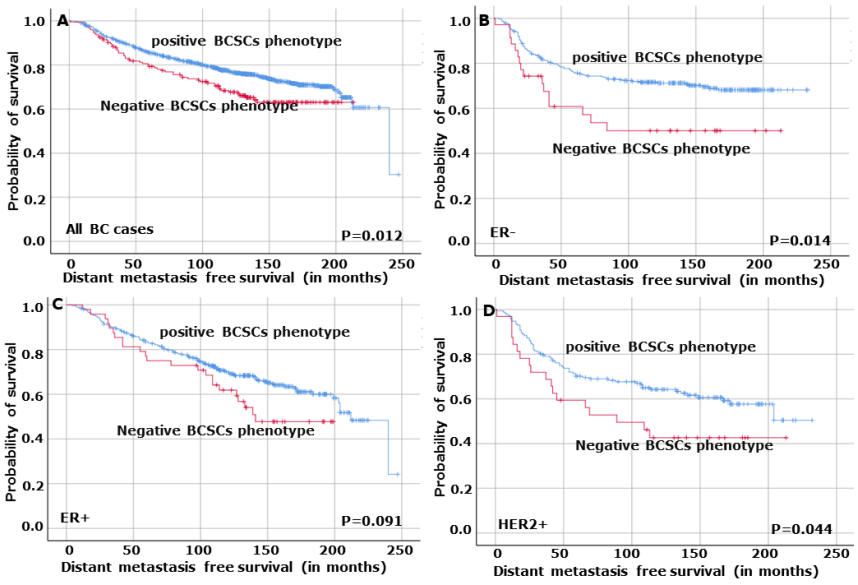


Figure 5.14: Associations between the ALDH1A1+SOX9+EpCAM+ cell phenotype and distant metastasis-free survival in (A) the entire cohort, (B) oestrogen receptor negative (ER-) tumours, (C) oestrogen receptor positive (ER+) tumours and (D) human epidermal growth factor receptor 2 positive (HER2+) tumours.

Table 5.9: Multivariate Cox regression hazard model of the associations between the ALDH1A1+SOX9+EpCAM+ cell phenotype and other prognostic clinicopathological features in the entire cohort and oestrogen receptor negative (ER-) and human epidermal growth factor receptor 2 positive (HER2+) subgroups.

	ALDH1A1+SOX9+EpCAM+ phenotype in entire cohort			**ALDH1A1+SOX9+EpCAM+ phenotype in ER- tumours**			**ALDH1A1+SOX9+EpCAM+ phenotype in HER2+ tumours**		
Parameters	***P*-value**	**HR**	**95% CI**	***P*-value**	**HR**	**95% CI**	***P*-value**	**HR**	**95% CI**
ALDH1A1+SOX9+ EpCAM+	0.005	1.533	1.14–2.06	0.011	1.964	1.16–3.30	0.026	1.83	1.07–3.12
Tumour size	0.694	1.043	0.84–1.28	0.490	1.143	0.78–1.73	0.098	1.45	0.93–2.25
Tumour grade	< 0.0001	1.888	1.59–2.23	0.812	1.066	0.62–1.81	0.012	2.04	1.17–3.57
Patient age (years)	0.978	0.978	0.80–1.22	0.369	1.189	0.81–1.73	0.534	1.19	0.74-1.78

5.4 Discussion

Despite receiving endocrine and chemotherapy regimens, 20–30% of patients with early-stage IBC subsequently develop metastatic disease [383, 384]. Tumour metastasis is a complex process initiated by aggressive tumour cells that can invade and migrate, survive in the circulation and form new lesions; not all tumour cells are able to metastasise [141]. BCSCs have attracted a lot of attention as they help to sustain tumour growth and are the driving force behind tumour invasion, metastasis, relapse and conventional therapy failure [385-387]. Scientists have used a variety of methods to identify and purify cells with BCSC phenotypes, such as flow cytometry, fluorescence-activated cell sorting, polymerase chain reaction and immunofluorescent staining [388-390]. However, the clinical utility of IHC analysis of patient tissue samples to detect BCSCs is still unclear.

Therefore, this study aimed to evaluate the expression of BCSC markers and detect cells with a BCSC phenotype in BC tissue samples using IHC. The expression of a panel of BCSC markers (ALDH1A1, CD133, CD44, CD24, EpCAM, SOX10, SOX9 and BMI-1) were evaluated individually. ALDH1A1 and SOX9 exhibited the highest expression and were detected in 71% and 52% of the study cohort, respectively. Jong *et al.* (2016) reported ALDH1A1 was expressed in 43% of BC cases, with expression ranging from weak to strong with H-scores of 51 to 300; tumours with negative expression had H-scores from 0 to 50 [46]. Similarly, high expression of ALDH1A1 ranging from weak to strong intensity was detected in BC cells in this thesis, as described in Chapter 3. Moreover, Bo Lei *et al.* (2012) reported SOX9 was highly expressed in 34% of their BC cohort;

positive tumours were considered to exhibit moderate to strong staining intensity in ≥ 10% of the tumour cells [47].

In contrast to previous studies, high BMI-1, SOX10 and CD44 were associated with a better prognosis in this BC cohort [291].

The BCSC markers tested in this study are highly expressed in tumour tissues, and BMI-1 and CD44 were associated with a better prognosis and less aggressive tumour features. The current study also provided some evidence to explain why these markers in combination, rather than individually, may possibly detect the BCSC phenotype in BC tissues. Several *in vitro* studies have demonstrated that combining BCSC markers (CD44+/CD24-/CD49f+ or $CD44^{+}/CD24^{-}$ $CD133^{+}$) provides a more accurate way to detect BCSCs than the individual markers alone [356, 391]. However, co-expression of CD44+CD24- was studied using IHC in BC samples to identify the BCSC phenotype with contradictory results, as the CD44-/CD24+ phenotype was prognostic for poor outcomes in early invasive BC [291].

This study revealed a positive correlation between the expression of ALDH1A1, EpCAM and SOX9 in BC, and indicates that these BCSC markers are co-expressed and could represent a BCSC phenotype. In addition, combined co-expression of ALDH1A1, SOX9 and EpCAM was more significantly associated with aggressive tumour features and poorer outcomes than the individual markers alone.

ALDH1A1 is a metabolic stem cell marker linked to many solid tumours and is involved in self-renewal, differentiation and evasion of the immune system [274, 275, 392, 393]. The role of ALDH1A1 is discussed in depth in Chapter 3.

EpCAM is a 314-amino acid polypeptide that contains a large N-terminal extracellular domain and short C-terminal cytoplasmic domain of 26 amino acids. EpCAM is an adhesion molecule that facilitates cell-cell interactions; however, EpCAM is involved in malignant transformation and overexpression of EpCAM can downregulate E-cadherin [394]. One study demonstrated that knockout of EpCAM significantly reduced cell migration and motility and increased adhesiveness in dendritic Langerhans cells, suggesting that EpCAM functions as a negative regulator of adhesion [395]. EpCAM has also been reported to be involved in tumour formation, tumour cell proliferation and differentiation, and plays important roles in stem cell regulation and interacts with β-catenin to activate W*nt* signalling [396-398]. Overexpression of EpCAM is associated with more advanced features and poorer outcomes in BC [399, 400].

SOX9, a member of the SOX family of transcription factors, belongs to the E protein subgroup, along with SOX8 and SOX10. SOX9 localises to both the cytoplasm and nucleus, and contains a DNA-binding high-mobility group (HMG) that enables nuclear translocation of SOX during development [401]. SOX9 contains two other functional domains, a dimerization domain (DIM) involved in homodimerisation with the HMG box of other SOX members and a transactivation domain (TAC) or C-terminal domain, which cooperates with transcription signalling or co-activators of transcription [402-404]. *In vitro* studies reported that knockdown of *SOX9* reduces the tumorigenicity of lung cancer cell lines [405]. Furthermore, overexpression of SOX9 enhances BC cell proliferation in highly metastatic BC [296, 406]. SOX9 also acts as a stem cell marker, promotes Wnt signalling and is associated with tumour progression in BC [407-409].

Therefore, it is important to explore the co-localisation of these markers to more precisely characterise the BCSC subpopulation and examine the potential role of BCSC phenotypes in different subtypes of BC. The double IHC staining protocol used in the current study showed that cells co-expressing ALDH1A1, SOX9 and EpCAM were present in 12% of the BC specimens overall. In addition, < 1% to 10% of the tumour cells were ALDH1A1+SOX9+EpCAM+, which aligns with the theory that BCSCs represent a small population of tumour cells [271, 356, 410], which suggest these cells presented the BCSC in the tissue sample. Moreover, the presence of the ALDH1A1+SOX9+EpCAM+ BCSC subpopulation was associated with advanced tumour stage and the highly aggressive ER- and HER+ subtypes, which confirms the BCSC properties. Matvey *et al.* (2019). suggested that the presence of ALDH+, CD44+ or CD24− cells and the CD44+CD24−CK18+EpCAM+CK19+ BCSC phenotype were strongly associated with metastasis in BC [411].

The ALDH1A1+SOX9+EpCAM+ BCSC subpopulation was associated with poorer BCSS, DFI and DMFS in BC patients, especially in the ER- and HER+ subgroups. Thus, the current study confirms the relationship between HER2 and CSC signalling pathways, which was reviewed by Deep *et al*. (2016) [412]. Osipo *et al.* (2008) demonstrated that inhibiting HER2 activated Notch 1 and resulted in trastuzumab resistance [413]. Moreover, Wnt signalling is persistently activated in HER2+ tumours [414]. In agreement with these reports, Sonic hedgehog (Shh) was expressed at high levels in the HER2+ subtype and Gli-1 was detected at high levels in the luminal B HER2+ subtype [156, 415].

In addition, among the subgroup whose tumours contained ALDH1A1+SOX9+EpCAM+ cells, patients who received chemotherapy and those

who did not receive chemotherapy had similar outcomes, which indicates that the ALDH1A1+SOX9+EpCAM+ BCSC subpopulation exhibits resistance to chemotherapy. Although the mechanisms that lead to chemoresistance in BC are poorly understood, the heterogeneity of BC cells produced by BCSCs could be one factor that promotes chemotherapy resistance [376, 416, 417].

5.5 Conclusion

This chapter indicates that BCSC phenotypes can be detected in tissue samples using IHC based on the co-expression of more than one BCSC marker. Furthermore, the ALDH1A1+SOX9+EpCAM+ BCSC phenotype may have prognostic value and be associated with tumour progression in BC, especially with regards to metastasis and recurrence of aggressive tumour subtypes. Overall, the assessment of co-expressing ALDH1A1+SOX9+EpCAM+ BCSCs during the treatment of BC may have significant predictive value.

Chapter 6

Refining Oestrogen Receptor (ER) Positivity in Invasive Breast Cancer

6.1 Introduction

The oestrogen receptor alpha (ERα) is a clinically useful prognostic and predictive marker in invasive breast cancer (IBC) [418]. Approximately 70–80% of BC tumours are ER-positive (ER+), and the standard systemic treatment for ER+ tumours is endocrine therapy [419, 420]. In routine clinical practice, the optimal threshold for ER positivity in IBC specimens has been adjusted over the years. A wide range of cut-off values have been historically used to define positive ER expression in IBC (0%, 1%, 5%, 10% and 20% ER+ cells); these optimal cut-offs were selected based on studies of the efficacy of endocrine therapy and patient outcomes [421-423].

Needle core biopsy (NCB) is currently taken to clinically assess ER expression using IHC. Hodi *et al.* (2007) reported strong agreement in the ER status of core biopsy and subsequent surgical excision specimen (SES) for 379 invasive tumours using a cut-off of 10% for ER positivity [418]. However, the highest level of concordance was observed for tumours with intermediate and high ER+ expression [424].

The 2010 guidelines from the American Society of Clinical Oncology (ASCO) and College of American Pathologists (CAP) recommended that the ER status of NCBs should be considered positive if 1% or more of the tumour cells

exhibit positive nuclear staining in IHC. In addition, tumours with ≥ 1% ER+ and PgR+ cells should be considered hormone receptor-positive (HR+) [425]. This low cut-off for hormone receptor-positive tumours may allow a larger number of patients to benefit from endocrine therapy [426]. However, many clinicians still employ a cut-off of 10% ER+ cells as the eligibility criterion for endocrine treatment [427].

The distribution of ER+ expression is bimodal, as the majority of invasive tumours show either completely negative ER staining (ER-negative) or ER positivity in more than 10% of cells [428]. Tumours with low ER+ expression (i.e., 1–9% ER+ cells) are rare, representing only 3% of BC cases [427]. However, this subgroup is clinically important, and there is controversy whether low ER+ tumours are more similar to luminal ER+ or basal ER- tumours [429]. The recently updated ASCO-CAP guideline recommends that a new reporting category, ER Low Positive, should be added to clinical assessment of ER status [430]. However, there are limited data regarding the pathological assessment of this subgroup and whether the 1% cut-off applies to NCBs or SESs. Thus, additional evidence is required before changing ER assessment in clinical practice.

6.1.1 Hypothesis

Although the current optimal threshold for positive ER immunostaining in IBC is 1% of tumour cells with ER-positive nuclei in NCBs or SESs, this study hypothesises that low ER+ expression (1-9%) in NCBs leads to misclassification of ER-negative tumours as ER+ tumours.

6.1.2 Aims of the chapter

The aims of this study are to assess low ER+ expression (1–9%) at both the protein (IHC) and mRNA (RNAscope) levels in NCBs and the corresponding SESs from BC patients.

6.1.3 Objectives

The objectives of this chapter are to:

(i) Assess ER immunostaining in low ER+ tumours (1–9%) and other comparison subgroups, including tumours with 10-20% ER+ cells and ER-negative tumours (< 1%), using NCBs and the corresponding SESs by IHC.

(ii) Examine the agreement between ER expression in NCBs and SESs in all subgroups of ER status.

(iii) Investigate interobserver agreement during assessment of ER+ expression between pathologists.

(iv) Define criteria for false positive patterns of ER+ expression in IHC.

(v) Assess ER expression in NCBs and SESs by IHC using different antibody clones.

(vi) Assess PgR expression in various ER subgroups.

(vii) Quantify *ESR1* and *PgR* expression using RNA *in situ* hybridisation and explore the associations with ER and PgR expression in IHC.

(viii) Compare the clinicopathological characteristics of low ER+ (1–9%) tumours with ER-negative (< 1%) and ER+ tumours (≥ 10%), including ER+ (10-20%) and ER+ (21 to 100%) tumours.

6.2 Materials and Methods

6.2.1 Patients and data collection

A patient database was used to identify patients with primary IBC treated at Nottingham City Hospital with known ER status based on analysis of NCBs, and undergo surgical excision. Patients with recurrent or metastatic BC were excluded.

A total of 10,550 patients with primary IBC were consecutively reported in routine practice. Of these 10,550 patients, 7,753 (73%) were ER+ (based on a cut-off of ≥ 1%) and 2,797 (27%) were ER-negative. Tumours with low ER expression (1–9%) represented 178 (2.2%) of the 7,753 ER+ cases (included patients who received neoadjuvant therapy).

Based on the availability of tissue specimens, a total of 1,743 patients with primary IBC were included in this study: 28 cases were ER-negative (< 1%), 77 cases had 1-9% ER+ cells and 1,638 cases had ≥ 10% ER+ cells, including 60 cases with 10–20% ER+ cells. None of the patients received neoadjuvant chemotherapy. Data on the clinicopathological parameters, patient outcomes and adjuvant therapy were retrieved for the study cohort; further details are provided in Chapter 2.

6.2.2 Pathological evaluation of ER expression

IHC staining was routinely performed by NHS laboratory staff at the Histopathology Laboratory at Queen's Medical Centre, Nottingham, which is certified and accredited by the CAP.

NCBs and SESs were routinely fixed in formalin for a maximum of 8 h and 48 h, respectively, according to the ASCO/CAP guidelines [431]. IHC for ER was performed on representative 4 μm-thick formalin-fixed paraffin wax-embedded (FFPE) sections. The EP1 rabbit monoclonal primary antibody was used from 2008 to 2013 and then replaced by the SP1 rabbit monoclonal primary antibody; the same immunohistochemical procedures were employed for both antibodies. IHC for PgR was conducted using the 1E2 rabbit monoclonal primary antibody clone.

An external control positive control (normal breast tissue) and negative control (tonsil) were performed for every run, and the run was repeated if any negative results appeared in the external positive control. The study also followed the standard practice of repeating the staining for ER-negative tumours if the internal control showed negative staining.

Multiple pathologists were involved in the ER and PgR assessments. ER and PgR expression were evaluated based on the percentages of positive tumour cell nuclei. The internal control was also reported, whether positive, negative or not presented. The ER and PgR scores from the original pathology reports were used in this study.

For this study, the percentages of cells expressing ER and PgR in the NCBs and SESs of cases with < 1-20% ER+ cells were re-reviewed by an assessor blinded to the original scores.

6.2.3 IHC staining

IHC staining and scoring for ER and PgR were repeated for a subset of the SESs (n = 80) depending on the availability of tissue blocks. Prior to IHC staining,

FFPE tissues were sectioned at 4 μm, and the internal control (normal breast tissue) was included on the SuperFrost Plus Slides.

IHC staining for ER and PgR were carried out by NHS laboratory staff at the Histopathology Laboratory of Queen's Medical Centre, Nottingham, using the VENTANA BenchMark ULTRA system (Tucson, AZ, USA) with VENTANA anti-ER (SP1) and anti-PgR (1E2) rabbit monoclonal primary antibodies, following the recommended protocol described in Chapter 2.

New sections were prepared for a subset of unselected cases (n = 10) with 1-20% ER+ cells and stained using two ER antibody clones (EP1 clone and SP1 clone); the SP1 antibody was also tested with and without the casein protein block step.

6.2.4 IHC assessment

ER and PgR were assessed by estimating the percentages of tumour cells exhibiting nuclear staining and classifying the pattern of staining, which included homogenous nuclear staining and granular/punctate nuclear staining. The intensity of staining (very weak, weak, moderate or strong) was also considered. The existence of non-specific staining, including background, cytoplasmic and stromal cell staining, were also reported, along with the internal control status.

6.2.5 RNAscope in situ hybridisation assay

RNAscope is an *in situ* hybridisation method used to detect specific RNA molecules in individual cells. In brief, 4 μm-thick FFPE tissue sections were dewaxed, rehydrated and probes for the target genes (*ESR1* and *PGR*) and

positive control (*PPIB*) and negative control (*dapB*) were hybridised at 40 °C for 2 h. A preamplifier was applied, followed by amplifiers, and the signals were detected using fast red chromogen, followed by haematoxylin counterstaining. Further details of the RNAscope conducted by Advanced Cell Diagnostics (ACD Bio) assay are provided in Chapter 2.

6.2.6 RNAscope assessment

The numbers of *ESR1* and *PgR* mRNA transcripts were determined by scoring the number of dots per cell at 20X magnification, and each cell was scored based on the specific criteria presented in Table 1.

The average score for approximately 100 randomly selected invasive tumour cells in each microscopic images was taken as the final score. The positive cut-off was a score ≥ 1, according to the manufacturer's guideline (https://acdbio.com/technical-support/solutions). Figure 6.1 shows representative images of tumours with each score.

Table 6.1: Criteria for RNAscope scoring.

Score	Criteria
0	No staining or <1 dot/10 cells
1	1-3 dots/cell
2	4-9 dots/cell; no or very few dots in clusters
3	10-15 dots/cell and < 10% dots in clusters
4	>15 dots/cell and > 10% dots in clusters

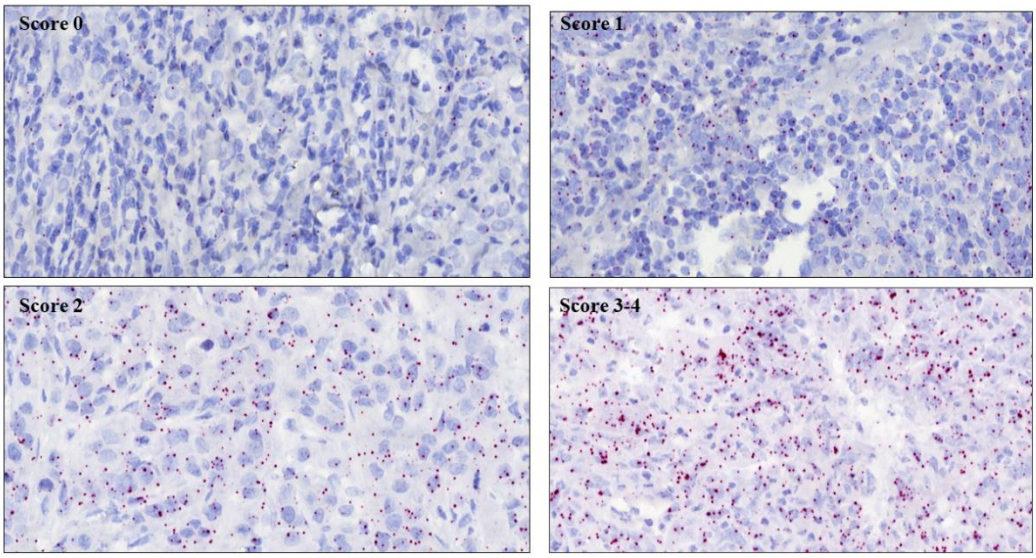


Figure 6.1: Shows representative images of ESR1 expression (score 0 to score 4) in breast cancer (BC) tissue sample using RNA-scope (40X).

6.2.7 Statistical analysis

IBM SPSS 24.0 (Chicago, IL, USA) software was used for the statistical analysis, as described in Chapter 2. The associations between the categorical groups of ER expression and clinicopathological parameters were analysed using the Chi-square test, while the associations with patient outcomes were assessed using Kaplan-Meier curves and the log-rank test. Inter-observer agreement was determined using the intra-class correlation coefficient (ICC).

6.3 Results

6.3.1 ER expression in NCBs and SESs

Tumours classified as having 1-9% ER+ cells in NCBs frequently showed differential expression in the SESs (i.e., ≥ 10% or 0% ER+ cells). Sixty cases were classified as having 1–9% ER+ cells in the NCBs. Of these, 26 (43%) were re-classified to 0% or < 1% ER+ cells and 34 (57%) were upgraded to ≥ 10% ER+ cells based on the SESs.

The majority of tumours (15/21; 71%) classified as < 1% ER+ cells in NCBs were reclassified as 0% based on the SESs; 11 (26%) of tumours with 10–20% ER+ cells in the NCBs were reclassified to ≥ 1% ER+ based on the SESs.

Table 2 summarises the ER expression scores for the NCBs and SESs for the following groups based on the NCBs: ER-negative (< 1%), low ER expression (1-9%), and 10-20% ER+.

Table 6.2: Oestrogen expression (ER) (percentage) in needle core biopsies (NCBs) and the matched surgical excision specimens (SESs).

	% ER+ cells in surgical excision specimens			
ER expression (%) in NCBs	**0% or <1%**	**1-9%**	**10-20%**	**> 20%**
< 1%	15 (71)	3 (14)	2 (10)	1 (5)
1-9%	26 (43)	20 (33)	7 (12)	7 (12)
10-20%	11 (26)	9 (21)	16 (37)	7 (16)

*The number of cases in each subgroup depended on the availability of clinical pathology scores for NCBs and SESs.

6.3.2 Interobserver agreement in ER scoring between pathologists

Diagnostic ER immunisations slides, including the NCBs (n=85), were re-evaluated, and SESs (n=85) were re-stained for ER and PgR. For cases with < 1–20% ER+ cells in the original NCBs, the ICC for agreement among the two pathologists was 0.910 and 0.844 for NCBs and SESs, respectively.

When the cases were divided into subgroups based on the percentage of ER+ cells in the NCBs, the ICC for NCBs was 0.407 for the low ER+ subgroup (1-9%) and 0.839 for the intermediate ER+ subgroup (10-20%). The ICCs indicated almost perfect agreement for the SESs in the 1–9% and 10–20% subgroups (0.920 and 0.989, respectively) and 100% agreement was achieved between the NCBs and SESs of the tumours with < 1% ER+ cells.

6.3.3 Characteristics of IHC staining

When the slides were evaluated, 39 of 68 (57%) cases with < 1–20% ER+ cells in the NCBs were reclassified to ER-negative (0% or < 1%) based on the SESs. More than half (65%) of the low ER+ tumours were reclassified to ER-negative (0 or < 1%) based on the SESs. Table 6.3 outlines the number of cases in each group reclassified as 0 or < 1%.

Interestingly, for the cases reclassified to 0% or < 1% based on the SESs, a variety of different staining patterns were observed in the NCBs, including specious/abnormal staining such as very weak to weak granular nuclear staining, or extremely faint nuclear staining with no obvious pattern (invisible pattern) in the presence or absence of nonspecific background staining, such as perinuclear cytoplasmic and stromal cell staining.

Overall, 61% of all low ER+ tumours exhibited predominantly abnormal staining patterns in the NCBs. These patterns were mainly granular staining with very weak intensity or very faint nuclear staining with a barely detectable pattern.

Only three of the reclassified cases showed a pure homogenous pattern of staining; 69% of the reclassified cases had an abnormal pattern with very weak to weak intensities and intense non-specific staining. Moreover, 88% of the ER-negative subgroup (< 1%) showed abnormal staining patterns in the NCBs. Figures 6.1 & 6.2 illustrate the different patterns, intensities and nonspecific ER staining observed in the NCBs.

Table 6.3: Number of cases in each subgroup reclassified to 0 or < 1% based on the surgical excision specimens (SESs).

% ER+ cells in NCBs	**Cases reclassified to 0% or <1% based on SESs**
	Number of cases (%)
< 1%	8 out of 9 (89)
1-9%	24 out of 37 (65)
10-20%	7 out of 22 (32)

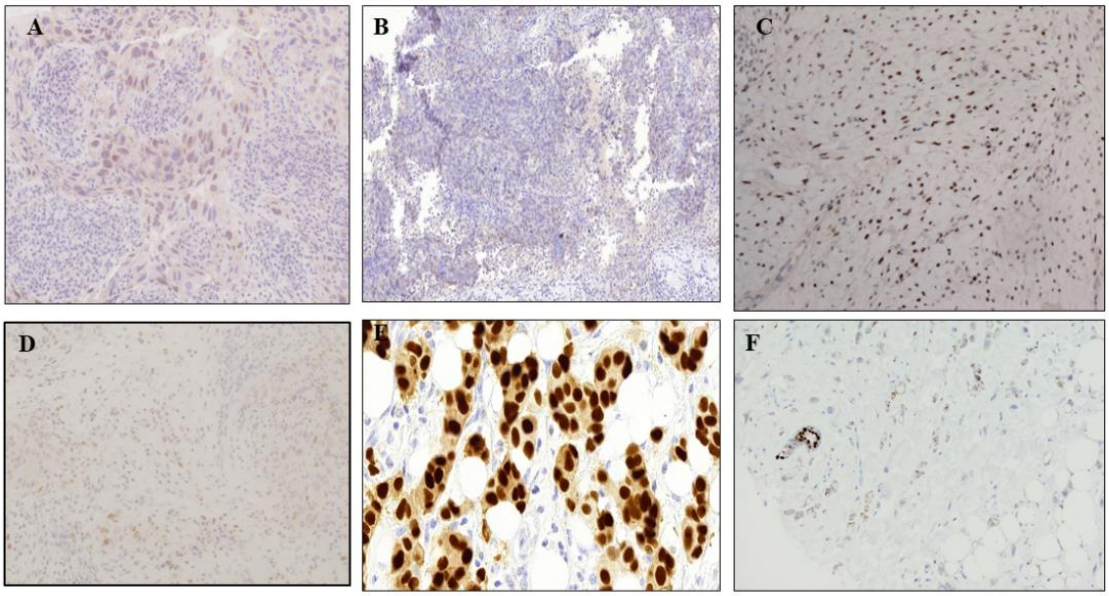


Figure 6.2: Representative examples of nonspecific oestrogen receptor (ER) staining. (A&B) Background staining with very weak to weak intensities. (C&D) Stromal cell staining. (E&F) Cytoplasmic staining (E an example for high nuclear staining with cytoplasmic staining of ER expression) (40×).

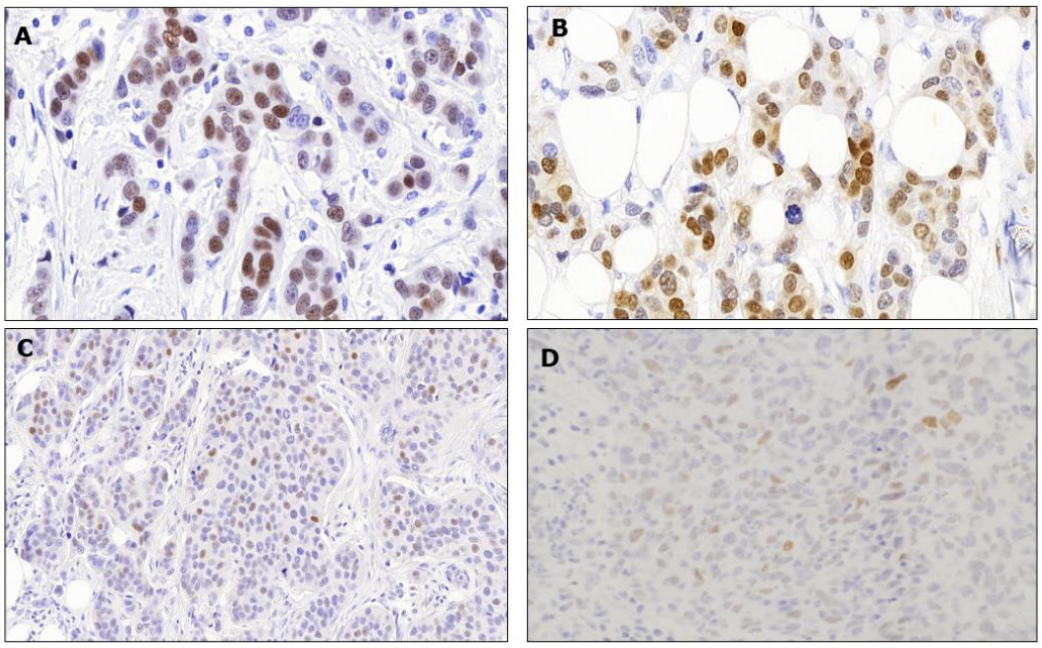


Figure 6.3: Representative examples of abnormal patterns of oestrogen receptor (ER) staining. (A&B) Granular nuclear pattern stain. (C&D) Very faint nuclear staining with no obvious pattern (40×).

6.3.4 Comparison of ER antibody clones

In order to investigate further, the NCBs and matching SESs of ten tumours were re-stained using two different ER antibody clones (EP1, SP1) and the SP1 antibody with and without a protein block.

A significant correlation in ER expression was observed between the two antibody clones in the NCBs and SESs (all $P < 0.0001$). Table 6.4 shows that the lowest percentages of ER+ cells were obtained for the SP1 antibody with a protein block, which indicates the protein block leads to lower levels of abnormal or nonspecific staining.; 66% of cases stained exhibited low percentages of ER+ cells with an abnormal pattern and non-specific staining.

In contrast, the samples stained with the EP1 antibody and the SP1 antibody without a protein block had higher expression scores and frequently exhibited abnormal staining patterns and nonspecific staining (Figure 6.3).

Table 6.4: Comparison of the numbers of ER+ cells for needle core biopsies (NCBs) and the matched surgical excision specimens (SESs) stained using the EP1 antibody and the SP1 antibody with/without a protein block.

NCBs, IHC category	NCBs *N* (%)			SESs *N* (%)		
	EP1	**SP1**	**SP1/protein block**	**EP1**	**SP1**	**SP1/protein block**
<1% or 0%	0 (0)	0 (0)	2 (22.2)	1 (11.1)	1 (11.1)	1 (11.1)
1-9%	2 (22.2)	4 (44.4)	3 (33.3)	1 (11.1)	4 (44.4)	4 (44.4)
≥ 10%	7 (77.7)	5 (55.5)	3 (33.3)	7 (77.7)	4 (44.4)	3 (33.3)

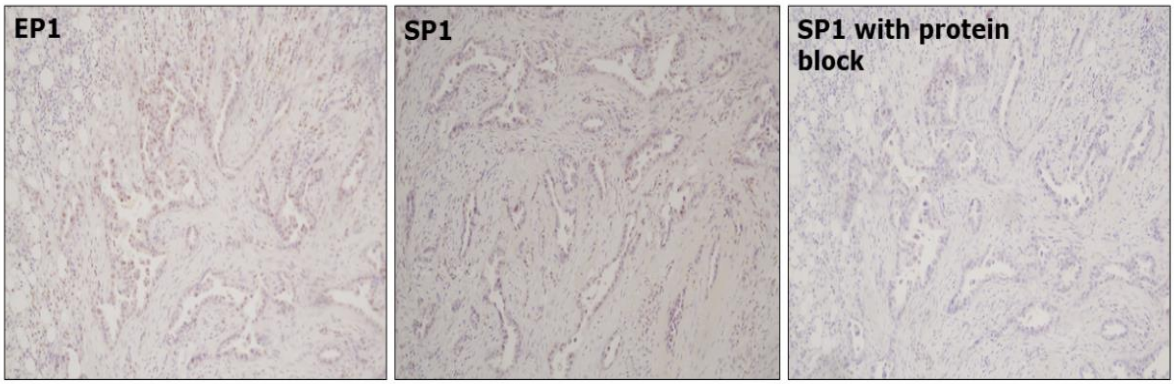


Figure 6.4: Representative examples of oestrogen receptor (ER) staining for different antibody clones (20×). The EP1 antibody resulted in weak cytoplasmic ER staining with some nuclear staining, the SP1 antibody resulted in low nuclear positivity, and the SP1 antibody with the protein block resulted in negative (ER) staining.

6.3.5 PgR expression in different ER+ subgroups

A high correlation was observed between the percentages of ER+ and PgR+ cells in the SESs ($P < 0.0001$). Overall, 92% of the low ER+ tumours (1-9% ER+ cells) were PgR-negative (0 or < 1% PgR+ cells).

Of the 31 cases reclassified to 0% or < 1% ER+ cells based on the SESs, 99% were PgR-negative; only one case was positive for PgR. Table 6.5 summarises PgR expression in each ER+ subgroup.

Table 6.5: Progesterone receptor (PgR) expression in different oestrogen receptor positive (ER+) subgroups.

ER expression, % cells in NCBs	**ER in SESs *N* (%)**			**PgR in SESs* *N* (%)**	
	<1% or 0	1-9%	≥ 10%	Positive	Negative
< 1%	5 (71.4)	2 (28.5)	0 (0)	2 (28.5)	5 (71.4)
1-9%	23 (92)	1 (4)	1 (4)	2 (8)	23 (92)
10-20%	8 (42.1)	5 (26.3)	6 (31.5)	5 (26.3)	14 (73.6)

*PgR cut-off (>1%)

6.3.6 Association between the RNAscope and IHC results for ER and PgR

There was a high level of agreement between the IHC and RNAscope scores of the SESs for both ER/PgR and *ESR1/PgR* ($P < 0.0001$). However, one case was negative for ER in IHC (0%) and *ESR1*-positive in the RNAscope assay (≥ 2 score). A total of 71% low ER+ (1-9%) tumours were *ESR1*-positive, while 100% of the reclassified cases were ESR1-negative. Table 6.6 summarises the agreement in the assessment of ER status between the RNAscope assay and IHC evaluations.

Table 6.6. Comparison of the immunohistochemistry analysis of oestrogen receptor expression and RNAscope assay of *ESR1* status for surgical excision specimens (SESs).

NCBs, IHC category	SESs, IHC *N* (%)		SESs, RNAscope *N* (%)	
	Positive	**Negative**	**Positive**	**Negative**
< 1%	0 (0)	3 (100)	0 (0)	3 (100)
1-9%	13 (48.1)	14 (51.8)	14 (51.8)	13 (48.1)
10-20%	11 (84.6)	2 (15.3)	11 (84.6)	2 (15.6)

6.3.7 Comparison of the clinicopathological characteristics and outcomes of the low ER+ subgroup and other subgroups

Tables 6.7 and 6.8 compare the clinicopathological characteristics and outcomes of the patients with low ER+ (1%–9%) and ER-negative (< 1%) or ER+ (≥ 10%) tumours based on analysis of the NCBs. There was no significant difference in any clinicopathological feature of the patients with low ER+ tumours (1–9%) and ER- tumours (< 1%), including patient age ($P = 0.362$),

tumour size (P = 0.141), tumour grade (P = 0.822), NPI (P = 0.971), histological type (P = 0.335), nodal stage (P = 0.835), HER2 (P = 0.349) and PgR (P = 0.047).

However, significant differences were observed between the 1–9% ER+ subgroup and ≥ 10% ER+ subgroup. Compared to the ≥ 10% ER+ subgroup, the 1–9% ER+ subgroup were more likely to have larger tumours (7% versus 3%, P = 0.002), advanced histological grade (10% versus 0.9%, P < 0.0001), poor Nottingham prognostic index (5% versus 1%, P < 0.0001), mixed NST lobular carcinoma (83% versus 2%, P < 0.0001), HER2-positive tumours (55% versus 3%, P < 0.0001), and PgR-negative tumours (11% versus 1%, P < 0.0001).

To further investigate the associations between ER status and the clinicopathological features of BC, the ≥ 10% subgroup was divided into two subgroups: ER+ (10–20%) and ER+ (21–100%). No significant difference was detected for any clinicopathological parameter between the 10–20% and 21–100% ER+ subgroups, including patient age (P = 0.347), tumour size (P = 0.162), tumour grade (P = 0.322), NPI (P = 0.954), histological type (P = 0.335), nodal stage (P = 0.835) or LVI (P = 0.537).

When the analysis was limited to false positive cases, false positive ER expression was more frequently observed among older patients (66%), grade 3 tumours (87%), NST histological type tumours (68%), HER2-negative tumours (69%), and PR-negative tumours (93%).

Table 6.7: Comparison of the clinicopathological features of the < 1% Oestrogen positive (ER+) versus 1-9% Oestrogen positive (ER+) subgroups.

Characteristic	% ER+ cells in NCBs based on IHC		x^2	*P*-value
	<1%	**1-9%**		
Age at diagnosis (years)				
< 50	11 (32.4)	23 (67.6)	0.831	0.362
≥ 50	17 (23.9)	54 (76.1)		
Tumour size (cm)				
< 2	18 (32.7)	37 (67.3)	2.169	0.141
≥ 2	10 (20)	40 (80.0)		
Tumour grade				
Grade I	1 (25)	3 (75.0)	0.392	0.822
Grade II	4 (21.1)	15 (78.9)		
Grade III	23 (28)	59 (72)		
Mitotic count				
1	5 (25)	15 (75.0)	0.968	0.616
2	10 (33.3)	20 (66.7)		
3	13 (23.6)	42 (76.4)		
Nuclear pleomorphism				
1	0 (0.0)	1 (100)	0.501	0.778
2	2 (33.3)	4 (66.7)		
3	26 (26.5)	72 (73.5)		
Tubule formation				
1	1 (33.3)	2 (66.7)	3.106	0.212
2	2 (10.5)	17 (89.5)		
3	25 (30.1)	58 (69.9)		
Nottingham prognostic index				
Good Prognostic Group	3 (25)	9 (75)	0.059	0.971
Moderate Prognostic Group	20 (27.4)	53 (72.6)		
Poor Prognostic Group	4 (25)	12 (75)		
Histological types				
Ductal (mixed)	18 (25)	54 (75)	5.710	0.335
Lobular	1 (25)	3 (75)		
Miscellaneous	4 (66.7)	2 (33.3)		
Mixed NST and lobular	1 (16.7)	5 (83.3)		
Special type	4 (25)	12 (75.0)		
Tubular	0 (0)	1 (100)		
Axillary nodal stage				
Stage I	20 (27)	54 (73)	0.361	0.835
Stage II	4 (21.1)	15 (78.9)		
Stage III	3 (30.0)	7 (70.0)		
Lympho-vascular Invasion				
Negative/probable	24 (28.9)	59 (71.1)	1.202	0.548
Definite	4 (19)	18 (81.0)		

Table 6.8: Comparison of the clinicopathological features of the 1-9% Oestrogen positive (ER+) versus 10-100% Oestrogen positive (ER+) subgroups.

Characteristic	% ER+ cells in NCBs based on IHC			
	1-9%	**10-100%**	**x^2**	***P*-value**
Age at diagnosis (Years)				
< 50	23 (5.1)	429 (94.9)	0.485	0.282
≥ 50	54 (4.3)	1203 (95.7)		
Tumour size (cm)				
< 2	37 (3.4)	1067 (96.6)	**9.753**	**0.002**
≥ 2	40 (6.6)	563 (93.4)		
Tumour Grade				
Grade I	3 (0.9)	318 (99.1)		
Grade II	15 (1.9)	780 (98.1)	**63.327**	**<0.0001**
Grade III	59 (10)	532 (90)		
Mitotic count				
1	15 (1.5)	965 (98.5)	**58.822**	**<0.0001**
2	20 (6.0)	312 (94.0)		
3	42 (11)	340 (89.0)		
Nuclear pleomorphism				
1	1 (3.3)	29 (96.7)	33.243	**<0.0001**
2	4 (0.7)	599 (99.3)		
3	72 (6.8)	991 (93.2)		
Tubule formation				
1	2 (1.3)	148 (98.7)	**9.785**	**0.008**
2	17 (3.1)	529 (96.9)		
3	58 (77)	942 (94.2)		
Nottingham Prognostic Index	9 (1.3)	682 (98.7)	**27.099**	**<0.0001**
Good Prognostic Group	53 (6.8)	726 (93.2)		
Moderate Prognostic Group	12 (5.2)	217 (94.8)		
Poor Prognostic Group				
Histological types				
Ductal (mixed)	54 (3.8)	1364 (96.2)		
Lobular	3 (1.7)	169 (98.3)	122.414	**<0.0001**
Miscellaneous	2 (40)	3 (60)		
Mixed NST and lobular	5 (83.3)	1 (16.7)		
Special type	12 (11.9)	89 (88.1)		
Tubular	1 (25)	3 (75)		
Axillary nodal stage				
Stage I	54 (5)	1025 (95)	2.663	0.264
Stage II	15 (3.2)	460 (96.8)		
Stage III	7 (4.7)	141 (95.3)		
Lymphovascular invasion				
Negative/probable	59 (4.7)	1184 (95.3)	22.030	**<0.0001**
Definite	18 (3.7)	444 (96.3)		

6.3.8 Patient outcome analysis

The univariate analysis showed that patients with 1–9% ER+ cells had similar survival and recurrence rates as patients with ER-negative (< 1%) tumours (*P* = 0.946 and 0.140, respectively). However, patients with 1–9% ER+ cells had shorter recurrence-free survival (DFI) compared to patients with ≥ 10% ER+ cells (*P* = 0.034**;** Figure 6.4). There were no significant differences in BCSS and DFI (*P* = 0.455 and *P* = 0.287, respectively) between the low ER+ (1-9%) and ER+ (10-20%) subgroups.

Endocrine therapy or chemotherapy had no significant impact on BSCC or DFI in the low ER+ group; however, this analysis was based on a small subgroup of patients (*P* ≥ 0.05).

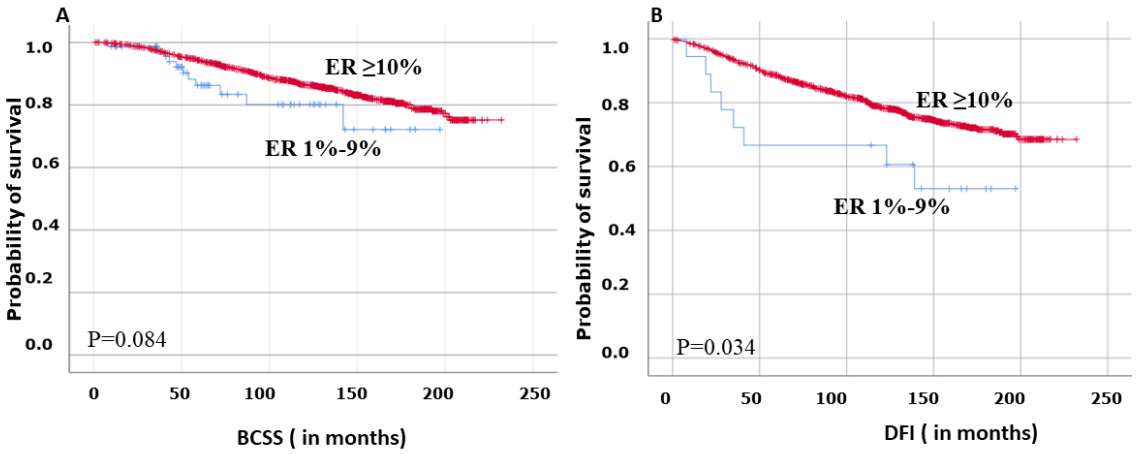


Figure 6.5: Kaplan-Meier survival plots of (A) breast cancer-specific survival (BCSS) and (B) recurrence free survival (DFI) for patients with 1-9% Oestrogen positive (ER+) cells and ≥ 10% Oestrogen positive (ER+) cells.

6.4 Discussion

Evaluation of ER status is a fundamental aspect of the pathological assessment of patients with BC. Decisions on BC treatment are determined based on ER status [432-434]. The current cut-off value for ER expression is 1% or more tumour cells exhibiting positive nuclear staining in IHC [425]. However, debate continues regarding whether tumours with 1–9% ER+ cells have greater molecular and clinical similarity to tumours with ≥ 10% ER+ cells or ER-negative (< 1%) tumours [435-437]. Several studies have reported assessment of low ER+ in SESs [429, 437], while other studies indicated that NCBs can be used to accurately evaluate ER status in clinical routine practice [438, 439]. However, Dekker *et al.* (2013) suggested tumours with ER-negative NCBs should be re-evaluated using SESs [440].

The current study indicates that only 2% of ER+ tumours contain 1–9% ER+ cells, in agreement with previous studies [427, 429]. This reflects the nature of ER expression, as the majority of tumour cells are either ER-negative or show strong positive immunoreactivity. Collins *et al.* (2005) reported that 99% of the BC cases in their study were either completely negative or strongly positive [441]. Muftah *et al.* (2016) provided further evidence of the bimodality of ER expression at both the mRNA and protein levels in invasive BC [442].

In this study, ER expression was examined in NCBs and the corresponding SESs. Low inter-observer agreement was observed between pathologists in the assessment of NCBs from low (1-9%) ER+ tumours. A previous study reported perfect agreement in evaluation of ER expression between four pathologists, though most cases exhibited strong, diffuse expression. However, two cases contained < 10% ER+ cells; while two pathologists scored these tumours as

positive (H-score = 1-7), the other pathologists scored these cases as negative (H score <1), which provides further evidence of the poor agreement among pathologists in the assessment of low ER expression [443].

Moreover, more than half of the low ER+ expressing NCBs in this study were reclassified as negative based on analysis of the corresponding SESs. A high incidence of specious staining giving the impression of positive staining was observed in the NCBs. Thus, to characterise the factors associated with a high probability of false positive staining in low ER+ tumours, the features of ER staining in NCBs associated with negative results in SESs were investigated. Rakha *et al.* (2012) reported 47% of low ER+ tumours were ER-negative (0 or < 1%) in the SESs, and they observed a high rate of false positive ER staining and nonspecific immunostaining in low ER+ tumours [444].

To further investigate this issue, this study assessed the results of different IHC methods in low ER+ expressing tumours using two antibodies and one antibody with and without a protein block. The EP1 antibody clone stained higher percentages of cells compared to the SP1 antibody with or without the protein block, in agreement with a previous study [445]. The EP1 clone has a high sensitivity for detecting the ER antigen. Badve *et al.* (2013) demonstrated the EP1 clone resulted in strong staining intensity and led to 10% higher percentages of positive tumour cells for the same tumours compared to the SP1 antibody [420]. However, the results presented in this chapter suggest that the high expression obtained using the EP1 antibody is nonspecific staining and false positive ER expression.

Use of the SP1 antibody clone with a protein block identified fewer ER+ cells, with less false positive staining, than the other antibody test conditions.

Protein blocks prevent nonspecific antibody binding. This study confirmed that false positive and nonspecific staining were blocked by using a protein block. Thus, the current study suggests that using the SP1 antibody clone with a protein block reduces false positive ER expression.

The IHC findings were validated using the RNAscope assay to detect *ESR1* mRNA molecules in individual cells and confirm the detection of positive ER expression using IHC. A high concordance was observed between the RNAscope analysis of *ESR1* and PgR mRNA and IHC analysis of ER and PgR expression. One case that was ER-negative in the IHC had a positive RNAscope score; this case was PgR-negative and the internal controls for both techniques were positive. It is possible the ER was expressed at the mRNA level but not at the protein level due to post transcriptional modifications of the ER. A previous study reported high agreement between the RNAscope analysis and IHC evaluation of the ER in 72 cases [446].

Moreover, the low ER+ cases reclassified as 0 or < 1% based on the SESs also exhibited negative results in the RNAscope assessment of *ESR1* mRNA in the SESs, which supports the IHC analysis of the SESs. This suggests that SESs provide a more reliable evaluation of low ER+ tumours than NCBs. Low ER+ tumours have been reported to exhibit heterogonous in biology and behaviour, and assessment of SESs was previously suggested to provide more accurate assessment of low ER tumours than assessment of NCBs [430].

Conversely, a previous study showed that NCBs provide more reliable assessment of ER status than SESs, due to pre-analytic factors that may lead to false negative results in SESs [447]. However, 97% of the SESs in this study

had a positive internal control (normal breast tissue) and the optimal tissue handling requirements described by the ASCO/CAP guideline were followed.

A previous study reported that 9% of ER+ tumours do not express the PgR [448]. However, this phenomenon is not fully understood. Horwitz *et al.* (2014) suggested that the absence of the PgR was due to loss of the ER [449]. In this study, the majority of tumours with low ER+ expression were PgR-negative, in agreement with another study [436]. The PgR is a well-characterised ER-dependent protein and the presence of the PgR indicates an intact ER response pathway [450]. A recent study demonstrated that tumours with 1–9% ER+ cells were associated with advanced clinicopathological features and poor survival [428]. Our findings confirm that tumours with 1–9% ER+ cells and ER-negative tumours exhibit similarly aggressive tumour features and poor survival rates. Moreover, patients with low ER+ expressing tumours had poorer tumour features and outcomes compared to patients with ≥ 10% ER+ tumours. Overall, this finding suggests that 1–9% ER+ tumours are more similar to ER-negative tumours and can be frequently misclassified as ER+ based on the NCBs.

6.5 Conclusion

Low ER (1–9%) BC NCBs have a high probability of being misclassified ER-negative tumours as ER-positive tumours. SESs are superior to NCBs for assessment of low ER+ expression. NCBs with false positive IHC staining patterns PgR-negative, and high-grade tumours should be re-evaluated and clinical decision-making should be based on assessment of the subsequent SESs.

Chapter 7: Quantification and clinical impacts of intratumor ER heterogeneity in invasive breast cancer

7.1 Introduction

Patients with ER immunopositive (ER+) tumours have a better outcome compared to patients with other molecular subtypes of BC. However, approximately 30-50% of patients with ER+ invasive BC (IBC) do not respond to endocrine therapy and suffer disease progression, and disease recurrence [451], and 30-40% of patients with ER+ IBC develop distant metastasis [452].

Current clinical ER immunohistochemistry (IHC) evaluations do not adequately consider the heterogeneity of ER expression and/or quantify the levels or patterns of ER expression [453]. The recent ASCO/CAP guideline for assessment of ER status only considers the percentage of ER+ tumour cells, and does not consider any further criteria beyond a threshold for positive ER expression [454].

Gulisa *et al* (2017) reported that the presence of varied tumour cell populations within and between BC tumours may affect the response to therapy and patient outcomes [455]. Specific cell populations within a tumour mass could undergo selection for outgrowth during treatment and emerge as a resistant tumour. Intratumoural heterogeneity within tumours that exhibit both ER+ and negative immunostaining may potentially affect tumour progression in primary BC [456]. The failure of endocrine therapy in ER+ BC may indicate that only the ER+ subpopulations of cells respond to endocrine therapy, whereas tumours with

substantial ER negative (ER-) subclones do not exhibit a complete clinical response [456, 457].

Inic *et al* (2014) reported the luminal A and luminal B subtypes of BC were both ER+; however, patients with the luminal A and luminal B BC subtypes exhibited different patterns of tumour progression and outcomes [458]. Using a whole genome method, Sherene *et al* (2008) discovered and validated a subset of genes involved in the risk of distant metastasis in ER+ tumours, which included high expression of cell cycle genes [459].

More detailed assessments of intratumoral variability in the intensity and patterns of ER expression in tumours with the same percentage of ER+ tumour cells may potentially provide evidence of the biological factors that underlie non-effective clinical responses to endocrine therapy in patients with BC [453].

The progesterone receptor (PgR) is a nuclear receptor, and a well-characterised ER-regulated gene. PgR interacts with the ER in BC cells and, when activated by its ligand, it can redirect ER chromatin binding [460]. PgR is also a fundamental diagnostic marker in BC [461].

Rakha *et al* (2007) reported that 71% of BC were ER+ and 59% were PgR+; 55% were both ER+/PgR+, 16% were ER+/ PgR-, and 3.4% were ER-/PgR+ [462]. Kurozumi *et al* (2017) stated that PgR expression could potentially enable identification of patients with ER+ BC who may have a poorer prognosis and outcome [463].

This evidence suggests that assessing the intratumoral heterogeneity of ER expression in BC could provide additional biological insight into BC, and could potentially improve clinical decision making in ER+ BC.

7.1.2 Hypothesis

Intratumoral heterogeneity in ER expression, and the patterns of ER+ expression within the same tumour has a biological impact on tumour behaviour and response to therapy in IBC.

7.1.3 Aims of the study

The aims of this chapter are to (i) assess the heterogeneity of ER expression in primary IBC and the paired lymph node metastases and (ii) explore the associations between intratumoral ER heterogeneity and tumour behaviour and patient outcomes.

7.1.4 Objectives of the study

The objectives of this chapter are to:

(i) Comprehensively assess intratumoral ER heterogeneity by determining the percentages of ER-negative subclones and ER-positive subclones with very weak, weak, moderate and strong staining intensities.

(ii) Classify the patterns of ER+ expression in primary BC specimens.

(iii) Assess the percentages of cells expressing the PgR, an ER-related marker, in the ER+ and ER- subclones of primary BC specimens.

(iv) Compare intratumoral ER heterogeneity in primary BC and matched lymph node (LN) metastases.

(v) Quantify intratumoural ER heterogeneity index for primary BC specimens and matched LN metastases.

(vi) Investigate the clinical impact of the intratumoural ER heterogeneity index in BC.

7.2 Materials and Methods

7.2.1. Study cohort

Data on patients with primary IBC for which the ER expression status and lymph node status had been pathologically assessed were extracted from the patient database of Nottingham City Hospital.

Two patient cohorts were retrospectively selected: (i) patients with primary tumours with 1-70% ER+ cells (cases more than 70% of ER expression showed homogenous of ER expression (ER=100), when they were restained, and the aim in this study is investigate the heterogenous tumour with high proportion of negative clone) (*n* = 111), for which paired lymph node metastases were available for a subgroup of patients (*n* = 37), and (ii) a comparison cohort of patients with primary tumours that contained 100% ER+ cells (*n* = 391).

For both cohorts, the age at diagnosis ranged from 26-85 years and the median age was 52 years. The majority of patients had grade 2 and 3 (74%) BC and 29% had nodal stage 2 or 3 disease. Most patients (92%) had ductal carcinoma of non-special type (ductal/NST). Overall, 30% of patients failed to respond to endocrine therapy and developed disease recurrence and/or distant metastasis. Table 3 (in Chapter 2) summarises the clinicopathological features of the patient cohorts assessed in this study.

7.2.2 IHC staining

IHC staining for ER and PgR were performed on the tumours with 1-70% ER+ cells (*n* = 111). IHC staining of sequential sections was carried out at the Histopathology Laboratory, Queen's Medical Centre, Nottingham, by the NHS laboratory technicians. The sections were stained using a VENTANA BenchMark

ULTRA system (Tucson, AZ, USA) with VENTANA anti-ER (SP1) and anti-PgR (1E2) rabbit monoclonal primary antibodies, following the recommended protocol described in detail in Chapter 2.

7.2.3 Assessment of intratumoural ER heterogeneity

The slides were scanned into high-resolution digital images at 20× magnification using 3D Histech Panoramic 250 Flash II scanner (3DHISTECH Ltd., Budapest, Hungary). The digital images were assessed for ER and PgR using CaseViewer software.

The assessments were conducted on full-face sections in which tumour cells represented at least 40% of the section surface area. The assessments were limited to invasive carcinoma; carcinoma *in situ* and normal ducts were excluded from the assessments.

Only nuclear ER and PgR immunostaining were assessed. Non-specific staining, including cytoplasmic and stromal cells (fibroblasts and immune cells) were not considered, if present.

To assess intra-slide (intratumoural) ER and PgR heterogeneity, the digital images were annotated to divide the tumour area into four equal quadrants (regions) using vertical and horizontal lines.

The percentages of ER- cells and ER+ cells were visually assessed in each of the four quadrants, as well the whole slide image (whole tumour).

Moreover, the percentages of positive cells with each staining intensity (very weak, weak, moderate and strong) were also evaluated. The thresholds for each staining intensity were determined prior to the assessment, based on a quick scan of all of the tumour slides (Figure 1).

The distribution and intensity of ER+ and PgR+ staining in each tumour and quadrant were also classified into one of five groups: **uniform pattern,** all tumour cells stained positive with the same intensity; **scatter pattern,** positive cells dispersed surrounded by negative cells; **cluster pattern**, positive cells with the same staining intensity clustered together; **focal pattern,** positive cells sharply separated from negative cells; or **diffuse pattern,** positive cells diffused throughout the whole slide or quadrant [464]. The patterns are clarified in figure 7.6 in result section below.

ER and PgR expression were assessed in parallel using sequential sections for each case. ER expression was compared to PgR expression at both the tumour and regional levels for all criteria, i.e., the percentages of ER- cells, very weak, weak, moderate and strong ER+ cells and the expression pattern.

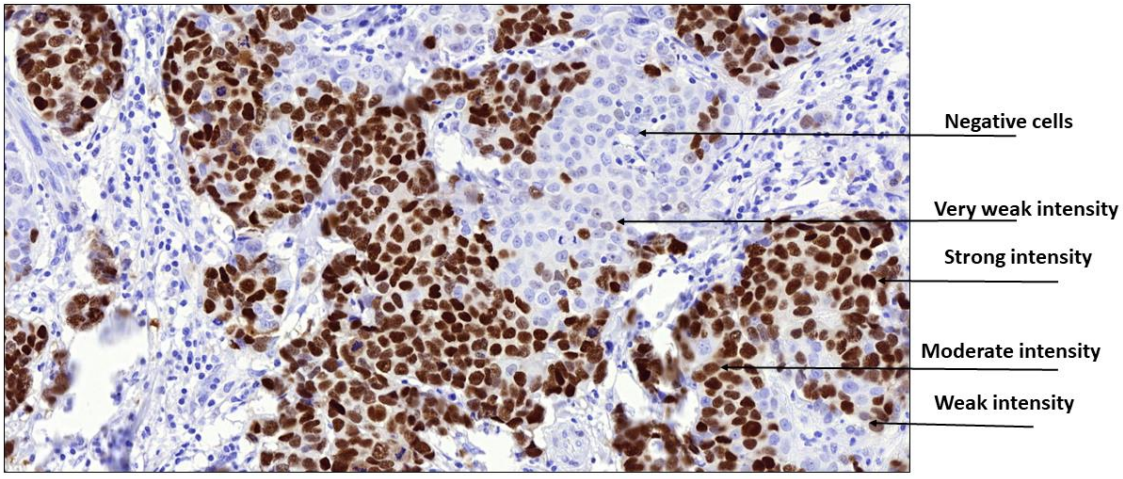


Figure 7.1: Intratumoral ER heterogeneity in primary BC. ER expression was visually determined. Distinct regions of ER- and ER+ cells were observed within individual tumours. Moreover, varied intensities of ER+ staining (very weak, weak, moderate, or strong) were observed within individual tumours.

7.2.4 Intratumoral ER heterogeneity index

Rao's quadratic entropy (QE; continuous score) based on the Simpson index was used to quantify the heterogeneity of ER expression at the tumour and regional (quadrant) levels [453, 465]. Rao's QE combined with a distance matrix is widely used to measure diversity, clonal diversity and genomic intratumour heterogeneity in cancer [466-468].

In the current study, the Simpson index was defined as the proportion of tumour cells that stained for ER at each intensity level: 0 (no stain/negative), 1 (very weak positive), 2 (weak positive), 3 (moderate positive) or 4 (strong positive).

The distance matrix was used to weigh the differences between the ER intensities within the same tumour or region. RStudio 3.6.0 software was used to calculate the QE. The index was performed by Dr Dimitrios, a bioinfomatician and one of our collaborator, in Nottingham Trent University.

7.2.5 Statistical analysis

IBM SPSS 25.0 software (Chicago, IL, USA) was used for statistical analysis, as described in Chapter 2. The associations between the categorical groups of ER heterogeneity and clinicopathological parameters were analysed using the Chi-square test, while the associations with patient outcome were assessed using Kaplan-Meier curves and the log-rank test. The Friedman test was used to compare the four regional scores, as well the expression of ER and PgR. Correlations between continuous variables were assessed using the Spearman's

rank test. The cut-offs were determined using x-tile Bioinformatics software version 3.6.1 (Yale University, USA).

7.3 Results

7.3.1 Intra-tumour heterogeneity of ER expression in BC

Intratumoral ER heterogeneity was assessed in full-face sections of 111 primary BC tumours (one section/case) and the corresponding 444 regions (quadrants) of the same tumours.

At the regional level, negative ER immunostaining was detected in at least in one region of every primary tumour; the percentage of ER- cells ranged from < 1% to 100%. Moreover, significant differences in the percentages of ER-, very weak ER+, weak ER+, moderate ER+ and strong ER+ cells were noted between the four regions ($P < 0.0001$). The heatmap in Figure 3 illustrates the variation in ER expression, including the negative and positive clones with different staining intensities between the four regions.

At the tumour level (whole slide), the majority (75%) of tumours with 1-20% ER+ cells showed weak and very weak ER staining intensities; strong and moderate staining intensities were not detected in this subgroup. The full spectrum of staining intensities, including strong, moderate, weak, and very weak ER+ staining, were observed within the majority (92%) of tumours with 21-70% ER+ cells. Figure 4 shows the distribution of the ER staining intensities in the two subgroups of tumours.

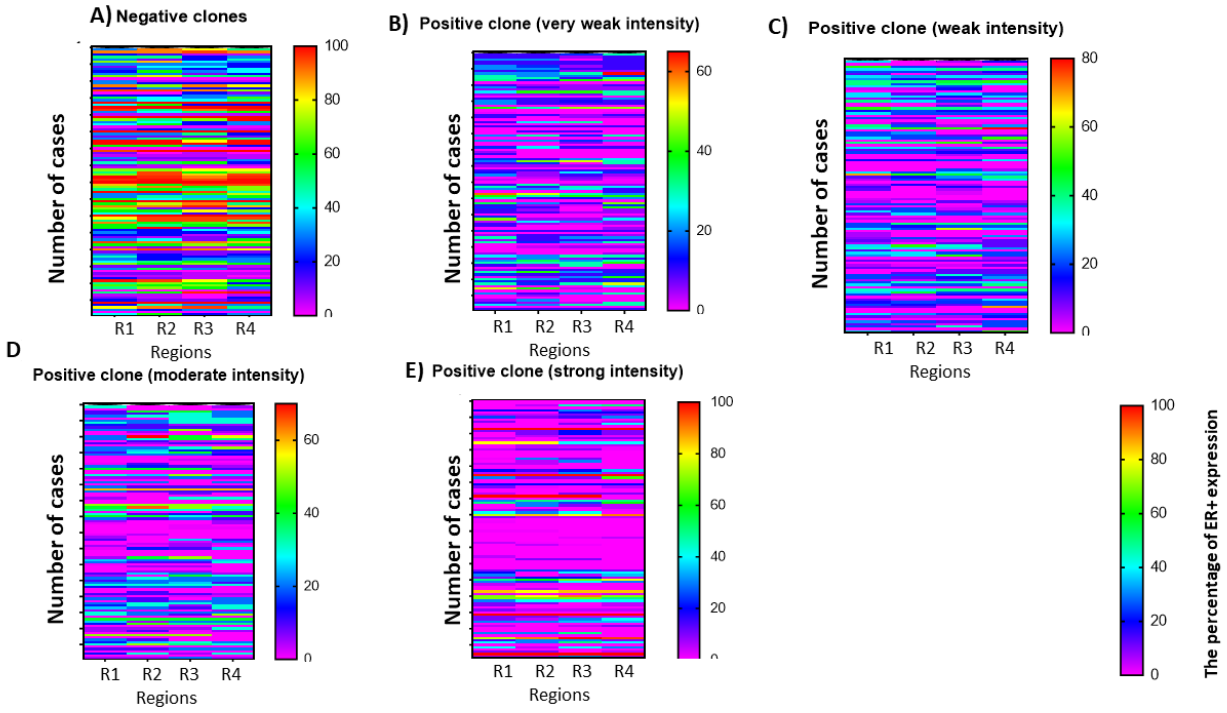


Figure 7.2: Heatmap showing the variation in oestrogen receptor (ER) expression at the regional level. A) negative ER expression, B) very weak ER+ expression, C) weak ER+ expression, D) moderate ER+ expression, and E) strong ER+ expression.

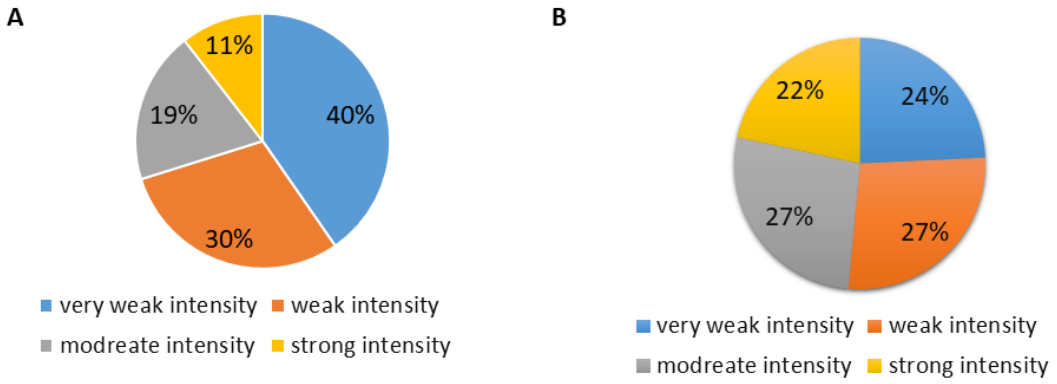


Figure 7.3: Percentage of tumours exhibiting each oestrogen receptor positive (ER+) staining intensity at the tumour level. A) The majority of tumours with 1-20% ER+ cells exhibited very weak or weak ER+ staining intensities. B) Tumours with 21-70% ER+ cells exhibited a range of ER+ staining intensities.

The ER staining distribution patterns were classified as uniform, diffuse, cluster, and scatter or focal, as shown in Figure 5.

At the tumour level, the focal pattern was observed in 14% (n=15) of cases; the majority of these cases (73%) contained 1%-20% ER+ cells. Tumours with the scatter and focal positive patterns showed high proportions of very weak and weak staining intensities (Figure 6). The diffuse pattern was detected in 94% of the tumours with 21-70% ER-positive cells.

At the regional level, a positive correlation between the uniform, diffuse, cluster or scatter distribution patterns of ER+ expression were detected between the four regions (all $P < 0.0001$; Figure 5). However, the strong and moderate intensities clustered together. The Spearman's rank test indicated a positive correlation between the strong and moderate intensities in all four regions (all P

< 0.0001) and a positive correlation between the weak and very weak intensities (all $P < 0.0001$; Table 2).

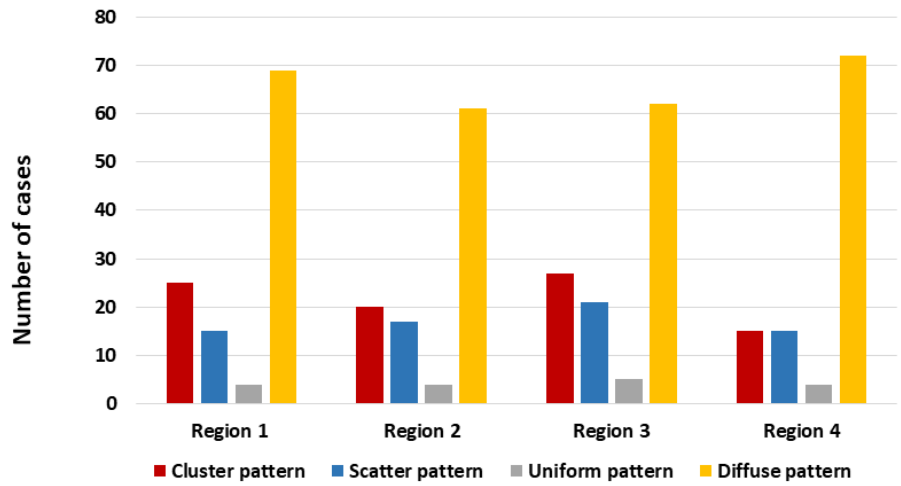


Figure 7.5: illustrates the number of cases with different patterns of ER distribution in each region.

Table 7.2: Spearman's rank analysis of the correlation between the percentages of cells with various ER+ staining intensities at the regional level in primary breast cancer.

Region	Correlation coefficient (*P*-value) for strong and moderate ER+ staining intensities	Correlation coefficient (*P*-value) for very weak and weak ER+ staining intensities
Region 1	0.60 (1.9×10^{-9})	0.49 (0.000003)
Region 2	0.55 (7.0×10^{-8})	0.55 (8.8×10^{-8})
Region 3	0.62 (4.2×10^{-10})	0.47 (0.000008)
Region 4	0.58 (7.9×10^{-9})	0.46 (0.000011)

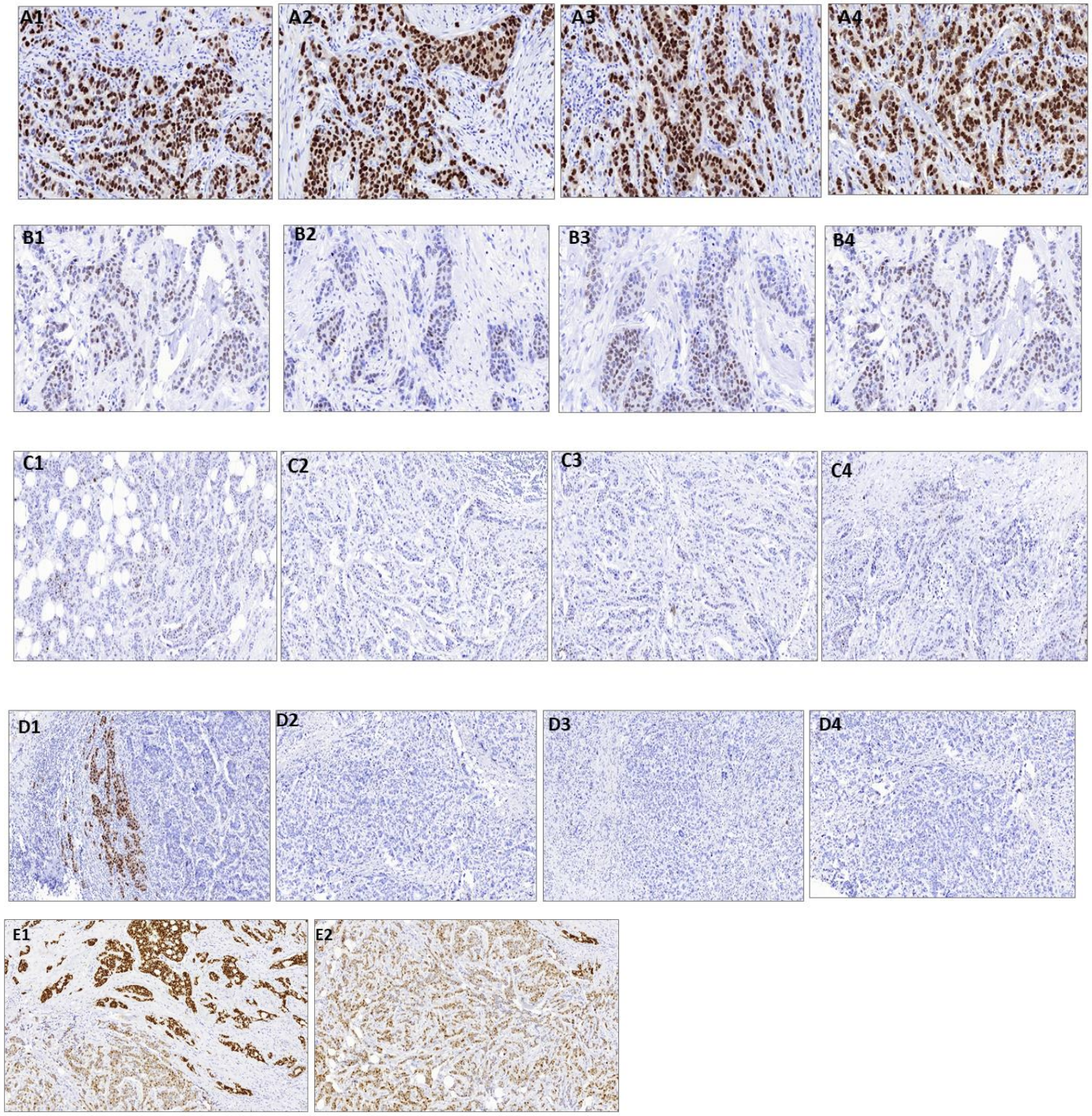


Figure 7.6: Representative images of the distribution patterns in different regions of the same tumour of ER expression in primary BC. (A1-4) Uniform pattern with strong intensity, (B1-4) diffuse, (C1-4) scatter, (D1-4), focal, and (E1&2) cluster patterns.

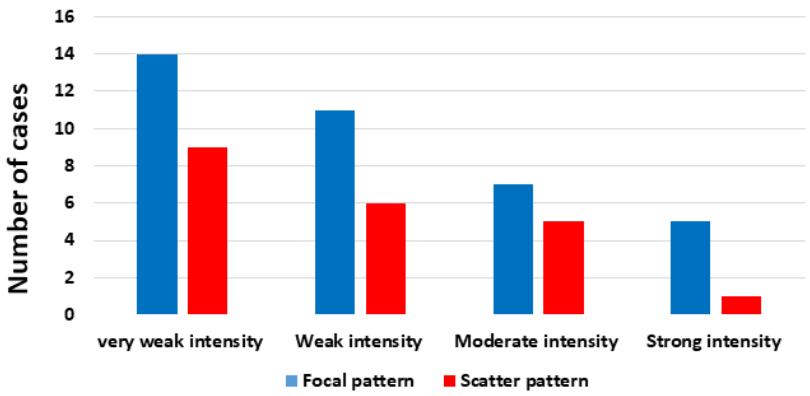


Figure 7.7: shows the high proportion of very weak and weak staining intensities in tumours with the focal and scatter distribution patterns of oestrogen receptor (ER) expression.

7.3.2 Co-expression of ER and PgR in primary IBC

Co-expression of ER and PgR was assessed in parallel using sequential sections of the same tumours (n=100).

At the tumour level, a high association between ER and PgR expression was observed (cut-off for PgR, ≥ 1% positive cells; $P < 0.0001$). Overall, 82% (81/100) of tumours were ER+ and PgR+, 18% (n = 18/100) were ER+ and PgR-, and 0% (0/100) ER- and PgR+.

In the ER+ and PgR+ subgroup, 43% (43/100) of the tumours exhibited a higher percentage of PgR+ cells than ER+ cells. Figure 8 shows an example of ER- tumour cells and the corresponding PgR+ cells, as well as an ER+ tumour with PgR+ cells in the same region.

At the regional level, variation in ER and PgR expression was observed. A higher percentage of cells exhibited strong (%) PgR+ expression than strong

ER+ expression, while a higher percentage of cells exhibited weak ER+ expression than weak PgR+ expression. Table 3 summarises the percentages of cells with negative, very weak, moderate and strong ER and PgR staining intensities.

There were no significant differences between the patterns of ER+ and PgR+ distribution at the regional or tumour level (all $P > 0.05$).

Table 7.4. Percentages of cells (median, mean and range) exhibiting strong, weak, very weak and negative oestrogen receptor (ER) and progesterone receptors (PgR) staining in the 111 primary breast cancer (BC) tumours.

	ER, percentage cells (%)			PgR, percentage cells (%)		
Staining intensity	**Median**	**Mean**	**Range**	**Median**	**Mean**	**Range**
Negative	40	45	1-30	57	56	1-30
Very weak +	10	12	0-45	0	4	0-30
Weak +	10	13	0-40	5	7	0-30
Moderate +	10	14	0-50	14	5	0-60
Strong +	4	14	0-100	5	18	0-100

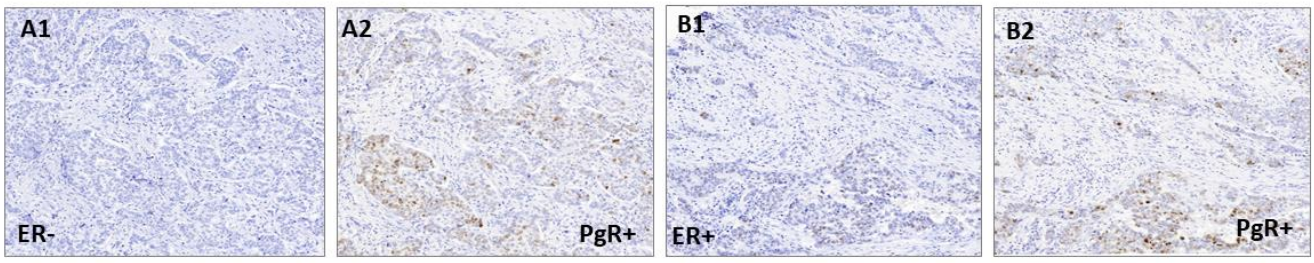


Figure 7.8 (A1-2): shows the same regions of an (A1) oestrogen receptor negative (ER-) and (A2) progesterone receptors positive (PgR+) tumour, (B1-2) shows the same regions of an (B1) ER+ and (B2) PgR+ tumour.

7.3.3 ER heterogeneity in primary tumours and the paired lymph node metastases

Overall, 40% (41/111) of the study cohort had lymph node (LN) metastases. Tissue blocks of the LN metastases were available for 37/40 of the patients with LN involvement and were subjected to staining for ER.

The percentage of ER+ cells was higher in LN metastases compared to the paired primary tumours ($P = 0.028$, Friedman test). However, two cases with ER+ primary tumours had ER- (0%) LN metastases. These primary tumours exhibited 25% and 20% ER+ cells, with focal and scatter patterns, respectively.

There was no significant difference in the ER expression pattern between the 40 primary tumours and the paired LN metastases ($P > 0.05$, Friedman test). Figure 9 presents representative examples of ER expression in the primary tumours and paired LN metastases.

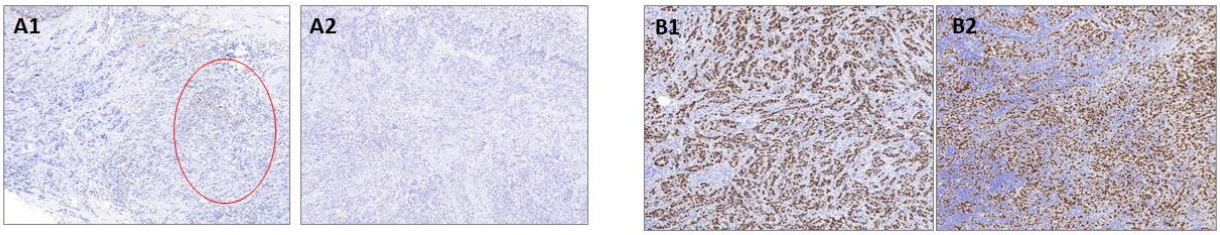


Figure 7.9: (A1 & A2) Primary BC with the focal pattern of oestrogen receptor positive (ER+) expression (A1) and the paired oestrogen receptor negative (ER-) LN metastasis (A2). (B1 & B2) Primary breast cancer with the diffuse pattern of ER+ expression (B1) and the paired LN metastases with the diffuse pattern of ER+ expression (B2).

7.3.4 Intratumor ER heterogeneity index

The intratumor ER heterogeneity index score was calculated at both the regional and tumour levels using Rao's quadratic entropy. The intratumor ER heterogeneity index score ranges from 0 to 1. The cut-off value for low and high intratumor ER heterogeneity was selected based on the continuous Rao's quadratic entropy score. Low intratumor ER heterogeneity was defined as a score of 0.20 or less and high intratumor ER heterogeneity as a score of greater than 0.20.

The maximum value for the cohort was 0.78. Overall, 78% (87/111) of tumours exhibited high intratumor ER heterogeneity and 21% (24/111) exhibited low intratumor ER heterogeneity. Figure 10 provides representative examples of tumours with high and low intratumor ER heterogeneity.

At the tumour level, the data showed that the majority (91%) of tumours with high ER heterogeneity had 21-70% ER+ cells, as shown in Figure 11.

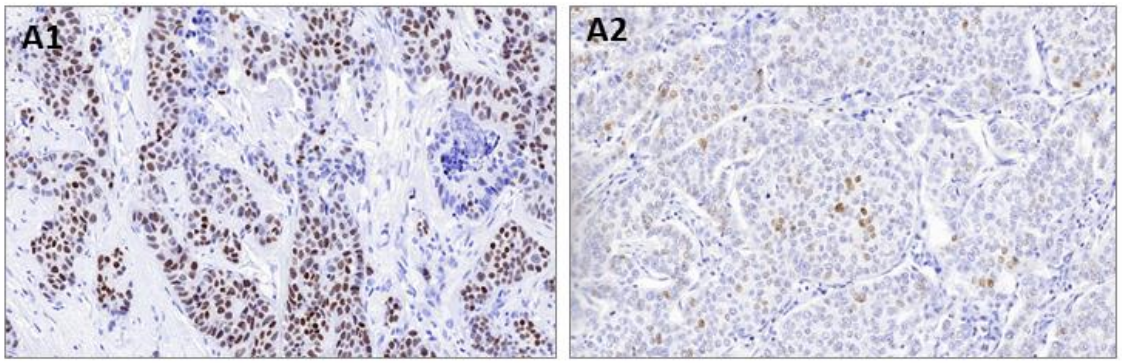


Figure 7.10: Representative examples of primary breast cancer (BC) specimens with (A) high and (B) low intratumor oestrogen receptor (ER) heterogeneity.

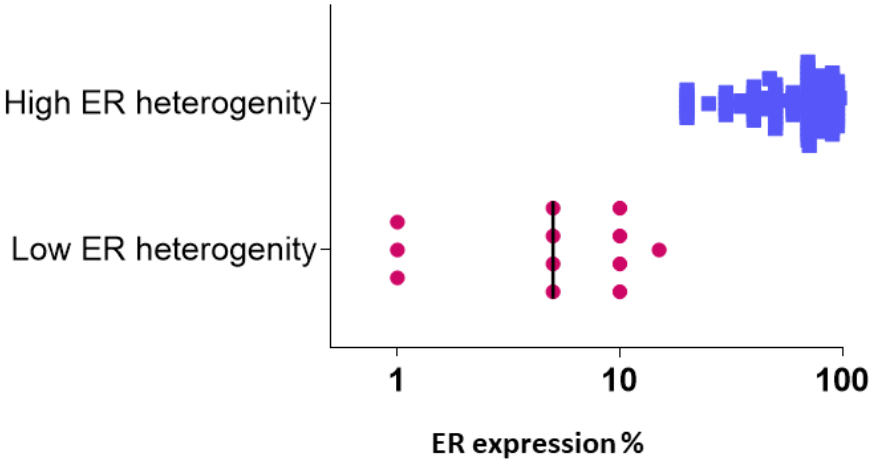


Figure 7.11: Dot plot of intratumor oestrogen receptor (ER) heterogeneity versus percentage of ER+ cells for the 111 primary breast cancer (BC) specimens.

There was a significant difference in the ER heterogeneity score between the four regions. The mean ranks were 3.80, 2.80, 2.15, and 2.37 for the four regions ($P < 0.0001$, Friedman test).

ER expression was more heterogeneous in the primary tumours (mean 0.53) than the LN metastases (mean=0.40; chi square test, $P = 0.025$).

Intratumoral PgR heterogeneity ranged from 0.019 to 0.78. PgR expression (mean= 0.37) was more heterogeneous than ER expression (mean=0.50; $P = 0.011$, chi square test).

7.3.5 Clinical significance of intratumour ER heterogeneity

We examined the clinical impact of intratumor ER heterogeneity by stratifying the 111 patients with 1-70% ER+ cells into two subgroups based on low and high tumour ER heterogeneity. Furthermore, the 111 tumours with high or low intratumoural ER heterogeneity were compared with homogenous ER tumours (i.e. the comparison cohort of $n = 391$ patients with 100% ER+ tumours, and H-score=300).

When the index score was examined, there was no significant association between high and low ER heterogeneity and any clinicopathological parameters (Table 5).

ER heterogenous tumours were associated with younger age ($P < 0.0001$), high tumour grade ($P < 0.0001$), advanced nodal stage ($P = 0.035$), poor NPI ($P < 0.0001$) and invasive carcinoma non-special histological type ($P < 0.0001$; Table 6) compared to the homogenous tumour (patients with 100% ER+ tumours, and H-score=300).

The Chi square test revealed high ER heterogeneity was significantly associated with the diffuse pattern, while low ER heterogeneity was associated with the focal and scatter patterns ($P = 0.046$).

Descriptive analysis showed that tumours with the focal pattern were mainly high grade (79%), 35% were LVI-positive, and 50% were large tumours. Of the tumours with the diffuse pattern, 48% were high grade, 20% were LVI-positive and 30% were large tumours.

Table 7.6: Clinicopathological features of the tumours stratified by the patterns of oestrogen receptor (ER) distribution in 111 cases of primary breast cancer (BC).

Clinicopathological feature	**Focal pattern N (%)**	**Diffuse pattern N (%)**	**Scatter pattern N (%)**
Tumour size (cm)			
< 2	7 (50)	50 (50)	5 (36)
≥ 2	7 (50)	25 (25)	4 (64)
Tumour Grade			
Grade I	1 (14)	10 (13)	1 (11)
Grade II	2 (28)	29 (39)	6 (67)
Grade III	4 (58)	36 (48)	2 (22)
Axillary nodal stage			
Stage I	5 (38)	50 (67)	6 (67)
Stage II	3 (24)	16 (21)	1 (11)
Stage III	5 (38)	9 (12)	2 (22)
Lymphovascular invasion			
Negative	9 (64)	60 (80)	5 (62)
Positive	5 (36)	15 (20)	3 (38)

Table 7.7: The association between intratumor oestrogen receptor (ER) heterogeneity and clinicopathological features in 111 cases of primary breast cancer (BC) (ER expression=1-70%)*****

Clinicopathological feature	**Low ER heterogeneity index score *N* (%)**	**High ER heterogeneity index score *N* (%)**	**x^2 *P*-value**
Patient Age (Years)			
< 50	6 (19)	26 (81)	0.088
≥ 50	17 (21)	63 (79)	0.767
Tumour size (cm)			
< 2	13 (19)	57 (81)	0.441
≥ 2	10 (24)	32 (76)	0.506
Tumour Grade			
Grade I	1 (6)	15 (93)	3.798
Grade II	12 (29)	30 (71)	0.150
Grade III	10 (18)	44 (82)	
Axillary nodal stage			
Stage I	12 (18)	56 (82)	1.896
Stage II	6 (23)	20 (77)	0.387
Stage III	5 (33)	10 (67)	
Lymphovascular invasion			
Negative	13 (18)	61 (82)	3.673
Positive	9 (36)	16 (64)	0.055
Nottingham Prognostic Index			
Poor Prognostic Group	4 (27)	11 (85)	1.199
Moderate Prognostic Group	14 (24)	45 (76)	0.549
Good Prognostic Group	5 (15)	28 (73)	
Histological tumour types			
Ductal (including mixed)	22 (24)	69 (76)	4.127
Special type	1 (13)	7 (87)	0.127
Lobular	0 (0)	12 (100)	

*Chi square test.

Table 7. 8. Association between intratumor oestrogen receptor (ER) heterogeneity (ER expression=1-70%) and ER homogeneity expression (ER=100%, and H-score=300) with the clinicopathological features of 502 cases of primary breast cancer (BC).

Clinicopathological feature	**ER homogenous N = 391 *N* (%)**	**ER heterogeneous N = 111 *N* (%)**	**X^2 *P*-value**
Patient age (Years)			
< 50	28 (49)	29 (51)	35.794
≥ 50	363 (83)	73 (17)	<0.0001
Tumour size (cm)			
< 2	277 (81)	65 (19)	1.929
≥ 2	114 (76)	37 (24)	0.165
Tumour Grade			
Grade I	116 (88)	16 (12)	27.773
Grade II	185 (84)	36 (16)	<0.0001
Grade III	90 (64)	50 (36)	
Axillary nodal stage			
Stage I	284 (82)	64 (18)	6.681
Stage II	87 (79)	23 (21)	0.035
Stage III	20 (63)	12 (38)	
Lymphovascular invasion			
Negative	308 (82)	69 (18)	0.067
Positive	83 (81)	20 (19)	0.796
Nottingham Prognostic Index			
Poor Prognostic Group	27 (46)	32 (54)	75.130
Moderate Prognostic Group	153 (75)	52 (25)	<0.0001
Good Prognostic Group	210 (94)	13 (6)	
Histological tumour type			
Ductal (including mixed)	312 (98)	7 (2)	
Special type	29 (100)	0 (0)	321.649
Lobular	39 (100)	0 (0)	<0.0001
Non-special type	1 (8)	12 (92)	
Tubular carcinoma	23 (25)	69 (75)	

*Chi square test

Univariate survival analysis showed that high intratumoral ER heterogeneity was also associated with a longer disease-free interval (DFI; *P* = 0.036; Figure 12) compared to low intratumoral ER heterogeneity, but not with BCSS (*P*>0.05). Compared to low intratumoral ER heterogeneity, high intratumoral ER heterogeneity was not significantly associated with BCSS or DFI in patients who received endocrine treatment or among patients who did not receive endocrine treatment (*P* > 0.05).

ER heterogeneity tumours (ER=1-70%) was associated with poorer BCSS in both patients who received endocrine treatment (*P* = 0.040) and those who did not receive endocrine treatment (*P*=0.001; Figure 13) compared to patients with ER homogenous tumours (tumour with ER=100% and H-score=300).

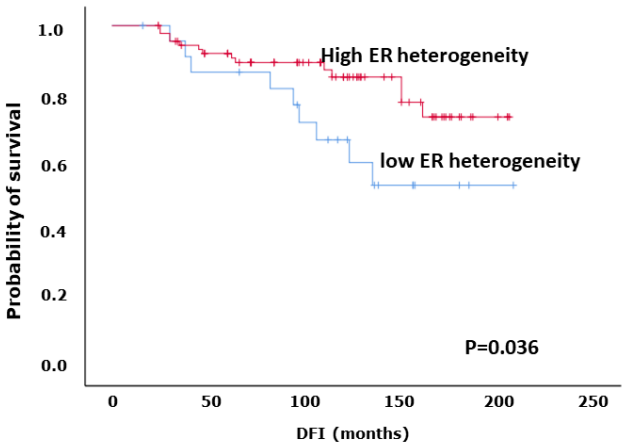


Figure 7.12: Kaplan-Meier survival plots of disease-free interval (DFI) for 111 breast cancer patients (ER expression=1-70%) stratified by high and low intratumoral oestrogen receptor (ER) heterogeneity.

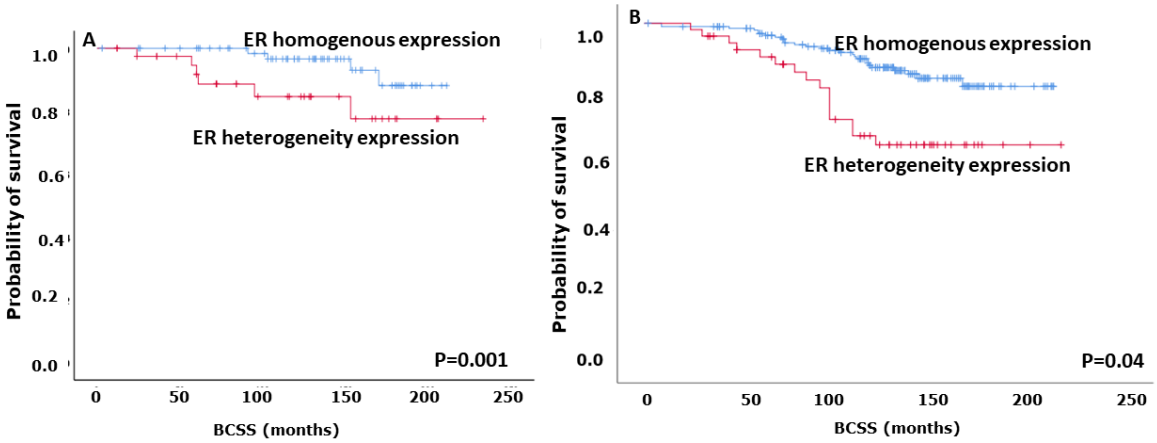


Figure 7.13: Kaplan-Meier survival plots of breast cancer specific survival (BCSS) for patients with oestrogen receptor (ER) heterogeneous tumours (high or low heterogeneity index score) and patients with ER homogenous tumours (ER expression =100% and H-score=300) who did not receive endocrine treatment (A) and who received endocrine treatment (B).

7.4 Discussion

Intratumor ER heterogeneity expression is widely observed in clinical BC samples. This investigation was based on the concept that ER heterogeneity may play a biological role on tumour behaviour and resistance to targeted therapy [469, 470]. ER heterogeneity was evaluated based on pathological assessment of IHC staining. Three measures were used to quantify ER heterogeneity at the regional and tumour levels: (i) the percentages of ER- cells and ER+ cells, including very weakly, weakly, moderately and strongly. (ii) the patterns of ER+ distribution, and (iii) Rao's quadratic entropy index [447, 464, 465, 471]. The three approaches were showed promising performance for assessment of ER heterogeneity.

These approaches confirmed spatial intratumoural heterogeneity in ER expression within different regions of the same tumour. ER+ clones exhibited distinct spatial locations in BC, in support of the suggestion that ER+ tumours are not a single biological system. Potts *et al* (2012) evaluated HER2 heterogeneity in BC at the cellular and tumour levels, and concluded that HER2 expression varied between neighbouring cells, as well as different regions of the tumours [453]. In the current study, Rao's quadratic entropy was applied to measure ER heterogeneity at the regional and tumour levels. Significant variations in ER expression were observed within the tumours and between regions, in agreement with Potts *et al* (2012). BC has the propensity to develop multiple subclones, which leads to intratumor heterogeneity. Phenotypic heterogeneity is thought to be biologically relevant in BC, as heterogeneous tumours can behave differently than homogenous tumours, and these differences can have a prognostic significance [472].

The clonal evolution hypothesis of cancer may explain ER heterogeneity and the existence of ER+ and ER- clones with different abilities to progress and metastasise. ER+ tumours can originate from a single genetically instable cell, followed by sequential selection of branching subclones [473, 474]. The clonal evolution hypothesis suggests that accumulation of genetic damage in somatic cells lead to tumour progression. Initially, the cellular composition of a tumour is homogeneous. Cells within the same subclone are assumed to have similar genetic profiles and equal oncogenic potential. However, acquisition of additional functional mutations leads to generation and divergence of tumour subclones with a competitive advantage, and leads to tumour heterogeneity. [475-477].

In further support of this result, it was observed that some ER-negative cells were PgR+. When assessed as a continuous value, the percentages of PgR+ cells were higher than the percentages of ER+ cells at both the regional and tumour levels. This observation suggests that ER expression reduces during the progression of breast cancer. PgR gene expression is dependent on the ER, thus PgR expression is thought to indicate the presence of an intact ER response pathway [450]. Although a lower proportion of BC tumour cells were PgR+ than ER+ [478], this result was based on an arbitrary cut-off whereas this study evaluated the percentages of ER+ and PgR+ cells as continuous variables. Although ER-/PgR+ BC is rare, Rhodes and Jasani (2009) reported that ER-/PgR+ represents a distinct biological phenotype, and more studies of this distinct subgroup are required [479]. The percentage of ER+ cells was higher in LN metastases BC than the paired primary tumours, which indicates the LN microenvironment does not downregulate the ER [480]. However, two cases with ER+ primary tumours and ER- LN metastases were detected in this cohort. These primary tumours had focal or scattered patterns of ER expression and low ER

heterogeneity indexes. In addition, it could be explain that the high proportion subclone of the tumour can drive the tumour to the LN. Taken together, the pattern of ER distribution should be considered when evaluating ER+ tumours with LN involvement.

The pattern of ER expression may have relevance in clinical practice. The ER expression pattern correlated with ER status in the metastases, and ER status strongly determines the response to endocrine therapy in breast cancer [481]. This study suggests patients with the focal or scatter pattern of ER expression should be treated with a combination of endocrine and chemotherapy.

For example, Harrell *et al* (2006) reported that the CD44+ subpopulation of tumour cells preferentially enter the tumour lymphatic system and metastasise to the LNs. Thus, targeting CD44+ primary tumour cells could potentially prevent LN metastasis [480].

The clinical significance of ER heterogeneity in BC was also assessed. ER heterogeneity was associated with aggressive tumour features and poorer outcomes compared to the cohort of patients with ER homogenous tumours, in agreement with a previous study [482]. The patients with ER+ heterogenous tumours were divided into two subgroups based on the intratumoral ER heterogeneity score. The low heterogeneous subgroup was defined by a heterogeneity index score ≤ 0.20, while tumours with a score > 0.20 were classified as highly heterogonous and exhibited variable diffuse intensities of ER staining. ER heterogeneous tumour (ER expression= 1-70%) was associated with more aggressive tumour features and resistance to endocrine therapy compared to ER homogenous tumour (ER expression= 100%, and H-score=300), where in the result the patients with ER heterogenous tumour who received endocrine

therapy have poorer outcome compared to homogenous tumour. The result indicates that ER heterogeneous tumours represent two distinct subgroups: one group with a spectrum of heterogeneity of ER expression containing multiple subclones with varied intensities of ER expression, and another group composed of two subclones of ER- cells and highly expressing ER+ cells.

In conclusion, this study confirms that ER heterogeneity occurs in BC and has a biological and clinical impact. Patients with intratumor ER heterogeneity and ER- subclones have significantly poorer outcomes compared to patients with homogenous ER expression. High intratumor ER heterogeneity scores were associated with longer DFI. Intratumor ER heterogeneity and the presence of ER- subclones was associated with resistance to endocrine therapy.

At present, endocrine therapy is recommended for patients with ER+ tumours and high NPI scores (> 3.4), whereas no adjuvant therapy is recommended for patients with excellent NPI scores (≤ 3.4). Premenopausal patients with moderate and poor NPI scores are candidates for chemotherapy, while postmenopausal patients with moderate or poor NPI scores are only given hormonal therapy [40].

This study suggests that patients with LN involvement and high intratumor ER heterogeneity scores can receive the same treatment as patients without LN involvement (endocrine treatment only), whereas patients with low heterogeneity ER scores should receive a combination of endocrine and chemotherapy.

Chapter 8

Oestrogen-regulated protein SLC39A6 is a prognostic marker associated with good outcome in luminal breast cancer

8.1 Introduction

Approximately 30–50% of patients with ER+ BC do not respond to endocrine therapy, which indicates ER+ BC is a heterogeneous subtype [483, 484]. Analysis of gene expression in BC and assessments of the prognostic and predictive value of novel biomarkers [485, 486] have highlighted the need to refine the classification of BC in order to more accurately reflect heterogeneity and tailor personalised therapeutic approaches [487, 488].

Several molecular biomarkers have been evaluated in luminal tumours, with a research emphasis on ER-related genes such the genes encoding the progesterone receptor (PgR), Forkhead box protein A1 (FOXA1), GATA binding protein 3 (GATA3) and Trefoil factor 1 (TFF1) [450, 489, 490]. Moreover, Andres *et al*. (2013) identified a signature of genes related to recurrence in BC, which included solute carrier family 39 member 6 (*SLC39A6*)[491]. SLC39A6 is an oestrogen-regulated gene that is upregulated in ER+ BC [492]. Moreover, expression of SLC39A6 has been associated with ER+ status [493].

SLC39A6, also known as LIV-1 and ZIP6, is encoded on chromosome 18q12.2 [494] and expressed at high levels in hormonally controlled tissues [495]. SLC39A6 is located on the endoplasmic reticulum (ER), while the N-terminal cleaved form is present on cell membranes [496].

SLC39A6 is a member of the ZIP family of transporters, which control zinc homeostasis by regulating zinc influx from extracellular to intracellular spaces [497]. The transport function of SLC39A6 plays an important role in cellular metabolism [496, 498, 499]. Zinc is required for a variety of cellular processes, including immune activity, protein synthesis, nucleic acid metabolism, cell proliferation, tissue repair and cell division [500], and low zinc levels can lead to metabolic disorder and inhibit cell growth [501].

Zinc is also involved in several signalling pathways in BC. For instance, zinc enters cells through SLC39A6 and activates Akt, which inhibits glycogen synthase kinase 3 beta (GSK-3β) [502]. In turn, deactivation of GSK-3β can affect Snail and downregulate E-cadherin (CDH1), epidermal growth factor receptor (EGFR) and MAPK [503].

Overexpression of SLC39A6 has been related to progression in several tumour types, including breast [504, 505], prostate [506], pancreatic [507], cervical [508] and liver cancer [509]. An *in vitro* study suggested that SLC39A6 is involved in metastasis in BC, as overexpression of SLC39A6 promoted the epithelial-mesenchymal transition (EMT) [503].

However, Kasper *et al.* (2005) reported that high SLC39A6 protein expression was associated with a better prognosis in BC (n = 111) [510]. Moreover, Taylaer *et al.* (2008) found that high SLC39A6 protein expression was associated with ER+ status and low grade tumours in 75 BC patients [511].

Therefore, the prognostic and predictive value of SLC39A6 expression at the proteomic and genomic levels in BC, especially in ER+ tumours, warrant further investigation.

8.1 1 Hypothesis

The oestrogen-regulated protein SLC39A6 has prognostic and predictive value in BC and may help to refine the classification of ER+ BC.

8.1.2 Aims

The aims of this study are to investigate the prognostic and predictive value of SLC39A6 in invasive BC, particularly in ER+ tumours.

8.1.3 Objectives

The objectives of this chapter are to:

(i) Assess SLC39A6 expression at the protein and mRNA levels in BC.

(ii) Investigate the associations between SLC39A6 protein and mRNA expression and the clinicopathological features and outcomes of patients with BC, especially ER+ tumours.

(iii) Explore the correlations between expression of SLC39A6 and other ER-related markers at both the protein and mRNA levels.

8.2 Materials and Methods

8.2.1 Study cohorts

SLC39A6 protein expression was evaluated in specimens from 670 patients with primary invasive BC (IBC) derived from a retrospective series of patients who presented at Nottingham City Hospital between 1989 and 1998. The Nottingham cohort is described in detail in Chapter 2; 62% of the patients were < 50-years-old, 12% and 59% had grade 1 and 3 tumours, respectively, and 75% had ER+ BC.

SLC39A6 mRNA expression and gene copy number (CN) alterations were assessed using the METABRIC cohort of 1,980 patients with BC; 20% of these patients were < 50-years-old, 9% and 50% had grade 1 and 3 tumours, respectively, and 76% had ER+ BC. The clinicopathological parameters of both study cohorts are summarised in Table 8.1.

Table 8.1: Summary of the demographic features of the cohorts assessed in this chapter.

Clinicopathological parameter	**METABRIC series *n* (%)**	**Nottingham series *n* (%)**
Age		
≥ 50 years	1526 (80)	251 (38)
< 50 years	377 (20)	410 (62)
Tumour size		
> 2 cm	1308 (68)	313 (48)
≤ 2 cm	605 (32)	344 (52)
Grade		
Grade 1	166 (9)	82 (12)
Grade 2	754 (41)	188 (29)
Grade 3	935 (50)	388 (59)
Tumour type		
Ductal (including mixed)	1545 (84)	536 (86)
Lobular	148 (8)	43 (6.6)
Medullary-like	32 (2)	19 (3)
Miscellaneous	12 (0.6)	3 (0.4)
Special type	113 (6)	23 (4)
Vascular Invasion		
Definite	Not available	421 (64)
Negative/Probable		236 (36)
Lymph node stage		
Stage 1	1016 (52)	405 (61)
Stage 2	609 (31)	192 (29)
Stage 3	311 (16)	62 (10)
Follow-up status		
Alive	1052 (68)	366 (55)
Died of BC	469 (24)	213 (33)
Died of other causes	N/A	81 (12)
Oestrogen receptor status		
Negative	470 (24)	195 (30)
Positive	1473 (76)	464 (70)
Progesterone receptor status		
Negative	922 (47)	294 (45)
Positive	1021 (53)	359 (55)
HER2 status		
Negative	1703 (88)	556 (85)
Positive	240 (12)	95 (15)

8.2.2 Assessment of SLC39A6 protein expression in the Nottingham cohort

Prior to performing immunohistochemistry (IHC), the specificity of the primary rabbit polyclonal anti-SLC39A6 antibody (#AA170320; LSBio, Cambridge, UK) was validated by western blotting of MCF7, MDA-MB-231 and HeLa cell lysates, as described in Chapter 2.

IHC staining was performed on 4 μm-thick full face sections and TMA sections using the Novolink Polymer Detection System (#RE7280-K; Leica, Newcastle, UK), as described in detail in Chapter 2. Antigen retrieval was performed in citrate buffer (pH 6.0) in a microwave (Whirlpool JT359 Jet Chef 1000 W) for 20 min. The SLC39A6 antibody was applied at the optimal dilution of 1:200 for 60 min at room temperature. Normal kidney tissue sections were used as a positive control, and the negative control was obtained by omitting the primary antibody.

SLC39A6 staining expression was semi-quantitatively evaluated using the modified histochemical scoring method (H-score), as described in detail in Chapter 2.

8.2.3 Existing immunohistochemical analysis for the Nottingham cohort

Immunohistochemical analysis of ER-related markers, including PgR, FOXA1, GATA3 and TFF1, had been previously performed for the Nottingham cohort; details of these markers are provided in Chapter 2.

8.2.4 Statistical analysis

Statistical analysis was performed as described in Chapter 2 using IBM SPSS 24.0 software (Chicago, IL, USA).

SLC39A6 expression was assessed at the proteomic and transcriptomic levels categorised using X-tile software (X-tile Bioinformatics Software version 3.6.1; Yale University; see Chapter 2).

8.3 Results

8.3.1 SLC39A6 protein expression in BC

The SLC39A6 primary antibody was used at a dilution of 1:1500 for Western blotting and detected a single, specific band at the predicted molecular weight of 85 kDa for SLC39A6, which confirmed the specificity of the antibody (Figure 8.1).

IHC analysis of full-face sections revealed homogenous immunohistochemical staining for SLC39A6 and validated the use of TMA cores to assess SLC39A6 expression in BC. SLC39A6 expression was observed as cytoplasmic and nuclear immunoreactivity in invasive epithelial tumour cells (Figure 8.2A, B & C).

High cytoplasmic SLC39A6 expression (cut-off, H-score > 160) and high nuclear SLC39A6 expression (cut-off, H-score > 0) were observed in 246/670 (37%) and 285/670 (43%) of the BC cases, respectively. SLC39A6 nuclear and cytoplasmic co-expression was observed in 140/670 (21%) of the entire Nottingham cohort. Moreover, a significant positive correlation was observed between SLC39A6 cytoplasmic and nuclear expression ($R = 0.422$, $P < 0.0001$).

Table 8.2 shows the mean, median and ranges of SLC39A6 cytoplasmic and nuclear expression in ER+ and ER- tumours. SLC39A6 protein was expressed at significantly higher levels in ER+ tumours than ER-negative tumours ($P < 0.0001$). Moreover, high SLC39A6 expression was significantly associated with the luminal A subtype, which is defined by ER+, PR+, HER2- and low Ki67 expression.

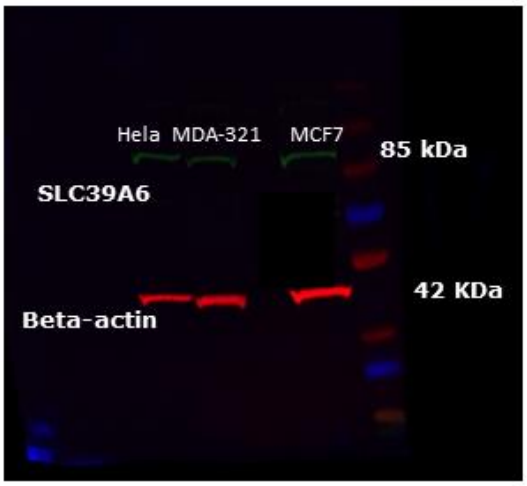


Figure 8.1: Western blot validation of the rabbit anti-SLC39A6 antibody. A single specific band (green) was detected at the expected molecular weight (85 kDa) in MCF-7, MDA-MB-231 and HeLa cell lysates. The red bands represent the β-actin positive control at 42 k Da.

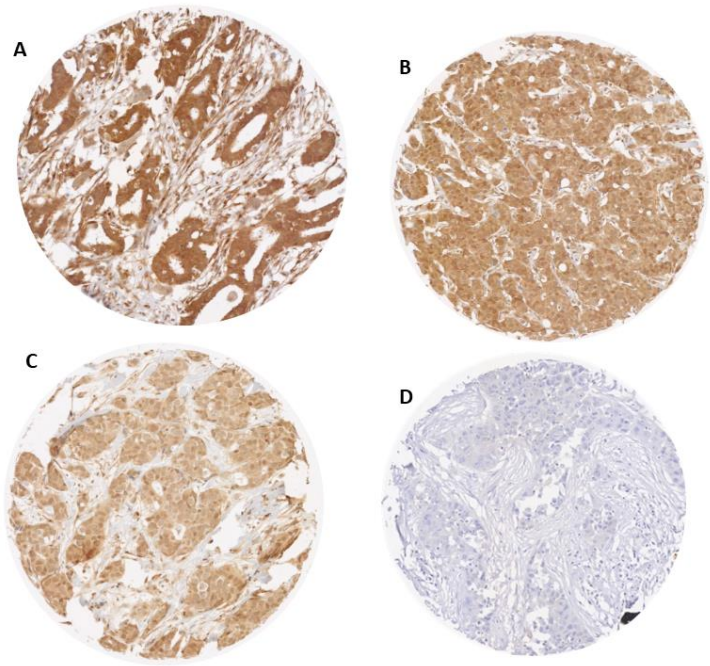


Figure 8.2: (A) Representative images of positive (strong intensity) immunohistochemical staining for SLC39A6 in the cytoplasm and nuclei of BC cells. (B) Positive (moderate intensity) immunohistochemical staining for SLC39A6 in the cytoplasm and nuclei of BC cells. (C) Positive (weak intensity) immunohistochemical staining for SLC39A6 in the cytoplasm and nuclei of BC cells. (D) Negative immunohistochemical staining for SLC39A6.

8.3.2 SLC39A6 mRNA expression in BC

High *SLC39A6* mRNA expression (8.5 log fold-change or greater) was observed in 981/1980 (50%) of the entire METABRIC cohort. Furthermore, high *SLC39A6* mRNA expression was observed in 1186/1473 of the ER+ tumours (79%; $P < 0.0001$) and only 21/449 (4%) of the ER-negative tumours.

SLC39A6 protein expression was not associated with *SLC39A6* mRNA expression ($P > 0.05$).

SLC39A6 CN gains were observed in 49/1980 (2.5%) and 35/1471 (2.3%) of the entire cohort and ER+ tumours, respectively, whereas 80/1980 (4.0%) and 66/1440 (4.4%) of the entire cohort and ER+ tumours exhibited CN loss, respectively.

Table 8.2: Mean, median and ranges of SLC39A6 mRNA and protein expression in the ER+ and ER- subgroups of BC.

SLC39A6 expression	**ER+ subtype**			**ER- subtype**		
	Mean	**Median**	**Range**	**Mean**	**Median**	**Range**
Cytoplasmic (H-score)	135	150	0-240	132	150	0-210
Nuclear (H-score)	44	0	0-200	25	0	0-200
mRNA (log fold change)	9.7	9.7	5.5-13.4	7.3	7.3	5.5-9.9

8.3.3 Associations between SLC39A6 expression and the clinicopathological features of BC

Both high nuclear and high cytoplasmic SLC39A6 protein expression were associated with features characteristic of a good prognosis in BC, including low tumour grade ($P < 0.0001$ and $P = 0.044$, respectively), low mitotic counts ($P < 0.0001$, only nuclear expression), low nuclear pleomorphism ($P < 0.0001$ and $P = 0.027$), good NPI ($P = 0.001$ and $P = 0.006$), and early nodal stage ($P = 0.004$, only nuclear expression). High SLC39A6 nuclear and cytoplasmic protein expression were not significantly associated with patient age, tumour size, LVI or histological subtype ($P \geq 0.05$; Table 8.3).

When the analysis was limited to luminal ER+ tumours, high nuclear and high cytoplasmic SLC39A6 protein expression were significantly associated with low tumour grade ($P < 0.0001$ and $P = 0.009$), low mitotic count ($P = 0.001$, only nuclear expression), a low nuclear pleomorphism score ($P < 0.0001$ and $P = 0.005$), early tumour stage ($P = 0.020$, only nuclear expression) and good NPI ($P < 0.0001$ and $P = 0.014$; Table 8.4).

Low *SLC39A6* mRNA expression was associated with high tumour grade ($P < 0.0001$ and $P = 0.006$) and poor NPI (all $P < 0.0001$) in both the entire BC cohort and the ER+ subgroup (Table 8.5).

Table 8.3: Associations between SLC39A6 cytoplasmic and nuclear protein expression and the clinicopathological features of breast cancer.

Clinicopathological parameters	**SLC39A6 cytoplasmic expression**			**SLC39A6 nuclear expression**		
	Low *N* (%)	**High** *N* (%)	**x^2** *P*-value	**Low** *N* (%)	**High** *N* (%)	**x^2** *P*-value
Patient age (years)						
< 50	270 (66)	140 (34)	2.49	240 (58)	170 (42)	0.635
≥ 50	150 (60)	101 (40)	0.114	139 (55)	112 (45)	0.426
Tumour size (cm)						
≤ 2	232 (67)	112 (33)	4.55	216 (63)	128 (37)	9.122
> 2	186 (59)	127 (41)	0.033	160 (37)	153 (49)	0.002
Tumour grade						
Grade 1	46 (56)	36 (44)	6.26	34 (42)	48 (58)	43.44
Grade 2	111 (59)	77 (41)	0.044	80 (43)	108 (57)	<0.0001
Grade 3	262 (67)	126 (32)		264 (68)	124 (32)	
Axillary nodal stage						
Stage 1	251 (62)	154 (38)	2.10	216 (53)	189 (47)	10.87
Stage 2	125 (65)	67 (35)	0.350	117 (61)	75 (39)	0.004
Stage 3	44 (71)	18 (29)		46 (74)	16 (26)	
Nottingham Prognostic Index						
Good NPI	91 (56)	72 (44)	10.33	65(40)	98 (60)	35.21
Moderate NPI	237 (63)	136 (37)	0.006	221(59)	152 (41)	<0.0001
Poor NPI	90 (74)	31 (26)		90(74)	31 (26)	
Vascular invasion status						
Negative						
Positive	268 (64)	153 (36)	0.001	236(56)	185 (44)	0.659
	150 (64)	86 (36)	0.980	140(59)	96 (41)	0.233

Table 8.4: Associations between nuclear and cytoplasmic SLC39A6 protein expression and the clinicopathological features of oestrogen receptor positive tumours.

Clinicopathological parameters	**SLC39A6 cytoplasmic expression**			**SLC39A6 nuclear expression**		
	Low No (%)	**High** No (%)	**x^2** *P*-value	**Low** No (%)	**High** No (%)	**x^2** *P*-value
Patient age (Years)						
< 50	206 (67)	102 (33)	**3.99**	175 (57)	133 (43)	**5.31**
≥ 50	89 (57)	66 (43)	**0.046**	71 (46)	85 (54)	**0.014**
Tumour size						
≤ 2	147 (69)	67 (31)	3.96	129 (60)	86 (40)	**8.08**
> 2	147 (60)	99 (40)	0.47	115 (47)	131 (53)	**0.003**
Tumour grade						
Grade 1	43 (54)	36 (46)	**9.52**	34 (43)	45 (57)	**26.88**
Grade 2	106(60)	72 (40)	**0.009**	75 (42)	103 (58)	**<0.0001**
Grade 3	145(71)	58 (29)		136 (68)	68 (33)	
Axillary nodal stage						
Stage 1	178 (63)	106 (37)	4.49	140 (49)	145 (51)	**7.81**
Stage 2	85 (62)	52 (38)	0.106	77 (56)	60 (44)	**0.020**
Stage 3	31 (80)	8 (20)		28 (72)	11 (28)	
Nottingham Prognostic Index						**19.26**
Good	86 (56)	68 (44)	**8.50**	62 (40)	92 (60)	**<0.0001**
Moderate	159 (66)	82 (34)	**0.014**	136 (56)	106 (44)	
Poor	49 (75)	16 (25)		46 (70)	19 (29)	
Vascular invasion status						
Negative	186 (65)	102 (35)	0.15	151 (52)	137 (47)	0.07
Positive	108 (63)	64 (37)	0.698	93 (54)	80 (46)	0.782

Table 8.5: Associations between *SLC39A6 mRNA* expression and the clinicopathological features of BC in the entire METABRIC cohort and oestrogen receptor positive tumours.

Clinicopathological parameters	***SLC39A6* expression in BC**			***SLC39A6* expression in ER+ BC**		
	Low No (%)	**High** No (%)	**x^2** *P*-value	**Low** No (%)	**High** No (%)	**x^2** *P*-value
Patient age (Years)						
< 50	202 (54)	175 (46)	50.27	48 (22)	167 (78)	1.19
≥ 50	516 (34)	1010 (66)	**<0.0001**	236 (19)	998 (81)	0.275
Tumour size (cm)						
≤ 2	239 (39)	370 (61)	0.77	98 (21)	366 (79)	1.035
> 2	486 (37)	822 (63)	0.380	187 (19)	805 (81)	0.172
Tumour Grade						
Grade 1	32 (19)	134 (81)	**192.28**	28 (17)	134 (83)	**10.39**
Grade 2	173 (23)	581 (77)	**<0.0001**	115(16.6)	576 (83)	**0.006**
Grade 3	500 (54)	435 (46)		131(24)	421 (76)	
Axillary nodal stage*						
Stage 1	375 (37)	641(63)	1.10	172(21)	632 (79)	4.22
Stage 2	231(38)	378 (62)	0.576	80(18)	371(82)	0.121
Stage 3	125 (40)	186 (60)		35(16)	181 (84)	
Nottingham Prognostic Index						
Good	163 (24)	505 (76)	83.90	111(18)	500 (82)	1.49
Moderate	471 (44)	609 (56)	**<0.0001**	156(21)	597 (79)	0.474
Poor	102 (52)	93 (48)		20(18)	89(82)	

8.3.4 Associations between expression of SLC39A6 and other ER-related biomarkers

The correlations between SLC39A6 protein and mRNA expression and the expression of other ER-related markers were examined using data available for the METABRIC cohort (mRNA expression) and the biomarker repository of the Nottingham BC cohort (for protein expression). The biomarkers selected for this analysis are associated with the ER in BC.

Significant positive correlations were identified between SLC39A6 expression and the expression of other ER-related markers, including PgR, FOXA1, GATA3, and TFF1, at both the protein and mRNA levels ($P < 0.05$; Tables 8.6 & 8.7).

Table 8.6: Associations between cytoplasmic and nuclear SLC39A6 protein expression and expression of other oestrogen receptor-related markers.

	SLC39A6 cytoplasmic expression			**SLC39A6 nuclear expression**		
Biomarkers status	**Low** No (%)	**High** No (%)	**X^2** *P*-value	**Low** No (%)	**High** No (%)	**X^2** *P*-value
PgR status						
Low	201 (68)	93 (32)	5.68	201 (68)	93 (32)	25.14
High	213 (59)	146 (41)	0.017	176(49)	184 (51)	<0.0001
FOXA1 status						
Low	184 (64)	104 (36)	1.06	198 (69)	90 (31)	44.21
High	104 (59)	72 (41)	0.301	66 (37)	111 (63)	<0.0001
GATA3 status						
Low	224 (67)	111 (33)	3.36	210 (63)	125 (37)	6.10
High	51 (56)	40 (44)	0.038	44 (48)	47 (52)	0.013
TFF1 status						
Low	146 (67)	71 (33)	7.35	146 (67)	72 (33)	18.99
High	100 (54)	85 (46)	0.007	84 (45)	101 (55)	<0.0001

Table 8.7: Associations between *SLC39A6* mRNA expression and the mRNA expression levels of other oestrogen receptor-related markers.

	***SLC39A6* mRNA expression**		
Biomarker status	**Low No (%)**	**High No (%)**	**X^2 *P*-value**
***PgR* status**			
Low	577 (63)	345 (37)	454.98
High	159 (16)	862 (84)	<0.0001
***FOXA1* status**			
Low	392 (88)	55 (12)	624.46
High	326 (22)	1137 (78)	<0.0001
***GATA3* status**			
Low	533 (76)	164 (24)	694.42
High	186 (15)	1011 (85)	<0.0001
***TFF1* status**			
Low	576 (60)	385 (40)	386.27
High	158 (16)	805 (84)	<0.0001

8.3.5 Outcome analysis

High nuclear SLC39A6 protein expression was associated with longer BCSS in the entire BC cohort (P = 0.001); however, no significant association was observed between cytoplasmic SLC39A6 expression and BCSS (P = 0.217; Figure 8.33). In the ER+ subgroup, both high nuclear and high cytoplasmic SLC39A6 expression were associated with significantly longer BCSS (P = 0.007 & P = 0.015, respectively; Figure 8.4). No associations were identified between SLC39A6 nuclear and cytoplasmic expression and BCSS in the ER- BC subgroup (P > 0.05; Figure 8.5).

The cases in the Nottingham cohort were classified into four groups based on the combination of SLC39A6 nuclear and cytoplasmic expression: (nuclear ^{low}/cytoplasmic ^{low} (n=149)), (nuclear ^{high}/cytoplasmic ^{low} (n=77)), (nuclear ^{high}/cytoplasmic ^{high} (n=78)), and (nuclear ^{low}/cytoplasmic ^{high} (n=47)). The (nuclear ^{high}/cytoplasmic ^{high}) group had better BCSS in comparison to all other subgroups in the entire BC cohort (P = 0.010) and ER+ subgroup (P = 0.014; Figure 8.6).

Similar results were obtained for *SLC39A6* mRNA*,* as high *SLC39A6* expression was associated with favourable BCSS in the entire BC cohort and ER+ tumours (P < 0.001 and P = 0.041; Figure 8.6), but not in ER- tumours (P > 0.05). *SLC39A6* CN loss was also associated with significantly poorer BCSS in the whole cohort and ER+ tumours (both P < 0.0001; Figure 8.7).

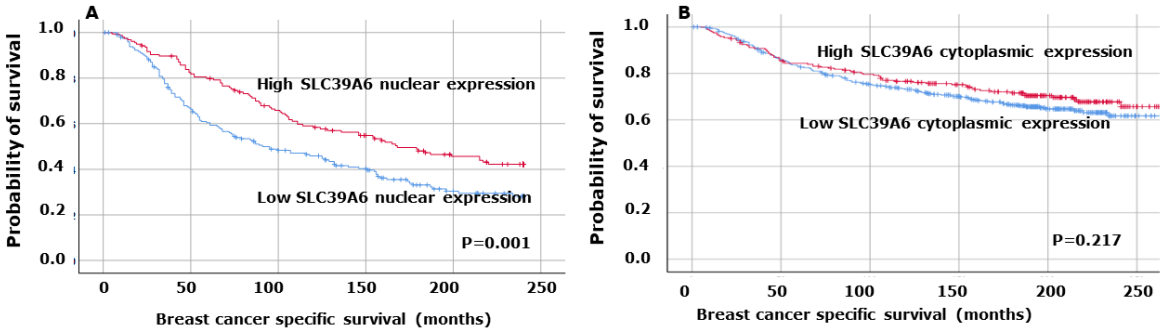


Figure 8.3: Kaplan-Meier survival plots of the associations between high SLC39A6 (A) nuclear and (B) cytoplasmic expression and breast cancer-specific survival (BCSS) in the entire breast cancer cohort.

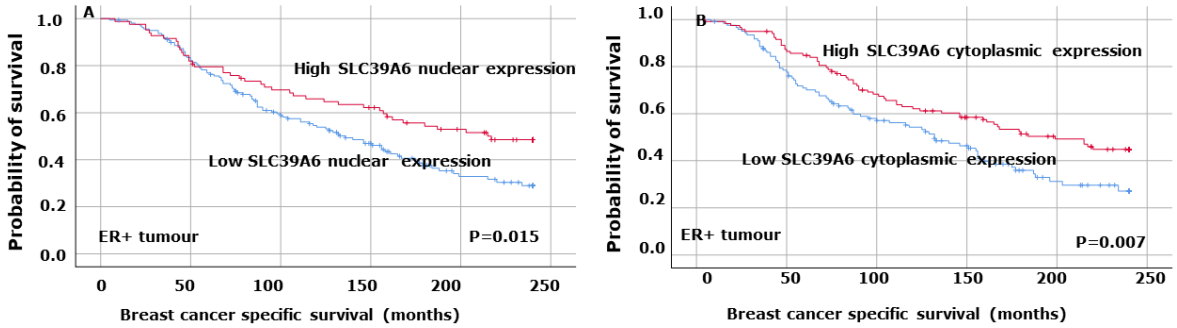


Figure 8.4: Kaplan-Meier plots of the associations between high SLC39A6 (A) nuclear and (B) cytoplasmic protein expression and breast cancer-specific survival (BCSS) in oestrogen receptor positive breast cancer.

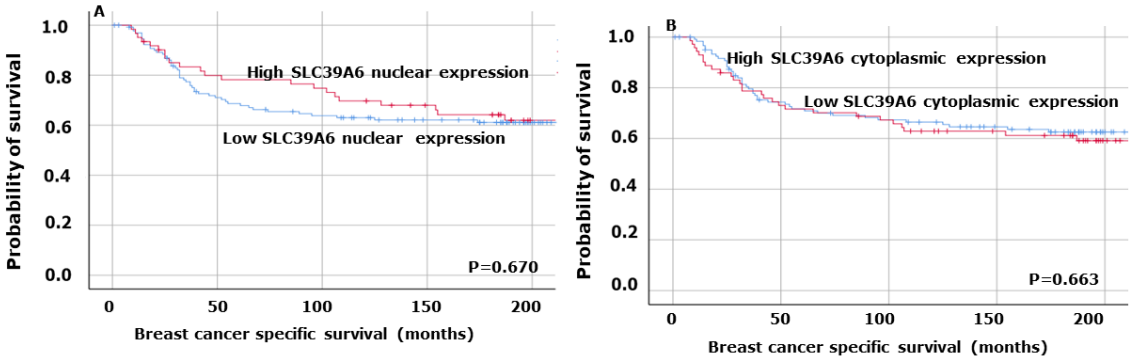


Figure 8.5: Kaplan-Meier plots of the associations between high SLC39A6 (A) nuclear and (B) cytoplasmic protein expression and breast cancer-specific survival (BCSS) in oestrogen receptor -negative breast cancer.

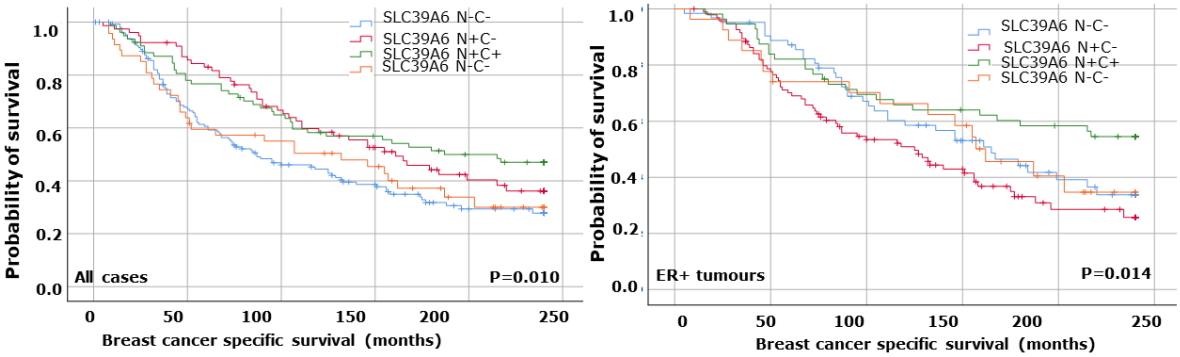


Figure 8.6: Kaplan-Meier survival plots showing the association between the combination of high SLC39A6 cytoplasmic and high SLC39A6 nuclear protein expression and breast cancer-specific survival (BCSS) in (A) the entire breast cancer cohort and (B) oestrogen receptor positive breast

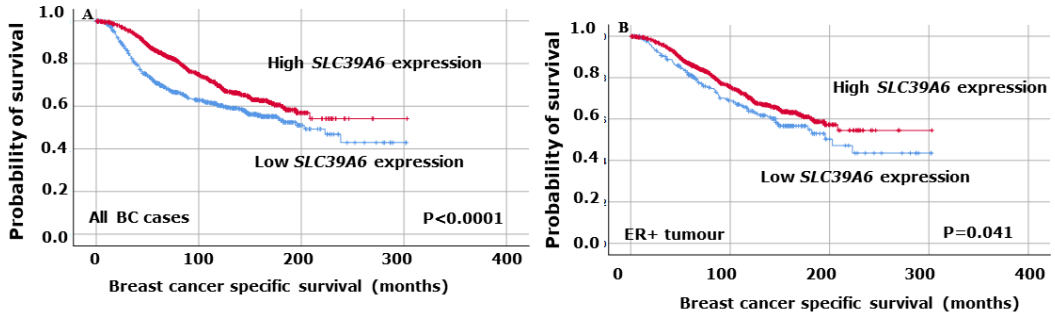


Figure 8.7: Kaplan-Meier survival plots showing the association between high *SLC39A6* mRNA expression and breast cancer-specific survival (BCSS) in (A) the entire BC cohort and (B) oestrogen receptor positive breast cancer.

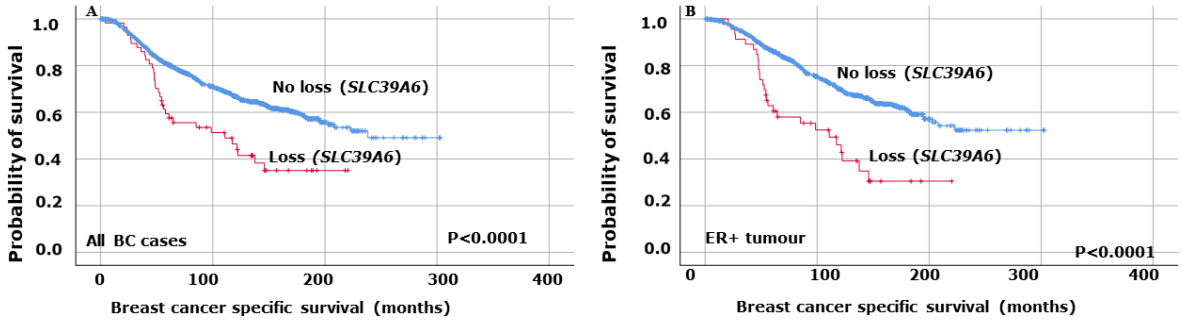


Figure 8.8: Kaplan-Meier survival plots showing the association between *SLC39A6* copy number loss and breast cancer-specific survival (BCSS) in (A) the entire breast cancer cohort and (B) oestrogen receptor positive breast cancer.

8.3.6 Multivariate analysis

In the multivariate Cox regression model including standard prognostic factors such as patient age, tumour grade and tumour size, high SLC39A6 nuclear expression had no significant prognostic value in the entire BC cohort (P = 0.05; Table 8.8). However, high SLC39A6 nuclear expression was an independent prognostic factor for better BCSS in ER+ tumours (P = 0.034, HR 0.678, 95% CI 0.472-0.972; Table 8.9). However, when the Cox regression model included tumour stage, SLC39A6 was not a significant prognostic factor for BCSS (P > 0.05).

Moreover, when the Cox regression model was restricted to ER+ tumours and incorporated other ER-related proteins, only nuclear SLC39A6 expression was an independent prognostic factor for BCSS (P = 0.002; Table 8.10).

In the multivariate Cox regression model including standard prognostic factors such as patient age, tumour grade and tumour size, *SLC39A6* mRNA expression was an independent prognostic factor for BCSS in the entire cohort (P = 0.001, HR = 0.727, 95% Cl = 0.598-0.884) and ER+ tumours (P = 0.029, HR = 0.913, 95% Cl = 0.841-0.991; Table 8.11). The prognostic value of *SLC39A6* mRNA for BCSS in the entire cohort was maintained when other ER-related markers were included in the Cox regression model (P = 0.034, HR = 0.918, 95% Cl = 0.847-0.994; Table 8.12).

Table 8.8: Multivariate Cox regression hazard model for breast cancer specific survival including prognostic clinicopathological parameters and SLC39A6 nuclear protein expression in the entire cohort.

	SLC39A6 nuclear expression		
Parameters	***P*-value**	**Hazard ratio**	**95% CI**
SLC39A6	0.051	0.751	0.56-1.001
Tumour size	<0.0001	1.683	1.261-2.246
Tumour grade	<0.0001	1.89	1.49-2.41
Patient age	0.015	1.41	1.07-1.88

Table 8.9: Multivariate Cox regression hazard model of for breast cancer specific survival including prognostic clinicopathological features and SLC39A6 protein expression in the oestrogen receptor positive subgroup.

parameters	**SLC39A6 nuclear expression**			**SLC39A6 cytoplasmic expression**		
	***P*-value**	**Hazard ratio**	**95% CI**	***P*-value**	**Hazard ratio**	**95% CI**
SLC39A6	0.034	0.678	0.472-0.972	0.050	0.684	0.467-1.001
Tumour size	0.016	1.564	1.088-2.248	0.005	1.662	1.165-2.371
Tumour grade	<0.0001	1.901	1.464-2.470	<0.0001	1.900	1.457-2.478
Patient age	0.005	1.676	1.166-2.410	0.008	1.620	1.135-2.312

Table 8.10: Multivariate Cox regression hazard model for for breast cancer specific survival including other ER-related markers and SLC39A6 protein expression in the oestrogen receptor positive subgroup.

	SLC39A6 nuclear expression			**SLC39A6 cytoplasmic expression**		
parameters	***P*-value**	**Hazard ratio**	**95% CI**	***P*-value**	**Hazard ratio**	**95% CI**
SLC39A6	0.002	1.726	0.233-0.723	0.054	0.568	0.320-1.010
PgR	0.141	0.648	0.364-1.155	0.356	0.760	0.424-1.1362
GATA3	0.650	0.875	0.493-1.556	0.827	0.938	0.530-1.661
FOXA1	0.066	1.726	0.966-3.085	0.508	1.196	0.705-2.028
TIFF1	0.166	0.693	0.413-1.164	0.258	0.743	0.444-1.243

Table 8.11: Multivariate Cox regression hazard model for breast cancer specific survival including other prognostic clinicopathological parameters and *SLC39A6* mRNA expression in the entire cohort and oestrogen receptor positive subgroup.

	***SLC39A6* mRNA in entire cohort**			***SLC39A6* mRNA in ER+ subgroup**		
parameters	***P*-value**	**Hazard ratio**	**95% CI**	***P*-value**	**Hazard ratio**	**95% CI**
SLC39A6	0.001	0.727	0.598-0.884	0.029	0.913	0.841-0.991
Tumour size	<0.0001	2.089	1.665-2.621	<0.0001	2.322	1.715-3.144
Tumour grade	<0.0001	1.353	1.145-1.597	0.001	1.409	1.158-1.715
Patient age	0.360	1.107	0.890-1.377	0.063	1.391	0.983-1.969

Table 8.12: Multivariate Cox regression hazard model for breast cancer specific survival including other oestrogen receptor related markers and *SLC39A6* mRNA expression in the oestrogen receptor positive subgroup.

	***SLC39A6* mRNA expression in ER+ subgroup**		
parameters	***P*-value**	**Hazard ratio**	**95% CI**
SLC39A6	0.034	0.918	0.847-0.994
PgR	<0.0001	0.587	0.468-0.737
GATA3	0.014	0.884	0.801-0.975
FOXA1	0.008	1.109	1.027-1.196
TIFF1	0.663	1.011	0.966-1.059

8.3.7 Associations between SLC39A6 and the response to endocrine therapy

As high SLC39A6 expression was associated with a good prognosis and outcome in patients with ER+ BC, we hypothesised that SLC39A6 may play an important role in the response to hormone therapy targeting the ER. The associations between SLC39A6 expression and RFS, DMFS and BCSS were examined in two subpopulations of patients with ER + BC: patients who received endocrine therapy versus endocrine-therapy naïve patients.

In patients with ER+ BC who did not receive endocrine therapy, high SLC39A6 protein and high *SLC39A6* mRNA expression were associated with longer BCSS (P = 0.001 & P =0.028; Figures 8.8 and 8.9) and DMFS (P = 0.027; only protein expression; Figure 8.10), compared to patients with ER+ BC with low SLC39A6 expression. No significant associations were observed between SLC39A6 protein or mRNA expression in patients with ER+ BC who received endocrine treatment (P > 0.05).

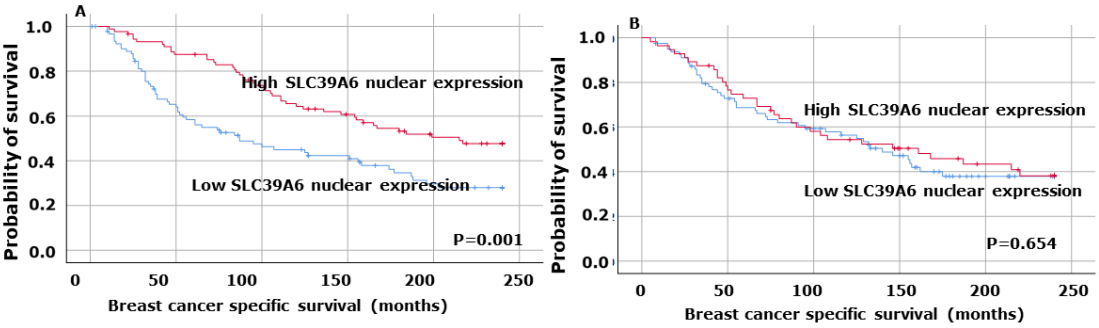


Figure 8.9: (A) Kaplan-Meier survival plots showing the association between high SLC39A6 nuclear expression and longer breast cancer-specific survival (BCSS) in endocrine-therapy naïve patients with ER + BC. (B) No significant association was observed between SLC39A6 nuclear expression and BCSS in oestrogen receptor positive subgroup patients who received tamoxifen.

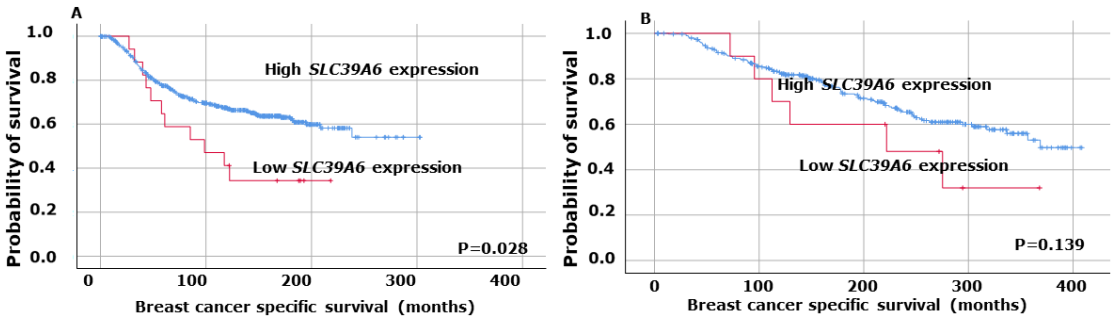


Figure 8.10: (A) Kaplan-Meier survival plots showing the association between high SLC39A6 mRNA expression and longer BCSS in endocrine-therapy naïve patients with oestrogen receptor positive tumours. (B) No significant association was observed between SLC39A6 mRNA expression and BCSS in ER+ patients who received tamoxifen.

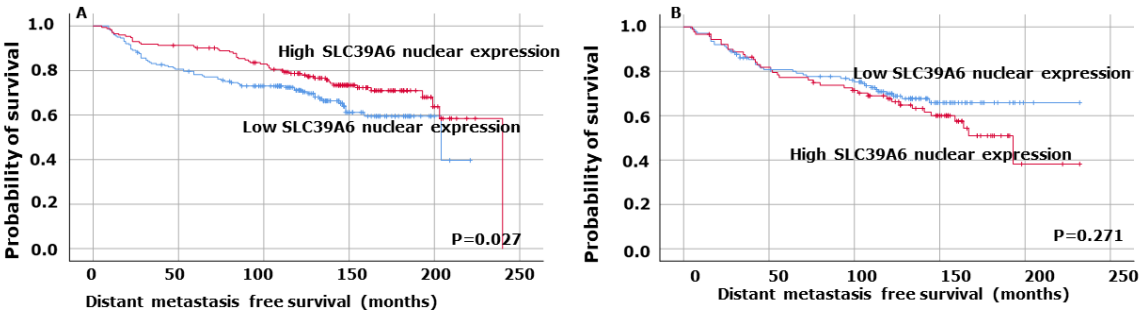


Figure 8.11: (A) Kaplan-Meier survival plots showing the association between high SLC39A6 nuclear expression and longer distant metastasis-free survival (DMFS) in endocrine-therapy naïve patients with oestrogen receptor positive breast cancer. (B) No significant association was observed between SLC39A6 nuclear expression and DMFS in ER+ patients who received tamoxifen, however at 100 month the high expression of SLC39A6 was associated with shorter DMFS.

8.4 Discussion

ER+ tumours constitute up to three-quarters of all cases of invasive BC, and exhibit morphological and molecular heterogeneity and have varied prognoses, therapeutic responses and overall survival rates [512-514]. Better understanding of the varied molecular and biological behaviour of ER+ BC may help to further refine prognostic models and more accurately predict the response of individual patients to adjuvant treatments. In this study, the ER-related marker SLC39A6 was quantified at the proteomic, transcriptomic and genomic levels in two large cohorts in order to assess its prognostic and predictive value in BC.

High cytoplasmic SLC39A6 protein expression, high nuclear SLC39A6 protein expression, and high *SLC39A6* mRNA expression were observed in 43%, 37%, and 50% of BC tumours overall, respectively. Compared to the current study, McClelland *et al.* (1998) reported a higher frequency of high SLC39A6 protein expression (70%) and a lower frequency of high *SLC39A6* mRNA expression (28%) in their BC cohort (n = 44) [515]. Another study (n = 111) detected SLC39A6 using IHC and *situ* hybridisation (I*S*H) in 28% and 53% of BC cases, respectively [510]. Such discrepancies could be due to the inclusion of limited numbers of cases compared to this study, as well as differences in the scoring systems and primary antibodies between studies.

Furthermore, no significant correlation was observed between SLC39A6 protein and mRNA expression in this study, which indicates post-transcriptional regulatory mechanisms are involved in the regulation of SLC39A6 protein expression. Our results regarding the nuclear expression of SLC39A6, and its association with cytoplasmic expression indicate that SLC39A6 functions as not only a zinc exchanger but also as a transcriptional regulator in cancer cells.

SLC39A6 is an oestrogen-inducible gene that is upregulated in ER+ BC [492]. A previous study reported that SLC39A6 expression was associated with ER+ status [493]. The current study confirmed that both high *SLC39A6* mRNA and high SLC39A6 protein expression were observed more frequently in ER+ tumours than ER-negative tumours.

Previous *in vitro* studies suggested that SLC39A6 is associated with progression in BC by facilitating zinc influx into tumour cells, which subsequently promotes tumour growth and the epithelial-to-mesenchymal transition [516, 517]. Hogstrand *et al.* (2013) predicted that the N-terminus of SLC39A6 is located in the endoplasmic-reticulum, and knockdown of *SLC39A6* using a siRNA upregulated STAT3 expression [496], which suggests that SLC39A6 located in the endoplasmic reticulum could play a distinct role in BC.

Moreover, this study indicates that high cytoplasmic and high nuclear SLC39A6 protein expression and high *SLC39A6* mRNA expression were associated with classical clinicopathological parameters related to lower tumour aggressiveness and longer survival, especially in the ER+ subgroup. These observations are not consistent with a previous report that low SLC39A6 protein expression was associated with larger tumour size, high grade and advanced tumour stage in BC [510]. Moreover, high nuclear SLC39A6 expression was an independent prognostic factor for a good outcome in this study.

Therefore, this work suggests cytoplasmic and nuclear SLC39A6 exert distinct roles, especially in ER+ tumours. For example, oestrogen has been suggested to activate and lead to N-terminal cleavage of SLC39A6, which enables trafficking of the protein to the cell membrane [503]. Moreover, high SLC39A6 expression may alter zinc homeostasis in BC cells, which may in turn facilitate tumour cell

metabolism and enable the development and progression of cancer. Overall, this study suggests that the ER inhibits the pro-tumorigenic and EMT-promoting effects of SLC39A6. However, when endocrine therapy inhibits the ER, then the pro-proliferative/pro-EMT effects of SLC39A6 are restored.

The current study also assessed the correlations between the expression of SLC39A6 and other well-known ER-related markers that have been identified as signature genes in ER+ tumours, such as FOXA1 and GATA3 [489, 513]. FOXA1 facilitates the interactions between transcription factors such as the ER and DNA, as FOXA1 can occupy compacted DNA in the absence of other interacting proteins [518]. A previous IHC study identified a strong association between high ER and high FOXA1 expression in BC [519]. Inhibition of GATA3 can direct the ER to DNA binding sites surrounded by FOXA1 [520], a downstream effector of GATA3 [521]. Moreover, TFF1 and PgR have both been identified as informative biomarkers in ER+ tumours and are associated with a better prognosis [522, 523]. This chapter further strengthens the associations between SLC39A6 expression and GATA3, TFF1 and PgR in BC.

This study indicates inhibit the ER expression on patients undergoing endocrine therapy may change the pro-tumorigenic role of SLC39A6 and also suggests that patients with ER+ tumours, and expressing high levels of SLC39A6 may not be suitable candidates for long-term endocrine therapy. McClelland *et al.* (1998) previously reported that SLC39A6 was associated with the response of ER+ tumours to endocrine therapy [515].

8.5 Conclusion

SLC39A6 may have prognostic value in BC, especially in ER+ tumours. Further assessment of SLC39A6 may help to more accurately stratify patients with ER+ BC and identify patients who may achieve a good outcome.

Chapter 9

Heterogeneity and prognostic significance of tumour infiltrating lymphocytes (TILs) in invasive breast cancer

9.1 Introduction

The breast cancer (BC) microenvironment is comprised of a complex mixture of cell types, including immune modulating cells that provide an opportunity for therapeutic modulation [524, 525]. Several studies have reported tumour-infiltrating lymphocytes (TILs)—as assessed using haematoxylin and eosin (H&E)-stained sections—have prognostic and predictive value in BC, regardless of the cellular subset [526-529].

Both intratumoral and stromal TILs have been assessed in the tumours of patients receiving neoadjuvant and adjuvant treatment. For instance, Denkert *et al.* (2010) demonstrated that assessment of intratumoural and stromal TILs in core biopsies could be used to predict the response to chemotherapy in invasive BC [530]. Ono *et al.* (2012) quantified stromal TILs in core biopsies by determining the percentages of TILs. High TILs was associated with a better response to chemotherapy in TNBC [531]. Moreover, Nummer *et al.* (2013) reported that intratumour, stromal and lymphocyte-predominant breast cancer (LPBC) TILs have predictive value for the response to anthracycline/taxane-based neoadjuvant chemotherapy (NACT) in BC [532]. Using samples from two national randomised clinical trials, Adams *et al*. (2014) demonstrated that stromal TILs have prognostic value for patients with TNBC treated with adjuvant chemotherapy [529].

Although TILs have been reported to have prognostic and predictive value in BC, a variety of methods have been used to assess TILs. Therefore, the International TILs Working Group guideline recommendations were devised to harmonise the assessment of TILs [533]. However, the impact of overall TIL heterogeneity, including the density of TILs in different regions of individual tumours and across the various subtypes of BC, as well as the precise roles of specific TIL subsets, remain to be defined.

The presence of varied subsets and spatial distributions of immune cells may contribute to spatial heterogeneity in BC. Moreover, intratumor heterogeneity with respect to the actual tumour cells may also influence the spatial distribution and subsets of immune cells through varied cytokine secretion and neoantigen presentation by different subpopulations of tumour cells [534].

The various subsets of TILs play different roles in tumour progression. Some TIL subsets may suppress tumour growth by destroying cancer cells, whereas other subsets can promote tumour progression by selecting tumour cells that could survive in an immunocompetent host [95, 535, 536]. For instance, T lymphocytes account for the majority of TILs in invasive BC and can activate host defence mechanisms [95], which are mainly mediated by cytotoxic CD8+ T cells. However, T regulatory cells (Tregs) play a vital role in immune tolerance to tumour cells and can protect against tumour development and progression [536-538].

The immune checkpoint proteins PDL-1 and PD1 have also recently received much attention in cancer [539-541]. However, the contribution of various immune cell subsets and expression of immune checkpoint proteins to

the prognostic value of the overall density of TILs assessed using H&E slides also remain unclear.

9.1.1 Hypothesis

The hypotheses tested in this chapter are:

1- Spatial and temporal TIL heterogeneity exist in BC.

2- TILs have varied prognostic value in different subtypes of BC, as well as in primary and recurrent tumours.

9.1.2 Aims

The aims of the current chapter are to:

(i) Assess the spatial and temporal distribution of TILs in invasive BC.

(ii) Assess the prognostic significance of TILs in different subtypes of invasive BC.

(iii) Assess the correlations between the overall percentage of TILs and the cell counts for various TIL subsets in BC.

9.1.3 Objectives

The objectives of this chapter are to:

(i) Assess the distribution of TILs in H&E-stained primary BC tissues and the corresponding recurrent tumours.

(ii) Determine the prognostic value of TILs in different molecular subtypes of BC.

(iii) Investigate the contribution of various TIL subsets based on expression of immune markers and immune checkpoint inhibitors as molecularly define

to overall TILs, and explore the prognostic significance of these subsets in various subtypes of BC.

9.2 Materials and Methods

9.2.1 Characteristics of the study cohorts

The study cohort was comprised of 700 patients with invasive BC who presented to Nottingham City Hospital between 1988 and 1998 (see Chapter 2 for details). Full-face H&E stained sections (n = 700) were assessed to determine the prognostic significance of TILs in BC.

Inter-slide TIL heterogeneity was investigated by assessing H&E-stained slides from multiple FFPE archival tissue blocks (3-4 blocks/case) for a subgroup of patients (n = 52/700). Intra-slide TIL heterogeneity (n = 230 slides) was assessed by examining different regions of the same tumour (i.e., in four quadrants of the same slide) for 230/700 patients.

The distribution of TILs in primary tumours and the corresponding recurrent tumours was investigated by assessing the tumour specimens for 44 patients for whom recurrent tumour blocks were available.

9.2.2 Availability of other data on immune markers for the cohort

Expression of the immune biomarkers CD3, CD8, FOXP3, CD20 and CD68 and were previously examined using IHC by Dr Mahmoud, a PhD student in our laboratory [231-234, 493]. Expression of the immune checkpoint inhibitors PDL-1 and PD1 were assessed using IHC in this study by Eleanor Grace Hodgson, a master student in our group. All markers were examined using the antibodies listed in Table 2.2.

The total numbers of each type of immune cell in the stroma surrounding the tumour were counted in each core. PD1 and PDL-1 expression were scored as the percentages of PD1+ TILs, PDL-1+ TILs, and PDL-1+ tumour cells. The

cut-offs were used (≥ 1%) based on previous work [542]. The immunostaining of PD1 and PDL-1 were performed according to the method described in chapter 2.

9.2.3 Assessment of TILs in H&E-stained sections

The tumour specimens were routinely fixed in 10% formalin, embedded in paraffin and full-face 4 µm-thick sections were prepared for H&E staining as described in detail in Chapter 2. All sections contained adequate tumour burden, with tumour cells occupying at least 40% of the surface area in each section. The H&E slides were scanned into high-resolution digital images at 20x magnification using a 3D Histech Pannoramic 250 Flash II scanner (3DHISTECH Ltd., Budapest, Hungary). The digital images were viewed using a 3D Histech Pannoramic Viewer (3DHISTECH Ltd., Budapest, Hungary).

TILs were semi-quantitatively assessed following the method recommended by the 2014 International TILs Working Group [533]. Each section was initially assessed at low magnification. The region of interest was determined as the border of the invasive tumour with the stromal compartment to view TILs located around and dispersed in the stroma between carcinoma cells, but not directly contacting tumour cells. TILs around crush artefacts, tumour necrosis, extensive central regressive hyalinisation or DCIS/normal lobules were not considered for quantification. All mononuclear cells, lymphocytes and plasma cells, if any, were scored. The number of TILs occupying the stroma was recorded as a percentage of all stromal cells, and termed average TILs (AV-TILs).

To further quantify intra-slide and inter-slide TIL heterogeneity, areas with dense TIL infiltrates were identified and designated 'hotspot-TILs' (HS-TILs). HS-TILs were defined as dense clusters of lymphocytes containing the highest number of lymphocytes. HS-TILs was expressed as a percentage of all stromal cells in the hotspot region. HS-TILs was only equal to AV-TILs in cases showing no or occasional scattered lymphocytes, or in the rare cases with dense lymphocytic infiltration throughout the entire tumour.

To assess intra-slide TIL heterogeneity, the digital images were annotated by dividing the tumour area into four quadrants using vertical and horizontal lines. AV-TILs and HS-TILs were assessed in each quadrant separately, and also in the whole slide.

All cases were scored independently by other observer who was blinded to the histopathological data and patient outcomes. The observers were trained by a consultant pathologist. Furthermore, to determine intra-observer reproducibility for the scoring system, 50% of cases were randomly selected and rescored by the same scorer.

9.2.4 Statistical analysis

The percentage of AV-TILs was categorised using a cut-off of 10%; scores < 10% were defined as negative/low and scores ≥ 10% were defined as positive/high [543]. In recurrent tumours, the median percentage of TILs (5%) was used as the cut-off. IHC staining for CD3, CD8, CD20, FOXP3, CD68, PDL-1 and PD1 were scored using the cut-offs defined in previous studies [231-234, 544].

Intra-observer agreement for TILs was determined using the Pearson correlation coefficient and inter-observer agreement was determined using the intra-class correlation coefficient. The Friedman test was used to assess inter-slide TIL heterogeneity (i.e., in four slides from the same case), intra-slide TIL heterogeneity (i.e., in four quadrants of the same slide) and TIL heterogeneity between primary and recurrent tumours.

The Spearman's rank correlation coefficient was used to assess the correlations between the percentage AV-TILs or HS-TILs and the numbers of TILs in various subsets.

The associations between the categorical groups for the percentages of TILs, PD1+ and PDL-1+ cells and the clinicopathological features of the patients were analysed using the Chi-square test. Associations with patient outcomes were assessed using Kaplan-Meier survival curves and the log-rank test. Cox proportional hazards regression models were built for multivariate survival analyses to estimate the hazard ratios (HR) for TILs and immune cell subsets when adjusted for other established prognostic factors. Two-tailed *P*-values less than 0.05 were considered significant for all tests.

9.3 Results

9.3.1 Heterogeneity of TILs at the morphological level

Inter-observer agreement for the percentage of AV-TILs between the two scorers for four slides from each patient was determined using the intra-class correlation coefficient. The intra-class correlation coefficients of 0.771 for slide 1, 0.580 for slide 2, 0.682 for slide 3 and 0.910 for slide 4 indicated good to excellent inter-observer agreement. The 4 slides were taken from different regions of the patient tumour for assessing TILs.

High reproducibility of TIL scoring was obtained for the same observer (Pearson correlation coefficient: 0.93 for AV-TILs and 0.97 for HS-TILs) and between the two observers (intra-class correlation coefficient: 0.79 for AV-TILs and 0.83 for HS-TILs; kappa agreement: 0.825 for AV-TILs and 0.785 for HS-TILs).

No significant differences in the percentage of AV-TILs were identified between slides from different tumour blocks/areas from the same patient (inter-slide variation, $P = 0.251$ for AV-TILs and $P = 0.162$ for HS-TILs; Figure 2). However, when each slide was divided into four quadrants and each quadrant was scored independently, the variation in the percentages of TIL was statistically significant (intra-slide variation, $P = 0.019$ for AV-TILs and $P < 0.0001$ for HS-TILs).

The mean, median and range of percentages of AV-TIL and HS-TIL for primary tumours of different molecular subtypes and a subgroup of primary tumours and the corresponding recurrent tumours are displayed in Tables 9.1 and 9.2. Luminal A tumours exhibited more homogeneity in respect of the inter-slide

AV-TIL than the other subtypes. The percentage of AV-TILs was higher in TNBC than in the other subtypes ($P < 0.001$).

Interestingly, trends towards a lower average percentage of TILs ($P = 0.052$; Table 9.1) and a more homogeneous distribution of TILs were observed in recurrent tumours compared with the corresponding primary tumours (Figure 9.2).

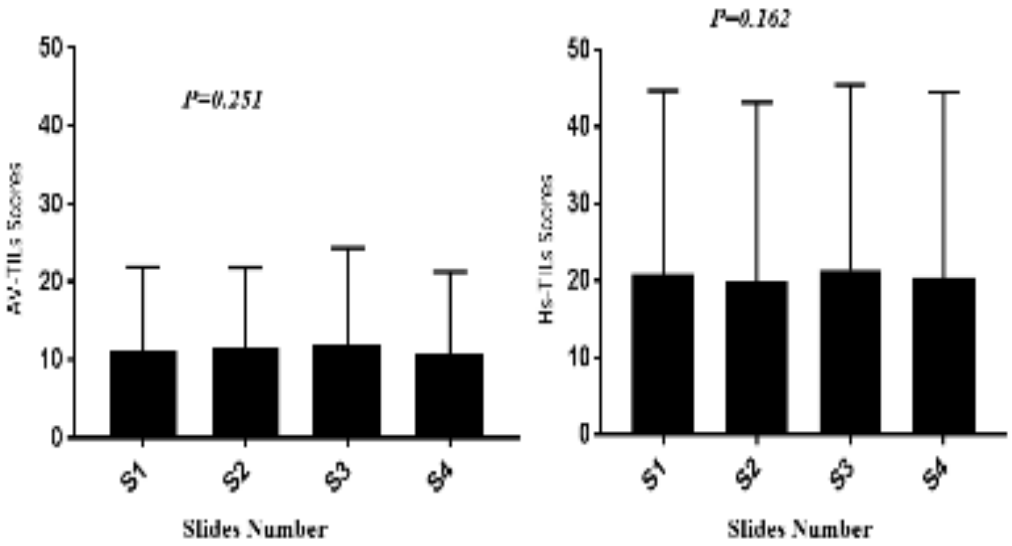


Figure 9.1: Average percentages of TILs (AV-TIL; left) and percentages of hotspot TILs (HS-TIL; right) for multiple tumour blocks. S1, 2, 3 & 4 represent sections 1, 2, 3, & 4 from four different tumour blocks of the same primary tumours.

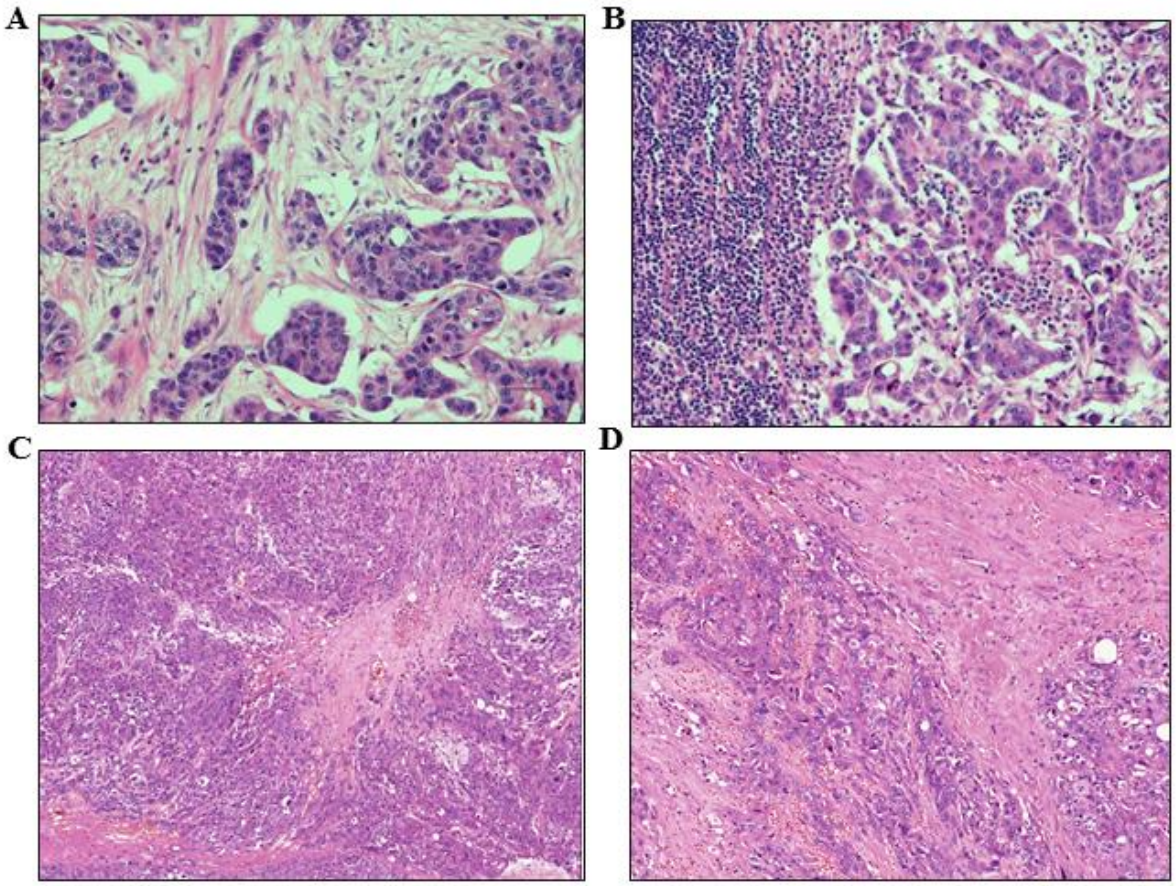


Figure 9.2: Representative photomicrographs of TILs infiltrating a primary tumour, and corresponding recurrent tumour from the same patient, which indicates the higher proportion of TILs in primary tumour compared to the corresponding recurrent. (A, B) show TILs in the two regions of primary tumour with same slide, including the cold (A) and hot (B) spots. (C, D) show the TILs in two regions of corresponding recurrent tumour within the same slide, with no differences of TILs in both regions.

Table 9.1: Mean, median and range of percentages of TILs in primary BC and the corresponding recurrent tumours for the subgroup of the study cohort with primary and recurrent tumour blocks (*n* = 44).

% Stromal TILs/tumour	Primary tumour			Recurrent tumour		
	Mean	**Median**	**Range**	**Mean**	**Median**	**Range**
AV-TILs	22	15	1-80	9	5	1-45
HS-TILs	38	28	1-90	13	7	1-60

Table 9.2: Mean, median and range of percentages of TILs in different molecular subtypes of BC.

% Stromal TILS/tumour	Luminal A			Luminal B			HER2-enriched			Triple negative		
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
AV-TILs	8	5	1-40	12	10	3-40	16	12	5-50	27	25	1-85
HS-TILs	16	8	2-90	22	12	3-90	33	20	6-90	45	40	1-90

9.3.2 Immune cell subset counts and overall TILs in BC

Expression of the lymphocytic biomarkers CD3, CD8, CD20 and FOXP3 and macrophage/histiocytic biomarker CD68 were previously evaluated in the whole cohort. This data was used to further assess specific subsets of TILs in BC. Table 9.3 summarizes the numbers of cells expressing each immune marker for the whole cohort and within different molecular subtypes of BC.

The T lymphocyte population, as demonstrated by the pan-T cell marker CD3, was the main component of TILs, followed by macrophages (CD68+ cells). B lymphocytes (CD20+ cells) represented the smallest component of TILs. Among the T lymphocyte subset, CD8+ cells were more frequent than FOXP3+ cells.

Though the TNBC subtype contained higher numbers of all subsets of TILs than other subtypes of BC, the ratio of immune cells in each subtype of BC appeared similar across the whole cohort, with the CD3+ subset being the major component in all subtypes of BC (Figure 9.3).

Multiple significant correlations were identified between the percentages of AV-TILs and the cell counts for the TIL subsets expressing each immune cell marker (all $P < 0.0001$; Figure 9.3)

Table 9.3: Mean, median and range of immune cell subsets in core diameters, assessed as absolute numbers of cells in each subset, for the entire cohort and different molecular subtypes of BC.

Stromal immune cell density	Entire cohort			Luminal A			Luminal B			HER2-enriched			Triple negative		
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
CD3+	55	33	0-450	34	22	0-258	44	25	0-317	73	51	0-425	93	67	0-450
CD8+	25	11	0-242	16	8	0-142	22	10	0-317	31	17	0-173	40	23	0-242
FOXP3+	14	7	0-153	6	2	0-55	11	5	0-68	19	13	0-153	27	21	1-128
CD20+	13	1	0-415	7	0	0-211	11	1	0-253	22	3	0-415	24	5	0-255
CD68+	32	24	0-252	18	15	0-150	28	23	1-215	38	32	3-252	51	42	1-225

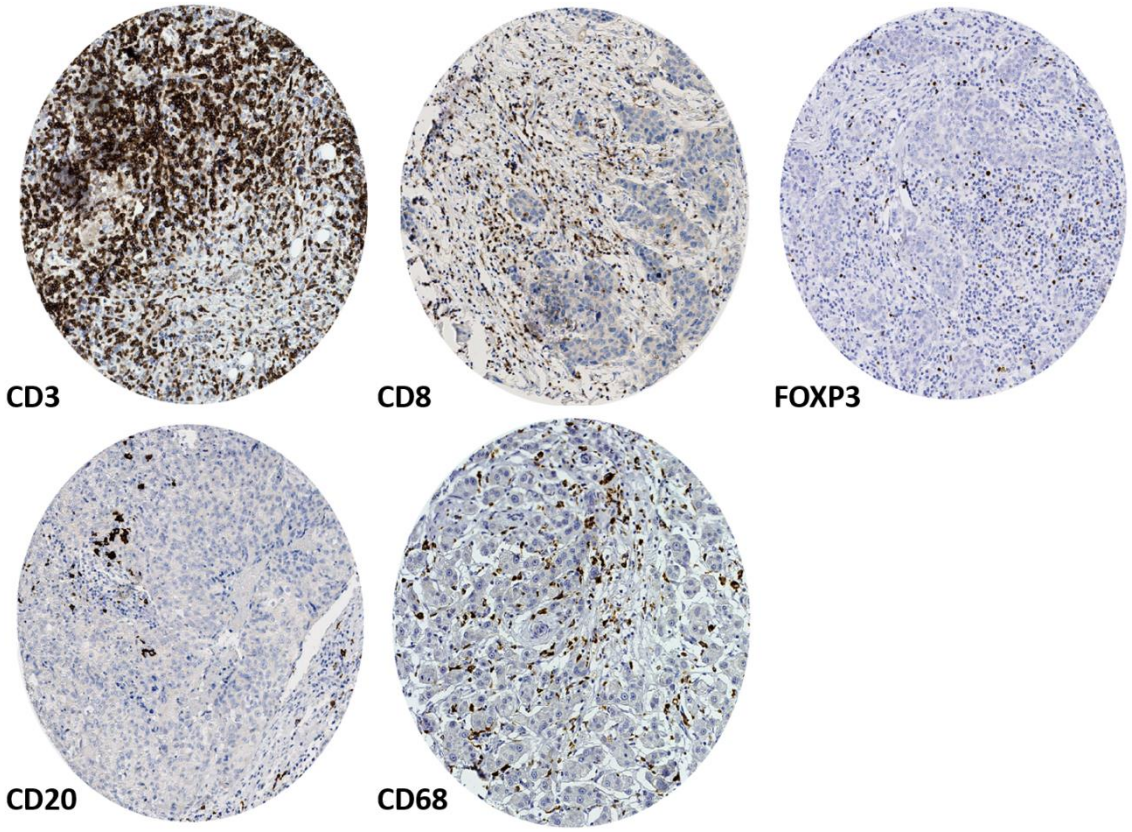


Figure 9.3: Representative photomicrographs of the expression of the immune cell markers CD3, CD8, FOXP3, CD20, and CD68 in BC.

9.3.3 PDL-1 and PD1 expression and TILs in BC

High PDL-1 expression (≥ 1%) was observed in the cytoplasm of tumour cells in 863/1061 (81%) cases and in the cell membrane of tumour cells in 47/1061 (4%) cases in the entire cohort. PDL-1+ was also expressed in TILs: high PDL-1 in TILs was detected in 178/1061 (16%) of the BC cases. PD1+ expression was only observed in TILs; 601/1242 (48%) of the BC cases expressed high levels of PD1 in TILs.

With respect to the molecular subtypes of BC, high cytoplasmic and membrane expression of PDL-1 in tumour cells were more frequently observed in TNBC (mean, 90%; mean, 7%) than the other subtypes. Moreover, the TNBC subtype exhibited high percentages of PD1+ TILs (mean, 30%) and PDL-1+ TILs (mean, 19%), while the luminal A subtype contained lower percentages of PD1+ TILs (mean, 10%) and PDL-1+ TILs (mean, 6%). Tables 9.4 and 9.5 show the mean, median, and range of percentages of PDL-1+ in tumour cells, and TILs as well as PD-1+ TILs.

Both the percentage of HS-TILs and the percentage of AV-TILs correlated significantly with the percentage of PD1+ TILs ($P = 0.031$ and $P = 0.032$), while only the percentage of AV-TILs correlated with the percentage of PDL-1+ TILs ($P < 0.0001$). PDL-1+ tumour cells were not significantly correlated with either HS-TILs or AV-TILs ($P > 0.05$). The percentage of PDL-1+ tumour cells (cytoplasmic expression) was positively correlated with the cell counts for CD3+, CD8+, CD68+, CD20 and FOXP3+ immune cells (all $P < 0.0001$). Moreover, PD1+ and PDL-1+ TILs correlated with the cell counts for CD3+, CD8+, CD68+, CD20+ and FOXP3+ immune cells (all $P < 0.0001$; Figure 9.3).

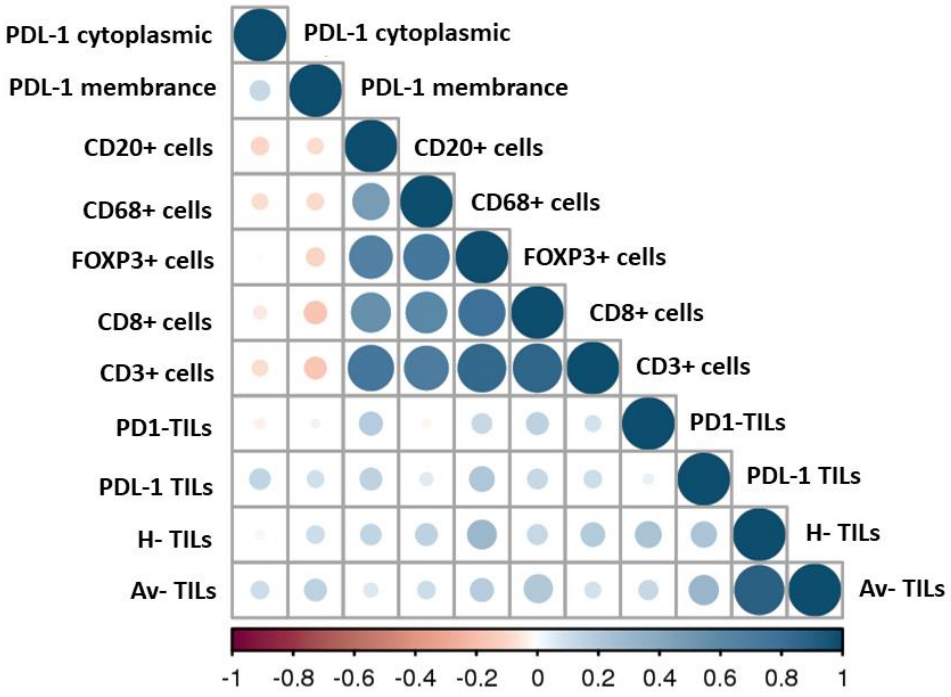


Figure 9.3. Correlation matrix summarising the correlations between the cell counts for immune markers and percentages of AV-TILs in the entire study cohort. Positive correlations are indicated in blue and negative correlations in red. The intensity of colour is proportional to the correlation coefficient.

Table 9.4: Mean, median and range of percentages of PDL-1+ tumour cells in different molecular subtypes of BC.

	% PDL-1+ tumour cells (cytoplasmic expression)			% PDL-1+ tumour cells (membrane expression)		
BC subtype	Mean	Median	Range	Mean	Median	Range
Whole cohort	78	80	0-300	3	0	0-200
Luminal A	66	70	0-255	2	0	0-150
Luminal B	81	80	0-280	1	0	0-70
HER2-enriched	85	80	0-300	4	0	0-120
Triple negative	90	90	0-280	7	0	0-200

Table 9.5: Mean, median and range of percentages of PDL-1+ TILs and percentages of PD1+ TILs in different molecular subtypes of BC.

	% PDL-1+ TILs			% PD1+ TILs		
BC subtypes	Mean	Median	Range	Mean	Median	Range
Whole cohort	21	10	0-100	11	5	0-85
Luminal A	10	7	0-50	6	0	0-75
Luminal B	20	10	0-90	10	5	0-70
HER2-enriched	25	15	1-100	13	5	0-70
Triple negative	30	25	2-90	19	15	0-85

9.3.4 Prognostic significance of TILs at the morphological level in BC

In primary BC, a high percentage of AV-TILs was associated with advanced tumour grade ($P < 0.0001$), advanced nodal stage ($P = 0.002$), poor NPI ($P < 0.0001$) and lymphovascular invasion ($P < 0.0001$), as well as the TNBC ($P < 0.0001$) and HER2+ subtypes ($P = 0.035$; Table 9.6). Moreover, a high percentage of AV-TILs was positively associated with tumour grade in TNBC ($P = 0.043$).

A high percentage of HS-TILs was significantly associated with advanced tumour grade ($P < 0.001$), advanced nodal stage ($P = 0.038$), poor NPI ($P < 0.0001$) and the TNBC subtype ($P < 0.0001$).

Univariate survival analysis showed no significant associations existed between high/low percentages of AV-TILs and breast cancer-specific survival (BCSS) or distant metastasis-free survival (DMFS; both $P \geq 0.05$) in the entire BC cohort (Figure 9.4).

However, a high percentage of AV-TILs was significantly associated with longer BCSS in the TNBC subgroup ($P = 0.026$). A high percentage of AV-TILs was also associated with longer DMFS in TNBC ($P = 0.017$; Figure 9.5), but not in the other subtypes of BC. No significant associations were identified between the percentage of HS-TILs and BCSS or DMFS in other molecular subtypes of BC ($P \geq 0.05$).

Interestingly, in recurrent tumours, a significant association was observed between a high percentage of AV-TILs and shorter post-recurrence patient survival ($P = 0.004$; Figure 9.6). However, in the subgroup of patients with recurrent tumours, no significant association was observed between the percentage of AV-TILs in the primary tumours and BCSS ($P > 0.05$).

As the percentage of AV-TILs was a significant prognostic factor for survival outcomes in TNBC, further analyses were limited to TNBC. Cox-

proportional models including other prognostic parameters such as patient age, tumour grade and nodal stage showed that a high percentage of AV-TILs was an independent prognostic factor for better BCSS (*P* = 0.004, HR 0.4, 95% CI 0.2–0.8; Table 9.7).

AV-TILs was also an independent prognostic factor for the response to adjuvant chemotherapy in TNBC (*P* = 0.02, HR = 0.438, 95 CI = 0.218-0.880).

Table 9.6: Associations between the percentage of average-tumour infiltrating lymphocytes (AV-TILs) and the clinicopathological features of primary BC.

Clinicopathological parameters	**Low TILs *N* (%)**	**High TILs *N* (%)**	x^2 ***P*-value**
Patient age (years)			
< 50	153 (57)	118 (43)	5.962
≥ 50	281 (66)	147 (34)	**0.015**
Tumour size (cm)			
< 2	272 (61)	171 (39)	0.208
≥ 2	161 (63)	94 (37)	0.649
Lymphovascular invasion			
Negative	327 (66)	167 (34)	12.420
Positive	106 (52)	98 (48)	**<0.0001**
Tumour Grade			
Grade I	65 (92)	6 (8)	127.131
Grade II	180 (85)	30 (14)	**<0.0001**
Grade III	45 (45)	229 (55)	
Axillary nodal stage *			
Stage I	276 (67)	139 (33)	12.107
Stage II	113 (59)	79 (41)	**0.002**
Stage III	43 (49)	47 (52)	
Nottingham Prognostic Index			
Good	157 (90)	17 (10)	86.578
Moderate	210 (57)	161 (43)	**<0.0001**
Poor	66 (43)	87 (57)	
Oestrogen Receptor			
Negative	70 (29)	168 (71)	163.702
Positive	364 (79)	97 (21)	**<0.0001**
Progesterone Receptor			
Negative	183(47)	203 (53)	79.107
Positive	249(80)	61 (20)	**<0.0001**
HER2 status			
Negative	306(64)	169 (36)	3.596
Positive	127(57)	96 (43)	**0.035**
Triple negative status			
Non-triple negative	400(74)	144 (26)	138.335
Triple negative	33(21)	121 (79)	**<0.0001**

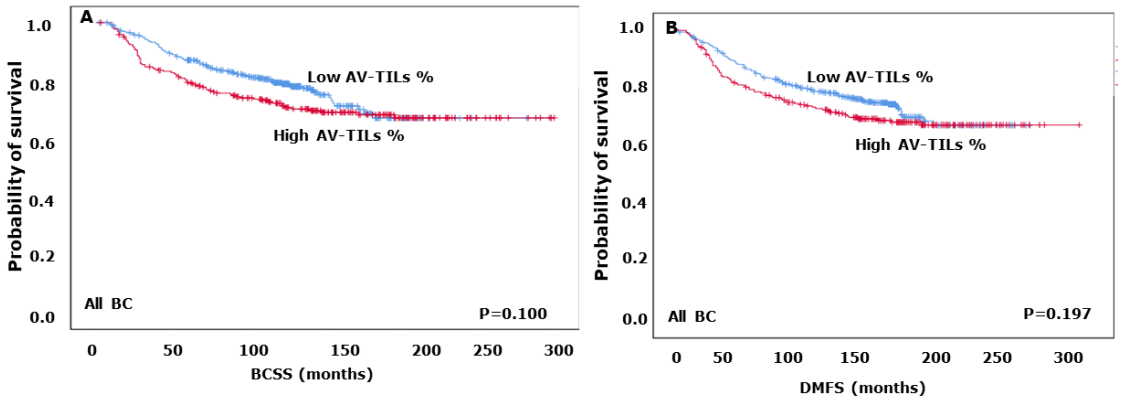


Figure 9.4: Kaplan-Meier survival curves of the associations between percentage AV-TILs and (A) breast cancer-specific survival (BCSS) and (B) distant metastasis-free survival (DMFS) in the entire cohort.

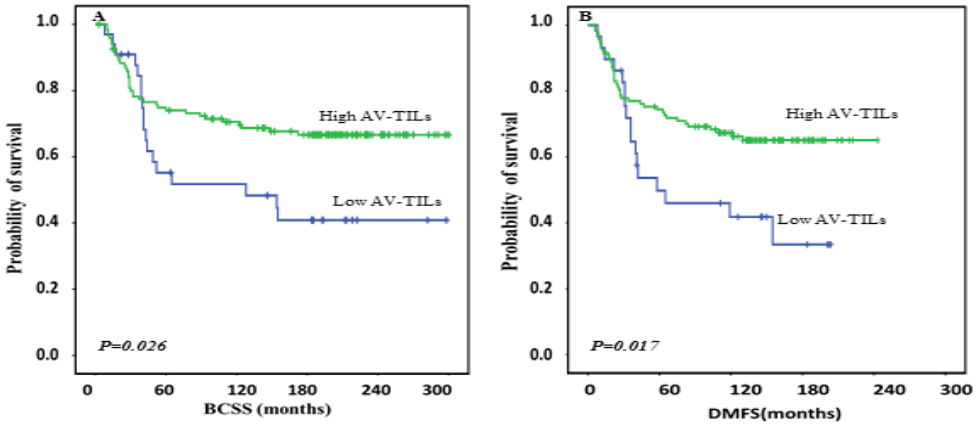


Figure 9.5: Kaplan-Meier survival curves of the association between percentage AV-TILs and (A) breast cancer-specific survival (BCSS) and (B) distant metastasis-free survival (DMFS) in triple negative breast cancer.

Table 9.7: Multivariate Cox regression hazard model including other prognostic clinicopathological parameters for the prognostic value of percentage AV-TILs and breast cancer-specific survival in triple negative breast cancer.

Parameters	Hazard ratio	95% Confidence interval		*P*-value
		Lower	Upper	
Patient age	0.7	0.6	4.2	0.169
Tumour grade	2.9	1.1	3.1	0.134
Nodal stage	2.0	0.5	2.2	<0.001
Tumour size	1.3	0.7	2.3	0.291
TILs-AV	0.4	0.2	0.7	0.004

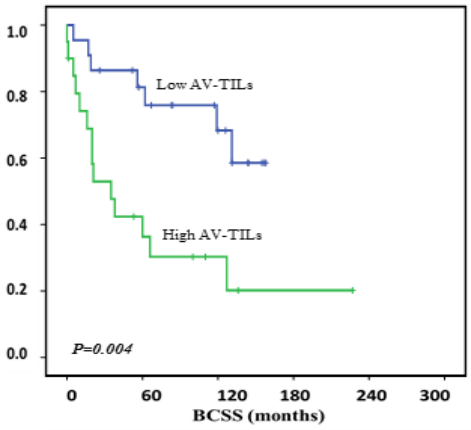


Figure 9.6: Kaplan-Meier survival curve of the association between percentage average (AV)-TILs in recurrent breast cancer and post-recurrence survival.

9.3.5 Prognostic significance of immune cell subsets in BC

Univariate survival analysis demonstrated that high FOXP3+ and CD68+ cell counts were significantly associated with shorter BCSS in the entire cohort (P = 0.006 and $P < 0.001$, respectively; Figure 9.7). However, a trend towards an association between high CD20+ cell counts and longer BCSS was observed in the entire cohort (P = 0.066).

High CD3+ (P = 0.021, P = 0.011), CD8+ (P = 0.017, P = 0.012) and CD20+ cell infiltrates (P = 0.037, P = 0.045) were associated with longer BCSS and DMFS in TNBC (Figure 9.8).

Additionally, a high number of histiocytic CD68+ cells was associated with poorer BCSS in the luminal B subtype (P = 0.04) and a high CD20+ cell count was associated with better BCSS in HER2+ tumours (P = 0.04; Figure 9.9).

Cox-proportional models including patient age, tumour grade, nodal stage and various immune cell subtypes showed only high CD8+ (P = 0.002, HR 0.5, 95% CI 0.3–0.8) and high CD3+ (P = 0.003, HR 0.3, 95% CI 0.2–0.7) cell counts were independent prognostic factors for a good outcome in TNBC, while the numbers of CD20+, CD68+ and FOXP3+ cells were not significantly associated with any outcome in TNBC (Table 9.8).

In the Cox proportional model including the percentages of TILs and various subsets of immune cells (CD3, CD8, FOXP3, CD20 and CD68), only the percentage of AV-TILs was significantly associated with BCSS (Table 9.6).

The prognostic value of combinations of high/low percentages of AV-TILs or HS-TILs with high/low numbers of cells expressing immune biomarkers was also assessed. However, no significant associations were observed between any combination for any outcome in the entire cohort or any subtype of BC.

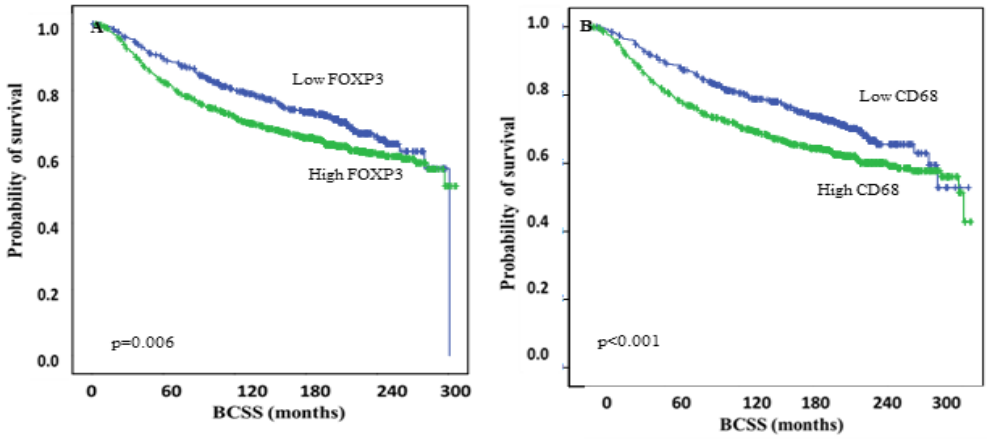


Figure 9.7: Kaplan-Meier survival curves of the associations between (A) the numbers of FOXP3+ lymphocytes and (B) CD68+ histiocyte and breast cancer-specific survival (BCSS) in the entire cohort.

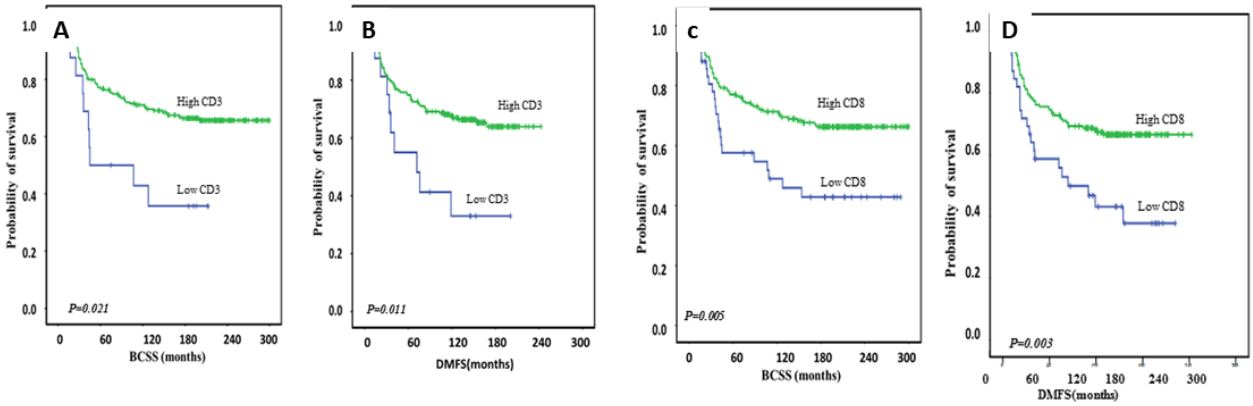


Figure 9.8: Kaplan-Meier survival curves of the associations between the number of CD3+ lymphocytes and (A) breast cancer-specific survival (BCSS) and (B) distant metastasis-free survival (DMFS), the numbers of CD8+ lymphocytes and (C) BCSS and (D) DMFS, and the numbers of CD20+ lymphocytes and (E) BCSS and (F) DMFS in triple negative breast cancer.

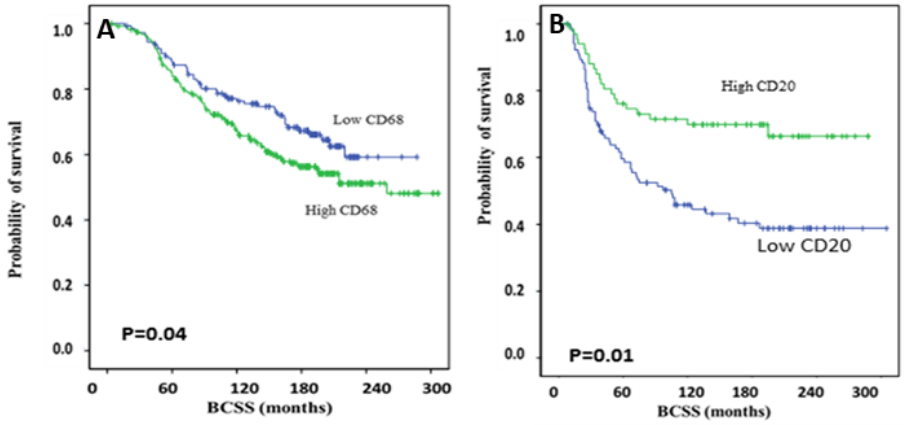


Figure 9.9: Kaplan-Meier survival curves of the associations between (A) the numbers of CD68+ cells and breast cancer-specific survival (BCSS) in the luminal B subtype and (B) the numbers of CD20+ lymphocytes and BCSS in the HER2+ subtype.

Table 9.8: Multivariate Cox regression hazard model including other prognostic clinicopathological parameters of the prognostic value of immunohistochemically defined CD3+ or CD8+ T lymphocyte counts for breast cancer-specific survival in TNBC.

Parameters	Hazard ratio	95% confidence intervals		*P*-value
		Lower	Upper	
CD3+ T lymphocytes				
Patient age	0.6	0.4	1.0	0.055
Tumour grade	1.9	0.6	5.8	0.252
Nodal stage	1.7	1.2	2.2	<0.001
Tumour size	1.5	0.9	2.4	0.081
CD8+	0.4	0.2	0.7	0.002
CD8+ T lymphocytes				
Patient age	0.6	0.4	1.0	0.08
Tumour grade	1.4	0.7	3.2	0.998
Nodal stage	1.6	1.2	2.1	0.001
Tumour size	1.4	0.9	2.3	0.097
CD3+	0.3	0.1	0.6	0.003

Table 9.9: Multivariate Cox regression hazard model including the numbers of various immunohistochemically defined immune cell subsets for the prognostic value of average TILs for breast cancer-specific survival in TNBC.

Parameters	Hazard ratio	95% Confidence intervals		*P*-value
		Lower	Upper	
CD3+	0.6	0.1	2.8	0.487
CD8+	0.5	0.2	1.1	0.088
FOXP3+	1.8	0.5	6.1	0.313
CD20+	1.1	0.5	2.5	0.724
CD68+	2.3	0.7	7.5	0.140
AV-TILs	0.3	0.1	0.8	0.023

9.3.6 Prognostic significance of PDL-1 and PD1 expression in BC

A high percentage of PD1+ TILs and high percentage of PDL-1+ TILs were significantly associated with larger tumour size ($P = 0.010$ and $P = 0.012$), advanced tumour grade (both $P < 0.0001$), NPI (both $P < 0.0001$), ER- status (both $P < 0.0001$), PgR- status (both $P < 0.0001$), HER2+ status ($P = 0.016$ and $P = 0.002$), and TNBC (both $P < 0.0001$; Table 9.10).

A high percentage of tumour cells expressing PDL-1+ in the cytoplasm was significantly associated with younger age ($P = 0.003$) advanced tumour grade ($P = 0.009$), advanced tumour stage ($P < 0.0001$) and NPI ($P < 0.0001$; Table 9.11). Univariate analysis showed that no significant association existed between PDL-1 or PD1 expression and any patient outcome in BC ($P \geq 0.05$).

However, when the percentage of PD1+ TILs was combined with the percentage of AV-TILs, univariate analysis showed a high percentage of PD1+ TILs combined with a high percentage of AV-TILs was significantly associated with longer BCSS ($P = 0.029$ and $P = 0.024$) and DMFS ($P = 0.002$ and $P = 0.022$) in both the entire BC cohort and TNBC (Figure 9.9). Moreover, the combination of a high percentage of PDL1+ TILs and high percentage of AV-TILs was significantly associated with longer BCSS in only TNBC (Figure 9.10).

Table 9.10: Associations between percentage PD1+ TILs and percentage PDL-1+ TILs and the clinicopathological features of breast cancer.

Clinicopathological parameters	% PD1+ TILs			% PDL-1+ TILs		
	Low No (%)	High No (%)	x^2 *P*-value	Low No (%)	High No (%)	x^2 *P*-value
Patient age						
< 50 years	399 (50)	395 (50)	3.217	340 (54)	290 (46)	11.253
≥ 50 years	200 (45)	245 (55)	0.073	153 (43)	204 (57)	0.001
Tumour size						
≤ 2 cm	313 (52)	288 (48)	6.616	256 (54)	218 (46)	6.366
> 2 cm	281 (45)	347 (55)	0.010	232 (46)	273 (54)	0.012
Tumour Grade						
1	127 (68)	61 (32)	90.807	84 (58)	62 (42)	47.595
2	237 (60)	156 (40)	0<0.0001	203 (63)	120 (37)	0.0001
3	231 (36)	414 (64)		200 (39)	308 (61)	
Axillary nodal stage						
1	369 (50)	366 (50)	2.412	313 (54)	271 (46)	9.790
2	186 (47)	209 (53)	0.299	143 (46)	170 (54)	0.007
3	42 (43)	56 (57)		31 (38)	50 (62)	
Nottingham Prognostic Index						
Good	219 (64)	135 (64)	50.281	176 (64)	98 (36)	38.373
Moderate	299 (44)	376 (56)	<0.0001	250 (47)	281 (53)	0.0001
Poor	76 (36)	124 (36)		62 (36)	112 (64)	
Vascular invasion						
Negative	397 (49)	406 (51)	1.258	327 (52)	307 (48)	2.719
Positive	193 (46)	226 (54)	0.262	157 (46)	184 (54)	0.099
Oestrogen receptor						
Negative	101 (30)	233 (70)	60.474	80 (32)	174 (68)	44.224
Positive	489 (55)	397 (45)	<0.0001	402 (56)	319 (44)	0.0001
Progesterone receptor						
Negative	192 (38)	311 (62)	25.771	154 (40)	228 (60)	22.826
Positive	377 (56)	299 (44)	<0.0001	314 (56)	245 (44)	0.0001
HER2 status						
Negative	510 (50)	505 (50)	5.807	422 (52)	397 (48)	9.550
Positive	64(40)	96 (60)	0.016	45 (37)	78 (63)	0.002
Triple negative status						
Non-triple negative	519 (53)	451 (47)	46.539	420 (54)	361 (46)	30.648
Triple negative	64 (28)	162 (72)	<0.0001	54 (31)	122 (69)	0.0001

Table 9.11: Associations between the percentage tumour cells expressing PDL-1+ in the cytoplasm and the clinicopathological features of breast cancer.

Clinicopathological parameters	**% PDL-1+ tumour cells**		
	Low *N* (%)	High *N* (%)	x^2 *P*-value
Patient age			
< 50 years	145 (21)	531 (79)	9.097
≥ 50 years	53 (14)	328 (86)	0.003
Tumour size			
≤ 2 cm	99 (18)	444 (82)	0.304
> 2 cm	99 (20)	407 (80)	0.581
Tumour grade			
1	35 (22)	128 (78)	9.419
2	81 (23)	273 (77)	0.009
3	80 (15)	450 (85)	
Axillary nodal stage			
1	144 (23)	485 (77)	18.240
2	40 (12)	289 (89)	<0.0001
3	12 (13)	78 (87)	
Nottingham Prognostic Index			
Good Prognostic Group	79 (26)	163 (89)	17.700
Moderate Prognostic Group	99 (18)	461 (82)	<0.0001
Poor Prognostic Group	20 (11)	227 (74)	
Vascular invasion status			
Negative	136 (20)	540 (80)	1.709
Positive	62 (17)	307 (83)	0.191
Oestrogen receptor			
Negative	44 (17)	223 (84)	0.873
Positive	148 (19)	629 (81)	0.350
Progesterone receptor status			
Negative	75 (19)	327 (81)	0.036
Positive	110 (18)	495 (82)	0.849
HER2 status			
Negative	160 (18)	722(82)	0.144
Positive	25 (20)	103(80)	0.704
Triple negative status			
Non-triple negative	162 (19)	680 (81)	2.634
Triple negative	26 (14)	158 (86)	0.105

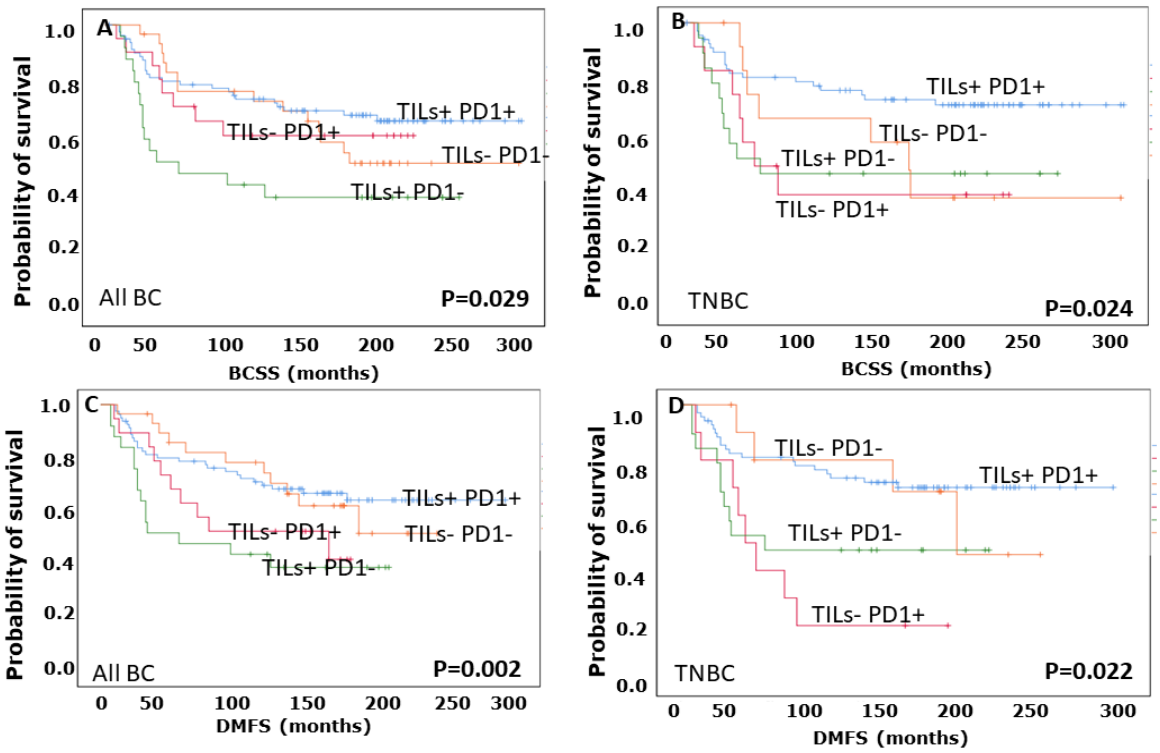


Figure 9.10: Kaplan-Meier survival curves of the associations between the combinations of percentage PD1+ TILs and percentage AV-TILs and breast cancer-specific survival (BCSS, A & B) and distant metastasis-free survival (DMFS, C & D) in the entire cohort (A & C) and TNBC (B & D). Percentages of cells were categorised as high (+) or low (-).

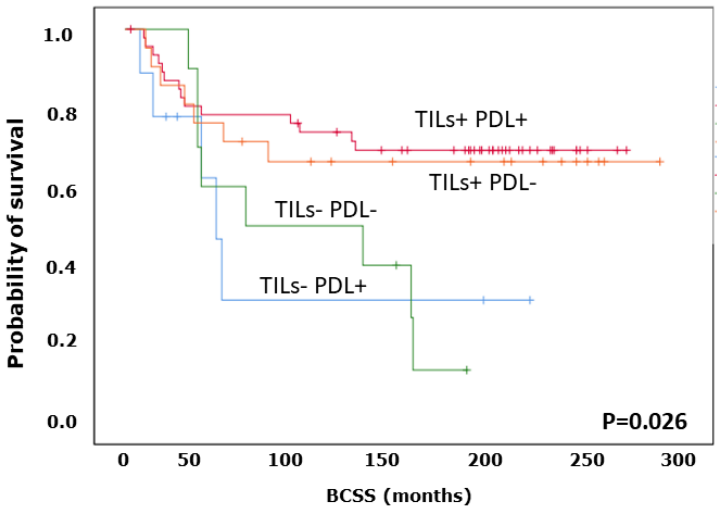


Figure 9.11: Kaplan Meier survival curves of the association between combinations of percentage PD1+ TILs and percentage AV-TILs and breast cancer-specific survival (BCSS) in TNBC. Percentages of cells were categorised as high (+) or low (-).

9.4 Discussion

There is growing interest in quantifying total TILs and various subpopulations of TILs in clinical practice due to their prognostic and possibly predictive value for the response to immunotherapies [545, 546]. In this study, the density of various infiltrating immune cells and their expression patterns were assessed in a large BC cohort with long-term clinical follow-up. The associations with patient outcome and tumour progression were also explored.

The International TILs Working Group stated that there is no obvious intertumour heterogeneity in the content of TILs in most tumours (i.e., between multiple tumour tissue sections) at the morphological level. However, no robust evidence was provided to support this observational statement [533]. To our knowledge, this is the first study to assess TIL heterogeneity in multiple tumour tissue blocks from the same primary tumours or to compare TIL heterogeneity in primary tumours and the corresponding recurrent tumorous.

For each sample in this study, the average percentage of TILs and percentage of TILs in hotspots were assessed to further characterise the distribution of TILs in different tumour areas. No significant differences in the percentages of AV-TILs or HS-TILs were observed between tumour blocks from the same cases, indicating that assessment of a single, full-face H&E-stained tissue section is sufficient to evaluate the density of TILs within a whole tumour.

However, when the analysis was limited to smaller tumour areas, i.e., different regions of the same slide (intra-slide heterogeneity), significant differences in the percentages of TILs were observed, particularly with regard to HS-TILs. This finding indicates the presence of substantial intra-slide TIL heterogeneity and therefore emphasises the variable nature of immune

infiltrating cells within different tumour cell populations. This observation is consistent with earlier reports, which found the CD8+, CD3+ and CD20+ TIL subpopulations in core biopsies are more heterogeneous across a single core than within multiple cores from the same tumour [547]. The variation in the percentages of CD3+, CD8+ and CD20+ TILs ranged from 30-33% between the core biopsies from the same tumour, while higher variation of 66-69% was observed between the same sections from the cores [548]. Another study reported that spatial heterogeneity of TIL subtypes could play a role in patient outcome. For instance, a high CD3+ T lymphocyte count in the invasive margin was associated with disease-free survival in BC, as high total and distant CD8+ T lymphocytes were both associated with better outcomes in BC [549].

In this study, both methods of assessing TILs, percentage AV-TILs and percentage HS-TILs, were significantly associated with poor prognostic features in the entire cohort, the possible explanation for this could be the high degree of genomic instability in high tumour grade triggers the tumour immunity [550]. In agreement with these findings, Kurozumi *et al.* (2019) reported that a high percentage of TILs was associated with ER- tumours and high grade tumours [551].

Interestingly, TILs were present at lower proportions and more homogenously distributed in recurrent tumours compared with the corresponding primary tumours. This is consistent with a previous report of lower percentages of TILs in recurrent metastatic TNBC and HER2+ lesions compared to the corresponding primary tumours [552]. Moreover, a high percentage of TILs was associated with significantly shorter post-recurrence survival, providing evidence that assessment of TILs in recurrent tumours could

represent a useful clinical variable to help predict post-recurrence outcome. Of note, the recurrent tumours assessed in this thesis included all subtypes of BC. Thus, the inverse association between TILs and patient outcomes in primary and recurrent cancer could be explained by the heterogeneity of BC, as well as the varied of roles of TILs in primary and recurrent tumours and within different subtypes of BC.

High AV TILs having associated with aggressive tumour feature in BC whereas in TNBC, high AV TILs have better BCSS, the possible explanation is the level of TILs present seems to be significantly associated with more aggressive features at diagnosis, and suggested that chemotherapy may service the release of several tumour associated neo-antigens, which follows cytotoxic-induced cell death that may suppress T reg cells, consequently inhibiting the antitumor response, and may restore cytotoxic T cells [553].

This study showed that varied TIL subsets are present in BC and that T cells, as defined by CD3+ cells, appear to dominate the TIL populations. This indicates TILs are a heterogeneous mixture of different immune cell types that are probably recruited to the vicinity of a tumour to perform specific functions [554, 555]. Additionally, this work showed that the numbers of CD3+, CD8+, FOXP3+, CD20+ and CD68+ immune cells were significantly associated with high percentages of AV-TILs and HS-TILs in TNBC, which is more aggressive than the other subtypes of BC. The percentage of average TILs in H&E-stained sections has been significantly associated with the numbers of CD20+ and CD45+ TILs in BC [556]. In this study, high CD8+ and CD3+ cellular infiltrates were associated with a good prognosis and had independent predictive value in TNBC, regardless of other clinicopathological parameters. These findings are concordant with other studies [234, 530, 557]. High CD20+ lymphocytes within

tumours was associated with better BCSS and DMFS in the TNBC and HER2+ subtypes in this study. This may indicate TNBC and HER2+ BC have a higher immunogenicity than Luminal BC due to higher production of antibodies against the tumour cells by B cells, in agreement with Buisseret and colleagues [89]. Several subsets of TILs have been described in TNBC and their clinical relevance has been explored [558], with the aim of identifying subtypes of TNBC that may have a better outcome or treatment responses. However, no significant association was observed between FOXP3+ cells and patient outcome in this study. In contrast, Yeong *et al.* (2017) reported that a high density of intratumoral FOXP3+ cells, but not stromal infiltrating cells, was associated with longer overall survival [559].

In contrast to TNBC, a high number of CD68+ macrophages was associated with a significantly poorer outcome in the Luminal B subtype, which may reflect, at least in part, the distinct roles of varied TIL subsets in different subtypes of BC [560]. These results also underscore the distinct tumour-stromal interactions in TNBC compared to non-TNBC tumours, as evidenced by their distinct inflammatory microenvironments. Tumour-associated macrophages differentiate into active classic (M1) and alternative (M2) macrophages. Alternative macrophages (M2) have been associated with cell growth, angiogenesis, migration and invasion in a range of cancers [561]. Taking together, TILs was behaved differently between BC subtypes, which may suggest the immune cells interact differently with different tumour types, and intra-tumour heterogeneity may affect the expression patterns of immune cells, thereby underlying the difference in the prognostic significance of TIL expression between ER-positive and ER-negative BC [551, 562]. In addition,

the immune microenvironment contribute to modulating the response to types of treatment [563].

The significant association between the percentage of AV-TILs and the various lymphocyte subsets, as well as the independent associations between AV-TILs and outcomes, observed in this study indicate assessing AV-TILs using H&E-stained sections can accurately represent the density of immune cells and has prognostic value in BC. Moreover, CD3 and CD8 lost their prognostic effect when AV-TILs were included in the same multivariate model. This provides evidence of the prognostic value of assessing AV-TILs in H&E-stained sections, especially in TNBC. In the current study, PDL-1+ cells (both in TILs and tumour cells) and PD1+ TILs were present at high levels in TNBC and correlated with more aggressive tumour features, in agreement with previous studies [528, 564, 565]. Moreover, the varied expression of immune checkpoint proteins such as PDL-1 suggests immunological infiltrates exert varied biological effects in different subtypes of BC.

Moreover, high PDL-1+ TILs combined with high AV-TILs+ and high PD-1+ TILs combined with high AV-TILs were associated with better patient outcomes, especially in TNBC, which suggests these TIL parameters are associated with a better response to therapy. A previous study reported that PDL-1+ TILs correlated with a high percentage of TILs, a high percentage of CD8+ TILs, ER- tumours and high tumour grade. PDL-1 expression was also reported as a predictive biomarker for the response to NAC with trastuzumab in HER2+ BC [566]. PDL1 expression on cancer cell membrane co-relates with CD8+ T cell infiltration may suggest, PD-L1/PD-1 blockade allows cytotoxic T cells to recognise and destroy tumour cells[567].

The current study also suggests clinical evaluation of TILs based on H&E staining and IHC analysis of PDL-1 and PD1 expression can be used to predict patient outcome and prognosis in TNBC.

9.5 Conclusion

TILs vary significantly between TNBC and non-TNBC, and TNBC contains a distinct subpopulation of immune cells. These differences may help to explain the varied prognostic utility of TILs and specific immune cell subsets in BC. Although the International TILs Working Group recognise heterogeneity is an issue, no strong guidance on this issue was provided. This study confirmed that assessment of average TILs in a single representative tumour block reliably reflects TILs across a whole tumour. This work demonstrates that assessment of the average percentages of TILs in H&E-stained slides using the TILs International Working Group Guidelines not only provides reproducible and independent prognostic information in TNBC, but also has equivalent prognostic value as analysis of immunohistochemically defined subpopulations of immune cells. Moreover, the temporal heterogeneity in TILs between primary tumours and recurrent tumours warrants further investigation. Furthermore, analysis of the percentages of PD-1+/PDL-1+ immune cells may play a vital role in identification of patients who may achieve a better outcome or require more intensive treatment.

Chapter 10

General Discussion

10.1 General Discussion

Breast cancer (BC) is a heterogenous disease and tumours are structurally complex systems [568]. Heterogeneity is observed in BC in terms of morphology, gene mutations, protein expression profiles, metabolic profiles, and various other components of the tumour microenvironment [124]. Intertumour heterogeneity (i.e., between tumours) also exists in BC, and patients are traditionally classified into homogenous subgroups based on several factors, including tumour size, tumour grade, tumour stage, histological type and molecular biomarkers [569].

BC is classically divided into several subtypes based on three molecular biomarkers: ER, PgR and HER2. These biomarkers play significant clinical roles in treatment decision-making for patients with BC [570]. These molecular classifications have been successfully employed to guide the administration of individualised therapies and improve BC-specific survival. However, a number of other molecular classifications actually exist. The clinical features and optimal treatments for these subtypes of BC remain to be clarified, and many of these subtypes fail to respond to therapy. For example, around 40% of patients with BC develop local and distant recurrence, which indicates systemic therapy fails in many subtypes of BC [571].

Intratumoural BC heterogeneity (i.e., within the same tumour) has been associated with aggressiveness and treatment failure [572]. Intratumor

heterogeneity can be viewed as cell populations with non-uniform phenotypes and variable expression of markers by different populations of cells within a tumour. These heterogeneity are exhibit distinct responses to treatment [573]. Phenotypic intratumour heterogeneity in BC develops due to accumulation of genetic mutations and/or epigenetic modifications [574], which can be explained by either the clonal evolution concept, the cancer stem cell concept, or both.

The clonal evolution concept suggests that neoplasms initially arise from a single cell, and then cancer cells with genetic instability acquire additional genetic alterations during tumour progression, which leads to evolution of more aggressive subclones within the tumour [204, 575]. The cancer stem cell concept proposes that a small population of tumorigenic cells with stem-like properties are responsible for tumour growth, disease progression, generation of intratumoral heterogeneity and resistance to therapy [576, 577]. Furthermore, variations in microenvironment components and the interactions of tumour cells with fibroblasts, stromal cells, infiltrating immune cells or the extracellular matrix can also lead to tumour variability [578].

Therefore, further refining the classical system of BC evaluation and classification by incorporating additional clinical and molecular parameters that reflect tumour heterogeneity may enable more accurate prognostication and guide more informed therapeutic decision making.

This thesis assessed heterogeneity at several aspect in BC using well-characterised patient cohorts. A variety of histopathologic analyses were employed to assess cellular and molecular heterogeneity that may potentially impact patient outcomes in BC. Molecular heterogeneity in terms of expression of BC stem cell (BCSC) markers, oestrogen receptor alpha (ER) and ER-related

biomarkers, and the density of tumour-infiltrating lymphocytes (TILs) exists within individual tumours and also between different tumours were assessed, and their association with patient outcomes in BC was examined.

Substantial evidence indicates that BC is hierarchically organised of heterogonous populations of tumour cells. BCSCs, the cells at the apex of this hierarchy, can self-renew into more differentiated cancer cells that contribute to tumour heterogeneity [579]. BCSCs contribute to tumorigenesis, metastasis and therapeutic resistance [391, 580-583]. Several markers can be used to identify BCSCs [147, 149, 584-586]; however, there is no universal marker for BCSCs. This thesis shows the prognostic value of the BCSC markers ALDH1A1 and BMI-1 vary in different molecular subtypes of invasive BC.

ALDH1A1 mRNA and protein expression were examined in two large BC cohorts. High ALDH1A1 protein expression was associated with advanced tumour features and poor outcome in the highly aggressive TNBC and Luminal B subtypes, which indicates ALDH1A1 may potentially represent a prognostic marker for these subtypes. In agreement these findings, Yao et al. (2017) reported ALDH1A1 was associated with a poor outcome, but only in TNBC. Additionally, Lizet et al. (2018) found high ALDH1A1 mRNA and protein expression were both associated with poorly differentiated histology and shorter overall survival in colorectal cancer [275]. Furthermore, expression of ALDH1A1 significantly correlated with the expression of other BCSC markers, suggesting ALDH1A1 could also be used to identify and isolate the stem cell subpopulation in BC tissues. Similarly, in vitro studies suggested that ALDH1A1 is associated with cancer stem cells and plays a role in self-renewal, differentiation and self-protection in various solid tumours [149, 275, 587, 588]. Overall, ALDH1A1—as assessed using IHC—appears to

represent a BCSC marker, and also has value as a prognostic marker in the Luminal B and TNBC subtypes.

The prognostic significance of BMI-1 expression also varied between BC subtypes. High BMI-1 expression was associated with poor BCSS in TNBC. However, high BMI-1 expression was prognostic for good BCSS and negatively associated with BCSCs in ER+ tumours. In agreement, Choi et al. (2008) reported high BMI-1 expression was associated with less advanced tumour features and low ALDH1A1 expression in ER+ tumours [348]. In contrast, Wang and colleagues (2012) suggested that high BMI-1 expression was associated with the basal-like phenotype of BC [346]. Study by Yu et al. (2011) in Head and Neck Squamous Cancer (HNSCC) was showed BMI-1 can regulate Snail and ALDH1; change the EMT-related genotypes of the $ALDH1^-$ cells; and modulate distant lung metastases, and overexpression of Bmi-1 could turn HNSCC-$ALDH1^-$ into cancer stem cell-like HNSCC-$ALDH1^+$ cells [589].

Overall, the data indicates high BMI-1 expression is associated with varied clinicopathological variables and outcomes in different molecular subtypes of BC. BMI-1 was associated with a good prognosis in ER+ tumours and a poor prognosis in TNBC, which suggests that expression of the ER modulates the biological function of BMI-1 in BC.

BCSC biomarkers are not exclusive to BCSCs and are also frequently expressed in normal and tumour cells [579], and have variable prognostic value in BC—as observed for ALDH1A1 and BMI-1 in this thesis. Thus, detection and isolation of BCSCs in tissue samples remains challenging. Co-localisation of expression of the BCSC markers ALDH1A1, SOX9 and EpCAM was observed in a

small population of tumour cells, and this cell population was associated with tumour progression, especially metastasis and recurrence in ER- and HER2+ tumours. These results indicate that BCSC phenotypes can be detected by assessing the co-localisation of multiple BCSC markers. De Beça et al. (2014) also indicated that the expression and distribution of BCSC markers varied between histological types, and demonstrated a higher percentage of cells with the $CD44^{high}/CD24^{low}/ALDH1A1^{high}$ phenotype in the no-special histological type of BC [590]. In contrast, Liu et al. (2014) demonstrated the lack of co-expression of the BCSC markers CD24, CD44, ALDH and SOX2 in vitro and in vivo. However, this thesis demonstrated no correlations between CD24, CD44 and ALDH1 expression, which may indicate these molecules are not good candidate stem cell markers for BC tissues. Moreover, Liu et al. (2014) only assessed a very small number of cases (n = 33) and detected a low percentage (1-10%) of tumour cells expressing CD24, CD44, ALDH1A1 and SOX2. Analysis of a larger cohort is required to confirm if BC cells co-express these markers [591], as the ALDH1A1+ SOX9+ EpCAM+ BCSC phenotype was only detected in 175 of 1550 cases in this thesis.

This study confirms that BCSC phenotypes can be evaluated in tissue samples using double IHC staining. Furthermore, BCSC phenotypes have prognostic value for tumour progression, especially the metastasis and recurrence of aggressive tumour subtypes. Overall, assessment of BCSCs could provide valuable clinical information to refine the treatment of BC.

In the second concept, ER expression was investigated. Assessment of ER expression is a fundamental component of pathological assessment and treatment decision making in BC. The ASCO/CAP recently recommended that ER

expression status should be considered positive if 1% or more of the tumour cells exhibit positive nuclear IHC staining [430]. Recently, our group confirmed the concept of bimodal ER expression in BC, with the majority of tumours exhibiting either ER-negative (< 1% ER+ cells) or ER positive (ER+ > 10% ER+ cells) cells [442]. This raises two questions. Firstly, are tumours with low ER+ expression (1-9% ER+ cells) truly ER+ or do they represent a misclassification of ER- tumours? Secondly, do cells expressing varied levels of the ER within the same tumour behave similarly or differently in the presence of ER- cells?

To address the first question, needle core biopsy (NCBs) and the corresponding surgical excision specimens (SESs) of tumours with low ER+ expression (1-9% ER+ cells) were re-evaluated using IHC and RNA in situ hybridisation. Overall, 65% of low ER+ (1-9%) NCBs were reclassified as ER-negative (0 or < 1% ER+ cells) based on the SESs; these cases were noted as false positives. Furthermore, the IHC results for the reclassified cases were validated by RNA in situ hybridisation. Rakha et al. (2012) reported 47% of low ER+ tumours were ER-negative (0 or < 1%) in SESs, and also observed a high rate of false positive and nonspecific ER immunostaining in low ER+ tumours [444].

Overall, this thesis indicates that tumours expressing low proportions of ER+ cells (1-9%) can exhibit false positive results in IHC analysis, thus NCBs with 1-9% ER+ cells that are PgR- and high-grade should be reassessed, and clinical decision making should be based on analysis of SESs.

To investigate whether intratumoral ER heterogeneity, including the existence of ER+ and ER- subclones within the same tumour, has a biological and

clinical impact on tumour behaviour and the response to therapy in BC, a comprehensive assessment of ER heterogeneity in primary tumours and the paired lymph node metastases was performed using a variety of approaches. This work confirmed that intratumoural heterogeneity in ER expression exists between different regions of the same tumour, and that variable patterns of ER distribution are observed.

Low ER heterogeneity scores were associated with more aggressive tumour features and a shorter disease-free interval, while a high heterogeneity score and the diffuse distribution pattern was associated with less aggressive features and a longer disease-free interval. Lindstrom et al. (2018) reported that high ER heterogeneity was associated with poorer survival in BC. However, their study was based on tumours containing 20-100% ER+ cells, while this study included tumours with 1-70% ER+ cells. Their heterogeneity index included tumours with 100% ER+ cells, which are likely to be less heterogeneous and are associated with good patient outcome [465]. Intratumoral ER heterogeneity with the existence of ER-negative clones has been associated with resistance to endocrine therapy, which suggests that endocrine treatment does not kill all tumour cells, and the bulk of the tumour survives and leads to relapse [456].

Moreover, this thesis indicates the pattern of ER distribution in the primary tumour can influence the behaviour of tumour metastases, which suggests ER+ tumours are not a single biological system. Taken together, this study demonstrates that intratumoural ER heterogeneity influences tumour progression, patient outcome and resistance to endocrine therapy. Thus, intratumoural ER heterogeneity should be taken into account when assessing clinical BC samples.

Furthermore, analysis of the variations in gene expression in ER+ BC can be used to further refine prognostic models and predict the response to adjuvant treatment[592]. This thesis demonstrates SLC39A6 is highly expressed in ER+ tumours. Furthermore, high SLC39A6 expression was associated with less advanced tumour features and good outcomes in ER+ tumours. Similarly, in vitro and ex vivo studies showed SLC39A6 is highly expressed in ER+ tumours and associated with less aggressive tumour features in ER+ BC [498, 500, 593]. However, other studies suggested that SLC39A6 was associated with disease progression and a poorer outcome in BC [505, 594]. Multiple regulatory pathways have been suggested to regulate SLC39A6, thus there may be an interaction between the ER and SLC39A6 [494]. High SLC39A6 expression also may signify a positive response to endocrine therapy. This finding is in agreement with another study, which showed 78% of tumours that failed to respond to endocrine therapy were SLC39A6-negative, whereas 61% of tumours that responded to treatment were SLC39A6-positive [498]. Overall, SLC39A6 may have prognostic value in BC, especially in ER+ tumours. Assessment of SLC39A6 could help to categorise and predict patients with ER+ BC who may achieve a good outcome.

The third concept assessed the spatial and temporal heterogeneity of tumour-infiltrating lymphocytes (TILs) in BC and explored their impact on tumour behaviour and patient outcome.

The International TILs Working Group reported no obvious variation in tumour-infiltrating lymphocytes (TILs) between tumour blocks from the same patient; however, no evidence was provided to support this statement [595]. This thesis confirmed that evaluation of a single tumour section is sufficient to

determine TILs in the whole tumour, and confirmed that spatial heterogeneity in the percentages of TILs can be observed within the same tumour section and temporal TIL heterogeneity exists between primary and recurrent tumours.

Moreover, high average TILs was associated with significantly longer BCSS in TNBC and shorter post-recurrence survival in recurrent tumours. Kurozmi et al. (2019) reported that the biological functions of overall TILs may vary in ER-positive and ER-negative primary BC, as high TILs were prognostic for a poor outcome in ER-positive tumours and a better outcome in ER-negative BC [551].

Collectively, despite the heterogeneity of immune cell type components present in BC, this thesis demonstrates that the average percentage of TILs in one full-face H&E-stained section reliably represents whole tumour TILs. Furthermore, the average percentage of TILs in the primary tumour is associated with outcome, particularly in TNBC, and holds prognostic significance in recurrent tumours.

10.2 Conclusions

Overall, this thesis characterised the impact of intertumour and intratumoural heterogeneity on the clinicopathological features and patient outcomes in invasive BC. The morphological and molecular analyses employed in this work provide detailed insight into the characteristics and prognostic and predictive value of heterogeneity in BC. Evaluation of BC heterogeneity in terms of the expression of BCSC markers and BCSC phenotypes, expression of the ER and TILs density may help improve clinical decision-making by refining prognostic predictions and helping to identify patients who may respond well to specific therapies. However, these results need to be validated before these strategies could be adopted in the clinic.

10.3 Limitations

The IHC protein expression patterns observed in this study were homogenous, and the use of TMA for large-scale analysis has been validated by various studies [127, 596]. However, analysis of at least two or three TMA cores per case may provide a more reliable assessment of protein expression in BC. Moreover, the IHC H-scoring method is a semi-quantitative assessment that can be influenced by the subjective judgment of the scorers; however, double blind scoring was used in this study.

Furthermore, the cohorts assessed in this thesis were from a single centre, thus there is a possibility of inaccurate results related to pre-analytical and analytical issues. The small numbers of patients in the ER cohort, recurrence cohort and LN metastasis subgroup also limited the ability to characterise heterogeneity.

10.4 Future recommendations

Further analysis of larger cohorts is required to validate the prognostic value of BMI-1 in TNBC, and functional studies are recommended to investigate the exact role of BMI-1 in different subtypes of BC.

Further validation studies are also recommended to confirm the co-localisation of ALDH1A1, SOX9 and EpCAM. Evaluation of full face sections is strongly recommended to assess the co-localisation of BCSC markers and study the distribution of these markers between different regions of tumours. Moreover, further analysis of the expression of BCSC markers in LN metastases and determination of the temporal heterogeneity between primary tumours and the corresponding recurrent tumours and metastases will provide valuable information on the heterogeneity of BCSCs in BC and their involvement in tumour invasion and metastasis. Multiplex staining is recommended to identify BCSCs in tissue samples.

Studies of larger cohorts and multi blocks are also recommended to confirm intratumoural heterogeneous ER expression. Furthermore, application of artificial intelligence to assess the heterogeneity of ER expression at the cellular level and investigate the interactions between tumour cells with varied ER expression levels could provide information on the biological impact of ER expression on tumour progression and patient outcome [597]. Moreover, analysis of validation cohorts from other laboratories is necessary to eliminate pre-analytical and analytical issues and confirm the results of this study. Functional studies are also necessary to clarify the role of SLC39A6 in ER+ BC. In vitro study in different breast cancer cell lines including ER+ (treated with and without endocrine therapy) and ER- cell lines should be taken in the consideration in the future study.

Finally, assessment of larger cohorts is required to further explore the clinical impact of TIL heterogeneity on other subtypes of BC, including ER+ and HER2+ tumours. Additionally, validation studies are necessary to confirm the clinical impact of TILs in recurrent tumours. Furthermore, application of artificial intelligence to quantify TIL heterogeneity is highly recommended to characterize the heterogeneous patterns of distribution of TILs and explore the precise interactions between various tumour cells and immune cells [598].

10.5 Ethical approval

This study was approved by the Nottingham Research Ethics Committee 2 under the title "Development of a molecular genetics classification of breast cancer" under the ethical approval number (RC2020313) and in compliance with the current ethical and legal guidelines of the United Kingdom. All basic and operational ethical principles for this study were conducted in accordance with the 1975 Helsinki Declaration, and its later amendments or comparable ethical standards. The identification and health-sensitive information for all patients in this study were anonymous.

References

1. Birnbaum D, Bertucci F, Ginestier C et al. Basal and luminal breast cancers: basic or luminous? (review). Int J Oncol 2004; 25: 249-258.
2. Cardiff RD. Are the TDLU of the human the same as the LA of mice? J Mammary Gland Biol Neoplasia 1998; 3: 3-5.
3. Javed A, Lteif A. Development of the human breast. Seminars in plastic surgery 2013; 27: 5-12.
4. Forsyth IA. The mammary gland. Baillieres Clin Endocrinol Metab 1991; 5: 809-832.
5. Hovey RC, Salomon D. The Journal of Mammary Gland Biology and Neoplasia into the future - the potential of plasticity and pluripotency. J Mammary Gland Biol Neoplasia 2015; 20: 1-3.
6. Feng Y, Spezia M, Huang S et al. Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. Genes & diseases 2018; 5: 77-106.
7. Bray F, Ferlay J, Soerjomataram I et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68: 394-424.
8. Bray F, Soerjomataram I. The Changing Global Burden of Cancer: Transitions in Human Development and Implications for Cancer Prevention and Control. In Gelband H, Jha P, Sankaranarayanan R, Horton S (eds): Cancer: Disease Control Priorities, Third Edition (Volume 3). Washington (DC): The International Bank for Reconstruction and Development / The World Bank

© 2015 International Bank for Reconstruction and Development / The World Bank. 2015.

9. Hortobagyi GN, de la Garza Salazar J, Pritchard K et al. The global breast cancer burden: variations in epidemiology and survival. Clin Breast Cancer 2005; 6: 391-401.
10. Youlden DR, Cramb SM, Dunn NA et al. The descriptive epidemiology of female breast cancer: an international comparison of screening, incidence, survival and mortality. Cancer Epidemiol 2012; 36: 237-248.
11. Jemal A, Bray F, Center MM et al. Global cancer statistics. CA Cancer J Clin 2011; 61: 69-90.
12. Ferlay J, Shin HR, Bray F et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010; 127: 2893-2917.
13. Saggu S, Rehman H, Abbas ZK, Ansari AA. Recent incidence and descriptive epidemiological survey of breast cancer in Saudi Arabia. Saudi medical journal 2015; 36: 1176-1180.
14. Assi HA, Khoury KE, Dbouk H et al. Epidemiology and prognosis of breast cancer in young women. J Thorac Dis 2013; 5 Suppl 1: S2-8.
15. Venkitaraman AR. Cancer suppression by the chromosome custodians, BRCA1 and BRCA2. Science 2014; 343: 1470-1475.
16. Welcsh PL, King MC. BRCA1 and BRCA2 and the genetics of breast and ovarian cancer. Hum Mol Genet 2001; 10: 705-713.
17. Petrucelli N, Daly MB, Feldman GL. Hereditary breast and ovarian cancer due to mutations in BRCA1 and BRCA2. Genet Med 2010; 12: 245-259.
18. Antoniou A, Pharoah PD, Narod S et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. Am J Hum Genet 2003; 72: 1117-1130.
19. Hulka BS, Moorman PG. Breast cancer: hormones and other risk factors. Maturitas 2008; 61: 203-213; discussion 213.
20. Kuchenbaecker KB, Hopper JL, Barnes DR et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. Jama 2017; 317: 2402-2416.

21. Mehrgou A, Akouchekian M. The importance of BRCA1 and BRCA2 genes mutations in breast cancer development. Medical journal of the Islamic Republic of Iran 2016; 30: 369-369.
22. Duffy SW, Morrish OWE, Allgood PC et al. Mammographic density and breast cancer risk in breast screening assessment cases and women with a family history of breast cancer. European journal of cancer (Oxford, England : 1990) 2018; 88: 48-56.
23. Baglietto L, Krishnan K, Stone J et al. Associations of mammographic dense and nondense areas and body mass index with risk of breast cancer. Am J Epidemiol 2014; 179: 475-483.
24. Chen CL, Weiss NS, Newcomb P et al. Hormone replacement therapy in relation to breast cancer. Jama 2002; 287: 734-741.
25. Beral V. Breast cancer and hormone-replacement therapy in the Million Women Study. Lancet 2003; 362: 419-427.
26. Zhou Y, Chen J, Li Q et al. Association between breastfeeding and breast cancer risk: evidence from a meta-analysis. Breastfeed Med 2015; 10: 175-182.
27. Hamajima N, Hirose K, Tajima K et al. Alcohol, tobacco and breast cancer--collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. Br J Cancer 2002; 87: 1234-1245.
28. Jung S, Wang M, Anderson K et al. Alcohol consumption and breast cancer risk by estrogen receptor status: in a pooled analysis of 20 studies. Int J Epidemiol 2016; 45: 916-928.
29. Makarem N, Chandran U, Bandera EV, Parekh N. Dietary fat in breast cancer survival. Annu Rev Nutr 2013; 33: 319-348.
30. Jones ME, Schoemaker MJ, Wright LB et al. Smoking and risk of breast cancer in the Generations Study cohort. Breast Cancer Res 2017; 19: 118.
31. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 1991; 19: 403-410.
32. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. Cancer 1989; 63: 181-187.
33. Narod SA. Tumour size predicts long-term survival among women with lymph node-positive breast cancer. Curr Oncol 2012; 19: 249-253.
34. Rakha EA, El-Sayed ME, Lee AH et al. Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. J Clin Oncol 2008; 26: 3153-3158.
35. Rakha EA, Reis-Filho JS, Baehner F et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. Breast Cancer Research 2010; 12: 207.
36. Schwartz AM, Henson DE, Chen D, Rajamarthandan S. Histologic grade remains a prognostic factor for breast cancer regardless of the number of positive lymph nodes and tumor size: a study of 161 708 cases of breast cancer from the SEER Program. Arch Pathol Lab Med 2014; 138: 1048-1052.
37. Michaelson JS, Silverstein M, Sgroi D et al. The effect of tumor size and lymph node status on breast carcinoma lethality. Cancer 2003; 98: 2133-2143.
38. Oluogun WA, Adedokun KA, Oyenike MA, Adeyeba OA. Histological classification, grading, staging, and prognostic indexing of female breast cancer in an African population: A 10-year retrospective study. International journal of health sciences 2019; 13: 3-9.
39. Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham prognostic index in primary breast cancer. Breast Cancer Research and Treatment 1992; 22: 207-219.
40. Fong Y, Evans J, Brook D et al. The Nottingham Prognostic Index: five- and ten-year data for all-cause survival within a screened population. Annals of the Royal College of Surgeons of England 2015; 97: 137-139.
41. Lee AH, Ellis IO. The Nottingham prognostic index for invasive carcinoma of the breast. Pathol Oncol Res 2008; 14: 113-115.
42. Blamey RW, Ellis IO, Pinder SE et al. Survival of invasive breast cancer according to the Nottingham Prognostic Index in cases diagnosed in 1990-1999. Eur J Cancer 2007; 43: 1548-1555.

43. He KW, Sun JJ, Liu ZB et al. Prognostic significance of lymphatic vessel invasion diagnosed by D2-40 in Chinese invasive breast cancers. Medicine (Baltimore) 2017; 96: e8490.
44. Ryu YJ, Kang SJ, Cho JS et al. Lymphovascular invasion can be better than pathologic complete response to predict prognosis in breast cancer treated with neoadjuvant chemotherapy. Medicine 2018; 97: e11647-e11647.
45. Osborne CK, Yochmowitz MG, Knight WA, 3rd, McGuire WL. The value of estrogen and progesterone receptors in the treatment of breast cancer. Cancer 1980; 46: 2884-2888.
46. Dowsett M, Houghton J, Iden C et al. Benefit from adjuvant tamoxifen therapy in primary breast cancer patients according oestrogen receptor, progesterone receptor, EGF receptor and HER2 status. Ann Oncol 2006; 17: 818-826.
47. Samaan NA, Buzdar AU, Aldinger KA et al. Estrogen receptor: a prognostic factor in breast cancer. Cancer 1981; 47: 554-560.
48. Gapinski PV, Donegan WL. Estrogen receptors and breast cancer: prognostic and therapeutic implications. Surgery 1980; 88: 386-393.
49. Torres-López L, Maycotte P, Liñán-Rico A et al. Tamoxifen induces toxicity, causes autophagy, and partially reverses dexamethasone resistance in Jurkat T cells. J Leukoc Biol 2019; 105: 983-998.
50. Williams C, Lin CY. Oestrogen receptors in breast cancer: basic mechanisms and clinical implications. Ecancermedicalscience 2013; 7: 370.
51. Li S, Han B, Liu G et al. Immunocytochemical localization of sex steroid hormone receptors in normal human mammary gland. J Histochem Cytochem 2010; 58: 509-515.
52. Palmieri C, Saji S, Sakaguchi H et al. The expression of oestrogen receptor (ER)-beta and its variants, but not ERalpha, in adult human mammary fibroblasts. J Mol Endocrinol 2004; 33: 35-50.
53. Manavathi B, Dey O, Gajulapalli VNR et al. Derailed estrogen signaling and breast cancer: an authentic couple. Endocrine reviews 2013; 34: 1-32.
54. Roger P, Sahla ME, Mäkelä S et al. Decreased expression of estrogen receptor beta protein in proliferative preinvasive mammary tumors. Cancer Res 2001; 61: 2537-2541.
55. Zhao C, Dahlman-Wright K, Gustafsson JA. Estrogen receptor beta: an overview and update. Nucl Recept Signal 2008; 6: e003.
56. Huang B, Warner M, Gustafsson J. Estrogen receptors in breast carcinogenesis and endocrine therapy. Mol Cell Endocrinol 2015; 418 Pt 3: 240-244.
57. Williams C, Lin C-Y. Oestrogen receptors in breast cancer: basic mechanisms and clinical implications. Ecancermedicalscience 2013; 7: 370-370.
58. Duffy MJ. Estrogen receptors: role in breast cancer. Crit Rev Clin Lab Sci 2006; 43: 325-347.
59. Siersbæk R, Kumar S, Carroll JS. Signaling pathways and steroid receptors modulating estrogen receptor α function in breast cancer. Genes Dev 2018; 32: 1141-1154.
60. Green S, Kumar V, Krust A et al. Structural and functional domains of the estrogen receptor. Cold Spring Harb Symp Quant Biol 1986; 51 Pt 2: 751-758.
61. Kach J, Conzen SD, Szmulewitz RZ. Targeting the glucocorticoid receptor in breast and prostate cancers. Sci Transl Med 2015; 7: 305ps319.
62. Bai Z, Gust R. Breast cancer, estrogen receptor and ligands. Arch Pharm (Weinheim) 2009; 342: 133-149.
63. Segovia-Mendoza M, Morales-Montor J. Immune Tumor Microenvironment in Breast Cancer and the Participation of Estrogen and Its Receptors in Cancer Physiopathology. Front Immunol 2019; 10: 348.
64. Yu L, Wang L, Mao C et al. Estrogen-independent Myc overexpression confers endocrine therapy resistance on breast cancer cells expressing ERαY537S and ERαD538G mutations. Cancer letters 2019; 442: 373-382.
65. Rondón-Lagos M, Villegas VE, Rangel N et al. Tamoxifen Resistance: Emerging Molecular Targets. Int J Mol Sci 2016; 17.

66. Yarden Y. The EGFR family and its ligands in human cancer. signalling mechanisms and therapeutic opportunities. Eur J Cancer 2001; 37 Suppl 4: S3-8.
67. Lee RJ, Albanese C, Fu M et al. Cyclin D1 is required for transformation by activated Neu and is induced through an E2F-dependent signaling pathway. Mol Cell Biol 2000; 20: 672-683.
68. Yen L, You XL, Al Moustafa AE et al. Heregulin selectively upregulates vascular endothelial growth factor secretion in cancer cells and stimulates angiogenesis. Oncogene 2000; 19: 3460-3469.
69. Tong ZJ, Shi NY, Zhang ZJ et al. Expression and prognostic value of HER-2/neu in primary breast cancer with sentinel lymph node metastasis. Biosci Rep 2017; 37.
70. Cianfrocca M, Goldstein LJ. Prognostic and predictive factors in early-stage breast cancer. Oncologist 2004; 9: 606-616.
71. Perou CM, Sørlie T, Eisen MB et al. Molecular portraits of human breast tumours. Nature 2000; 406: 747-752.
72. Yersal O, Barutca S. Biological subtypes of breast cancer: Prognostic and therapeutic implications. World J Clin Oncol 2014; 5: 412-424.
73. Green AR, Powe DG, Rakha EA et al. Identification of key clinical phenotypes of breast cancer using a reduced panel of protein biomarkers. British journal of cancer 2013; 109: 1886-1894.
74. Brenton JD, Carey LA, Ahmed AA, Caldas C. Molecular classification and molecular forecasting of breast cancer: ready for clinical application? J Clin Oncol 2005; 23: 7350-7360.
75. Curtis C, Shah SP, Chin SF et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012; 486: 346-352.
76. Cleator S, Heller W, Coombes RC. Triple-negative breast cancer: therapeutic options. Lancet Oncol 2007; 8: 235-244.
77. Prat A, Adamo B, Cheang MC et al. Molecular characterization of basal-like and non-basal-like triple-negative breast cancer. Oncologist 2013; 18: 123-133.
78. Cristofanilli M, Gonzalez-Angulo A, Sneige N et al. Invasive lobular carcinoma classic type: response to primary chemotherapy and survival outcomes. J Clin Oncol 2005; 23: 41-48.
79. Bollet MA, Savignoni A, Pierga JY et al. High rates of breast conservation for large ductal and lobular invasive carcinomas combining multimodality strategies. British journal of cancer 2008; 98: 734-741.
80. Song Y, Liu X, Zhang G et al. Unique clinicopathological features of metaplastic breast carcinoma compared with invasive ductal carcinoma and poor prognostic indicators. World J Surg Oncol 2013; 11: 129.
81. Makki J. Diversity of Breast Carcinoma: Histological Subtypes and Clinical Relevance. Clinical medicine insights. Pathology 2015; 8: 23-31.
82. Polyak K. Heterogeneity in breast cancer. The Journal of clinical investigation 2011; 121: 3786-3788.
83. Junttila MR, de Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response. Nature 2013; 501: 346-354.
84. Louault K, Bonneaud TL, Séveno C et al. Interactions between cancer-associated fibroblasts and tumor cells promote MCL-1 dependency in estrogen receptor-positive breast cancers. Oncogene 2019; 38: 3261-3273.
85. Dvorak KM, Pettee KM, Rubinic-Minotti K et al. Carcinoma associated fibroblasts (CAFs) promote breast cancer motility by suppressing mammalian Diaphanous-related formin-2 (mDia2). PLoS One 2018; 13: e0195278.
86. Liu T, Han C, Wang S et al. Cancer-associated fibroblasts: an emerging target of anti-cancer immunotherapy. Journal of Hematology & Oncology 2019; 12: 86.
87. Aboussekhra A. Role of cancer-associated fibroblasts in breast cancer development and prognosis. Int J Dev Biol 2011; 55: 841-849.
88. Natrajan R, Sailem H, Mardakheh FK et al. Microenvironmental Heterogeneity Parallels Breast Cancer Progression: A Histology-Genomic Integration Analysis. PLoS Med 2016; 13: e1001961.

89. García-Teijido P, Cabal ML, Fernández IP, Pérez YF. Tumor-Infiltrating Lymphocytes in Triple Negative Breast Cancer: The Future of Immune Targeting. Clin Med Insights Oncol 2016; 10: 31-39.
90. Dushyanthen S, Beavis PA, Savas P et al. Relevance of tumor-infiltrating lymphocytes in breast cancer. BMC Medicine 2015; 13: 202.
91. Kim R, Emi M, Tanabe K. Cancer immunoediting from immune surveillance to immune escape. Immunology 2007; 121: 1-14.
92. Vinay DS, Ryan EP, Pawelec G et al. Immune evasion in cancer: Mechanistic basis and therapeutic strategies. Semin Cancer Biol 2015; 35 Suppl: S185-s198.
93. Dunn GP, Bruce AT, Ikeda H et al. Cancer immunoediting: from immunosurveillance to tumor escape. Nature Immunology 2002; 3: 991-998.
94. Matsumoto H, Thike AA, Li H et al. Increased CD4 and CD8-positive T cell infiltrate signifies good prognosis in a subset of triple-negative breast cancer. Breast Cancer Research and Treatment 2016; 156: 237-247.
95. Wakabayashi O, Yamazaki K, Oizumi S et al. CD4+ T cells in cancer stroma, not CD8+ T cells in cancer cell nests, are associated with favorable prognosis in human non-small cell lung cancers. Cancer Sci 2003; 94: 1003-1009.
96. Vahidi Y, Bagheri M, Ghaderi A, Faghih Z. CD8-positive memory T cells in tumor-draining lymph nodes of patients with breast cancer. BMC Cancer 2020; 20: 257.
97. !!! INVALID CITATION !!! [189, 190].
98. Bailey SR, Nelson MH, Himes RA et al. Th17 cells in cancer: the ultimate identity crisis. Frontiers in immunology 2014; 5: 276-276.
99. Carvalho MI, Pires I, Prada J, Queiroga FL. A role for T-lymphocytes in human breast cancer and in canine mammary tumors. Biomed Res Int 2014; 2014: 130894.
100. Oda N, Shimazu K, Naoi Y et al. Intratumoral regulatory T cells as an independent predictive factor for pathological complete response to neoadjuvant paclitaxel followed by 5-FU/epirubicin/cyclophosphamide in breast cancer patients. Breast Cancer Res Treat 2012; 136: 107-116.
101. Baker K, Lachapelle J, Zlobec I et al. Prognostic significance of CD8+ T lymphocytes in breast cancer depends upon both oestrogen receptor status and histological grade. Histopathology 2011; 58: 1107-1116.
102. Liu S, Lachapelle J, Leung S et al. CD8+ lymphocyte infiltration is an independent favorable prognostic indicator in basal-like breast cancer. Breast cancer research : BCR 2012; 14: R48-R48.
103. Ali HR, Provenzano E, Dawson SJ et al. Association between CD8+ T-cell infiltration and breast cancer survival in 12,439 patients. Ann Oncol 2014; 25: 1536-1543.
104. West NR, Kost SE, Martin SD et al. Tumour-infiltrating FOXP3+ lymphocytes are associated with cytotoxic immune responses and good clinical outcome in oestrogen receptor-negative breast cancer. British Journal of Cancer 2013; 108: 155-162.
105. West NR, Kost SE, Martin SD et al. Tumour-infiltrating FOXP3(+) lymphocytes are associated with cytotoxic immune responses and good clinical outcome in oestrogen receptor-negative breast cancer. Br J Cancer 2013; 108: 155-162.
106. Mao Y, Qu Q, Chen X et al. The Prognostic Value of Tumor-Infiltrating Lymphocytes in Breast Cancer: A Systematic Review and Meta-Analysis. PLoS One 2016; 11: e0152500.
107. Liu S, Foulkes WD, Leung S et al. Prognostic significance of FOXP3+ tumor-infiltrating lymphocytes in breast cancer depends on estrogen receptor and human epidermal growth factor receptor-2 expression status and concurrent cytotoxic T-cell infiltration. Breast Cancer Research 2014; 16: 432.
108. Bates GJ, Fox SB, Han C et al. Quantification of regulatory T cells enables the identification of high-risk breast cancer patients and those at risk of late relapse. J Clin Oncol 2006; 24: 5373-5380.
109. Liu F, Lang R, Zhao J et al. $CD8^+$ cytotoxic T cell and $FOXP3^+$ regulatory T cell infiltration in relation to breast cancer survival and molecular subtypes. Breast Cancer Res Treat 2011; 130: 645-655.

110. Denkert C, Loibl S, Noske A et al. Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. J Clin Oncol 2010; 28: 105-113.
111. Hwang HW, Jung H, Hyeon J et al. A nomogram to predict pathologic complete response (pCR) and the value of tumor-infiltrating lymphocytes (TILs) for prediction of response to neoadjuvant chemotherapy (NAC) in breast cancer patients. Breast Cancer Res Treat 2019; 173: 255-266.
112. García-Teijido P, Cabal ML, Fernández IP, Pérez YF. Tumor-Infiltrating Lymphocytes in Triple Negative Breast Cancer: The Future of Immune Targeting. Clinical Medicine Insights. Oncology 2016; 10: 31-39.
113. Ahn SG, Jeong J, Hong S, Jung WH. Current Issues and Clinical Evidence in Tumor-Infiltrating Lymphocytes in Breast Cancer. Journal of pathology and translational medicine 2015; 49: 355-363.
114. Ladoire S, Arnould L, Apetoh L et al. Pathologic complete response to neoadjuvant chemotherapy of breast carcinoma is associated with the disappearance of tumor-infiltrating foxp3+ regulatory T cells. Clin Cancer Res 2008; 14: 2413-2420.
115. Reiman JM, Knutson KL, Radisky DC. Immune promotion of epithelial-mesenchymal transition and generation of breast cancer stem cells. Cancer Res 2010; 70: 3005-3008.
116. Tanei T, Morimoto K, Shimazu K et al. Association of breast cancer stem cells identified by aldehyde dehydrogenase 1 expression with resistance to sequential Paclitaxel and epirubicin-based chemotherapy for breast cancers. Clin Cancer Res 2009; 15: 4234-4241.
117. Singh A, Settleman J. EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. Oncogene 2010; 29: 4741-4751.
118. Creighton CJ, Li X, Landis M et al. Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. Proc Natl Acad Sci U S A 2009; 106: 13820-13825.
119. Sikandar B, Qureshi MA, Mirza T et al. Differential immune cell densities in ductal carcinoma In-Situ and invasive breast cancer: Possible role of leukocytes in early stages of carcinogenesis. Pak J Med Sci 2015; 31: 274-279.
120. Rakha EA, Aleskandarany M, El-Sayed ME et al. The prognostic significance of inflammation and medullary histological type in invasive carcinoma of the breast. Eur J Cancer 2009; 45: 1780-1787.
121. Lee AH, Happerfield LC, Millis RR, Bobrow LG. Inflammatory infiltrate in invasive lobular and ductal carcinoma of the breast. British journal of cancer 1996; 74: 796-801.
122. Lee AH, Happerfield LC, Millis RR, Bobrow LG. Inflammatory infiltrate in invasive lobular and ductal carcinoma of the breast. Br J Cancer 1996; 74: 796-801.
123. Beca F, Polyak K. Intratumor Heterogeneity in Breast Cancer. Adv Exp Med Biol 2016; 882: 169-189.
124. Januškevičienė I, Petrikaitė V. Heterogeneity of breast cancer: The importance of interaction between different tumor cell populations. Life Sci 2019; 239: 117009.
125. Dick JE. Stem cell concepts renew cancer research. Blood 2008; 112: 4793-4807.
126. Rakha EA, Ellis IO. An overview of assessment of prognostic and predictive factors in breast cancer needle core biopsy specimens. J Clin Pathol 2007; 60: 1300-1306.
127. Aleskandarany MA, Vandenberghe ME, Marchiò C et al. Tumour Heterogeneity of Breast Cancer: From Morphology to Personalised Medicine. Pathobiology 2018; 85: 23-34.
128. Ramón Y Cajal S, Sesé M, Capdevila C et al. Clinical implications of intratumor heterogeneity: challenges and opportunities. Journal of molecular medicine (Berlin, Germany) 2020; 98: 161-177.
129. Place AE, Jin Huh S, Polyak K. The microenvironment in breast cancer progression: biology and implications for treatment. Breast cancer research : BCR 2011; 13: 227-227.
130. Kim IS, Zhang XH. One microenvironment does not fit all: heterogeneity beyond cancer cells. Cancer Metastasis Rev 2016; 35: 601-629.
131. Arslan C, Sari E, Aksoy S, Altundag K. Variation in hormone receptor and HER-2 status between primary and metastatic breast cancer: review of the literature. Expert Opin Ther Targets 2011; 15: 21-30.

132. Cummings MC, Simpson PT, Reid LE et al. Metastatic progression of breast cancer: insights from 50 years of autopsies. J Pathol 2014; 232: 23-31.
133. Thomas P, Pranatharthi A, Ross C, Srivastava S. RhoC: a fascinating journey from a cytoskeletal organizer to a Cancer stem cell therapeutic target. Journal of Experimental & Clinical Cancer Research 2019; 38: 328.
134. Woodward WA, Hill RP. Cancer Stem Cells. Recent Results Cancer Res 2016; 198: 25-44.
135. Lin S, Fu S. The correlation of hematopoietic stem cells with cancer stem cells through the regulation of stromal cells in tumor microenvironment. Med Hypotheses 2013; 80: 494-497.
136. Visvader JE, Lindeman GJ. Cancer stem cells: current status and evolving complexities. Cell Stem Cell 2012; 10: 717-728.
137. Fessler E, Borovski T, Medema JP. Endothelial cells induce cancer stem cell features in differentiated glioblastoma cells via bFGF. Molecular cancer 2015; 14: 157-157.
138. Palomeras S, Ruiz-Martínez S, Puig T. Targeting Breast Cancer Stem Cells to Overcome Treatment Resistance. Molecules (Basel, Switzerland) 2018; 23: 2193.
139. Dick JE, Guenechea G, Gan OI, Dorrell C. In vivo dynamics of human stem cell repopulation in NOD/SCID mice. Ann N Y Acad Sci 2001; 938: 184-190.
140. Gupta PB, Fillmore CM, Jiang G et al. Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell 2011; 146: 633-644.
141. Geng S-Q, Alexandrou AT, Li JJ. Breast cancer stem cells: Multiple capacities in tumor metastasis. Cancer letters 2014; 349: 1-7.
142. Kim W-T, Ryu CJ. Cancer stem cell surface markers on normal stem cells. BMB reports 2017; 50: 285-298.
143. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nature Medicine 1997; 3: 730-737.
144. Goto H, Shimono Y, Funakoshi Y et al. Adipose-derived stem cells enhance human breast cancer growth and cancer stem cell-like properties through adipsin. Oncogene 2019; 38: 767-779.
145. Lee SY, Jeong EK, Ju MK et al. Induction of metastasis, cancer stem cell phenotype, and oncogenic metabolism in cancer cells by ionizing radiation. Mol Cancer 2017; 16: 10.
146. Eun K, Ham SW, Kim H. Cancer stem cell heterogeneity: origin and new perspectives on CSC targeting. BMB reports 2017; 50: 117-125.
147. Al-Hajj M, Wicha MS, Benito-Hernandez A et al. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A 2003; 100: 3983-3988.
148. Medema JP. Cancer stem cells: the challenges ahead. Nat Cell Biol 2013; 15: 338-344.
149. Ginestier C, Hur MH, Charafe-Jauffret E et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell stem cell 2007; 1: 555-567.
150. Chen K, Huang YH, Chen JL. Understanding and targeting cancer stem cells: therapeutic implications and challenges. Acta Pharmacol Sin 2013; 34: 732-740.
151. Matsui WH. Cancer stem cell signaling pathways. Medicine 2016; 95: S8-S19.
152. Riobo-Del Galdo NA, Lara Montero Á, Wertheimer EV. Role of Hedgehog Signaling in Breast Cancer: Pathogenesis and Therapeutics. Cells 2019; 8: 375.
153. Menzl I, Lebeau L, Pandey R et al. Loss of primary cilia occurs early in breast cancer development. Cilia 2014; 3: 7.
154. Di Mauro C, Rosa R, D'Amato V et al. Hedgehog signalling pathway orchestrates angiogenesis in triple-negative breast cancers. Br J Cancer 2017; 116: 1425-1435.
155. Bièche I, Lidereau R. Genetic alterations in breast cancer. Genes Chromosomes Cancer 1995; 14: 227-251.
156. O'Toole SA, Machalek DA, Shearer RF et al. Hedgehog overexpression is associated with stromal interactions and predicts for poor outcome in breast cancer. Cancer Res 2011; 71: 4002-4014.

157. ten Haaf A, Bektas N, von Serenyi S et al. Expression of the glioma-associated oncogene homolog (GLI) 1 in human breast cancer is associated with unfavourable overall survival. BMC Cancer 2009; 9: 298.
158. Jeng KS, Sheen IS, Jeng WJ et al. Activation of the sonic hedgehog signaling pathway occurs in the CD133 positive cells of mouse liver cancer Hepa 1-6 cells. Onco Targets Ther 2013; 6: 1047-1055.
159. Gu D, Liu H, Su GH et al. Combining hedgehog signaling inhibition with focal irradiation on reduction of pancreatic cancer metastasis. Mol Cancer Ther 2013; 12: 1038-1048.
160. Joost S, Almada LL, Rohnalter V et al. GLI1 inhibition promotes epithelial-to-mesenchymal transition in pancreatic cancer cells. Cancer Res 2012; 72: 88-99.
161. Buijs JT, Stayrook KR, Guise TA. The role of TGF-β in bone metastasis: novel therapeutic perspectives. Bonekey Rep 2012; 1: 96.
162. Takahashi T, Kawakami K, Mishima S et al. Cyclopamine induces eosinophilic differentiation and upregulates CD44 expression in myeloid leukemia cells. Leuk Res 2011; 35: 638-645.
163. Takebe N, Harris PJ, Warren RQ, Ivy SP. Targeting cancer stem cells by inhibiting Wnt, Notch, and Hedgehog pathways. Nat Rev Clin Oncol 2011; 8: 97-106.
164. Nusse R, Clevers H. Wnt/β-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. Cell 2017; 169: 985-999.
165. Clevers H, Nusse R. Wnt/β-catenin signaling and disease. Cell 2012; 149: 1192-1205.
166. Habas R, Dawid IB. Dishevelled and Wnt signaling: is the nucleus the final frontier? Journal of biology 2005; 4: 2-2.
167. Komiya Y, Habas R. Wnt signal transduction pathways. Organogenesis 2008; 4: 68-75.
168. Sun M, Zhao H, Xiao Q et al. Combined expression of aldehyde dehydrogenase 1A1 and β-catenin is associated with lymph node metastasis and poor survival in breast cancer patients following cyclophosphamide treatment. Oncol Rep 2015; 34: 3163-3173.
169. Jang GB, Kim JY, Cho SD et al. Blockade of Wnt/β-catenin signaling suppresses breast cancer metastasis by inhibiting CSC-like phenotype. Sci Rep 2015; 5: 12465.
170. Khramtsov AI, Khramtsova GF, Tretiakova M et al. Wnt/beta-catenin pathway activation is enriched in basal-like breast cancers and predicts poor outcome. Am J Pathol 2010; 176: 2911-2920.
171. DiMeo TA, Anderson K, Phadke P et al. A novel lung metastasis signature links Wnt signaling with cancer cell self-renewal and epithelial-mesenchymal transition in basal-like breast cancer. Cancer Res 2009; 69: 5364-5373.
172. Dey N, Barwick BG, Moreno CS et al. Wnt signaling in triple negative breast cancer is associated with metastasis. BMC Cancer 2013; 13: 537.
173. Shi W, Harris AL. Notch signaling in breast cancer and tumor angiogenesis: cross-talk and therapeutic potentials. J Mammary Gland Biol Neoplasia 2006; 11: 41-52.
174. Ju JH, Yang W, Oh S et al. HER2 stabilizes survivin while concomitantly down-regulating survivin gene transcription by suppressing Notch cleavage. Biochem J 2013; 451: 123-134.
175. Suman S, Das TP, Damodaran C. Silencing NOTCH signaling causes growth arrest in both breast cancer stem cells and breast cancer cells. British journal of cancer 2013; 109: 2587-2596.
176. Zhao D, Mo Y, Li MT et al. NOTCH-induced aldehyde dehydrogenase 1A1 deacetylation promotes breast cancer stem cells. J Clin Invest 2014; 124: 5453-5465.
177. Yuan X, Zhang M, Wu H et al. Expression of Notch1 Correlates with Breast Cancer Progression and Prognosis. PLoS One 2015; 10: e0131689.
178. Zhao J. Cancer stem cells and chemoresistance: The smartest survives the raid. Pharmacology & therapeutics 2016; 160: 145-158.
179. Sampieri K, Fodde R. Cancer stem cells and metastasis. Semin Cancer Biol 2012; 22: 187-193.

180. Dittmer J. Breast cancer stem cells: Features, key drivers and treatment options. Semin Cancer Biol 2018; 53: 59-74.
181. Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving paradigms. Cell 2011; 147: 275-292.
182. Fares J, Fares MY, Khachfe HH et al. Molecular principles of metastasis: a hallmark of cancer revisited. Signal Transduct Target Ther 2020; 5: 28.
183. !!! INVALID CITATION !!! [124, 125].
184. Ayob AZ, Ramasamy TS. Cancer stem cells as key drivers of tumour progression. Journal of biomedical science 2018; 25: 20-20.
185. Ayob AZ, Ramasamy TS. Cancer stem cells as key drivers of tumour progression. Journal of Biomedical Science 2018; 25: 20.
186. Ponta H, Sherman L, Herrlich PA. CD44: from adhesion molecules to signalling regulators. Nat Rev Mol Cell Biol 2003; 4: 33-45.
187. Cheng C, Yaffe MB, Sharp PA. A positive feedback loop couples Ras activation and CD44 alternative splicing. Genes Dev 2006; 20: 1715-1720.
188. Orian-Rousseau V, Chen L, Sleeman JP et al. CD44 is required for two consecutive steps in HGF/c-Met signaling. Genes & development 2002; 16: 3074-3086.
189. Sheridan C, Kishimoto H, Fuchs RK et al. CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. Breast Cancer Res 2006; 8: R59.
190. Balic M, Lin H, Young L et al. Most early disseminated cancer cells detected in bone marrow of breast cancer patients have a putative breast cancer stem cell phenotype. Clin Cancer Res 2006; 12: 5615-5621.
191. Charafe-Jauffret E, Ginestier C, Iovino F et al. Aldehyde dehydrogenase 1-positive cancer stem cells mediate metastasis and poor clinical outcome in inflammatory breast cancer. Clin Cancer Res 2010; 16: 45-55.
192. Marcato P, Dean CA, Pan D et al. Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. Stem Cells 2011; 29: 32-45.
193. Croker AK, Rodriguez-Torres M, Xia Y et al. Differential Functional Roles of ALDH1A1 and ALDH1A3 in Mediating Metastatic Behavior and Therapy Resistance of Human Breast Cancer Cells. Int J Mol Sci 2017; 18.
194. Li W, Ma H, Zhang J et al. Unraveling the roles of CD44/CD24 and ALDH1 as cancer stem cell markers in tumorigenesis and metastasis. Scientific reports 2017; 7: 13856-13856.
195. Luo M, Clouthier SG, Deol Y et al. Breast cancer stem cells: current advances and clinical implications. Methods Mol Biol 2015; 1293: 1-49.
196. Lagadec C, Vlashi E, Della Donna L et al. Survival and self-renewing capacity of breast cancer initiating cells during fractionated radiation treatment. Breast cancer research : BCR 2010; 12: R13-R13.
197. Phillips TM, McBride WH, Pajonk F. The response of CD24(-/low)/CD44+ breast cancer-initiating cells to radiation. J Natl Cancer Inst 2006; 98: 1777-1785.
198. Karimi-Busheri F, Rasouli-Nia A, Mackey JR, Weinfeld M. Senescence evasion by MCF-7 human breast tumor-initiating cells. Breast Cancer Res 2010; 12: R31.
199. !!! INVALID CITATION !!! [141].
200. Greaves M, Maley CC. Clonal evolution in cancer. Nature 2012; 481: 306-313.
201. Fujii H, Marsh C, Cairns P et al. Genetic divergence in the clonal evolution of breast cancer. Cancer Res 1996; 56: 1493-1497.
202. Polyak K, Shipitsin M, Campbell-Marrotta LL et al. Breast tumor heterogeneity: causes and consequences. Breast Cancer Research : BCR 2009; 11: S18-S18.
203. Merlo LMF, Pepper JW, Reid BJ, Maley CC. Cancer as an evolutionary and ecological process. Nature Reviews Cancer 2006; 6: 924-935.
204. Martelotto LG, Ng CK, Piscuoglio S et al. Breast cancer intra-tumor heterogeneity. Breast Cancer Res 2014; 16: 210.
205. Haque MM, Desai KV. Pathways to Endocrine Therapy Resistance in Breast Cancer. Front Endocrinol (Lausanne) 2019; 10: 573.

206. O'Leary B, Cutts RJ, Liu Y et al. The Genetic Landscape and Clonal Evolution of Breast Cancer Resistance to Palbociclib plus Fulvestrant in the PALOMA-3 Trial. Cancer Discov 2018; 8: 1390-1403.
207. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000; 100: 57-70.
208. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011; 144: 646-674.
209. Kwei KA, Kung Y, Salari K et al. Genomic instability in breast cancer: pathogenesis and clinical implications. Molecular oncology 2010; 4: 255-266.
210. Raynaud F, Mina M, Tavernari D, Ciriello G. Pan-cancer inference of intra-tumor heterogeneity reveals associations with different forms of genomic instability. PLoS Genet 2018; 14: e1007669.
211. Geyer FC, Weigelt B, Natrajan R et al. Molecular analysis reveals a genetic basis for the phenotypic diversity of metaplastic breast carcinomas. J Pathol 2010; 220: 562-573.
212. Navin N, Kendall J, Troge J et al. Tumour evolution inferred by single-cell sequencing. Nature 2011; 472: 90-94.
213. Goldhirsch A, Wood WC, Coates AS et al. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. Ann Oncol 2011; 22: 1736-1747.
214. Min Y, Bae SY, Lee H-C et al. Tubular carcinoma of the breast: clinicopathologic features and survival outcome compared with ductal carcinoma in situ. Journal of breast cancer 2013; 16: 404-409.
215. Ocaña A, Amir E, Pandiella A. HER2 heterogeneity and resistance to anti-HER2 antibody-drug conjugates. Breast Cancer Res 2020; 22: 15.
216. Brouckaert O, Paridaens R, Floris G et al. A critical review why assessment of steroid hormone receptors in breast cancer should be quantitative. Ann Oncol 2013; 24: 47-53.
217. Ahn S, Woo JW, Lee K, Park SY. HER2 status in breast cancer: changes in guidelines and complicating factors for interpretation. Journal of pathology and translational medicine 2020; 54: 34-44.
218. Seol H, Lee HJ, Choi Y et al. Intratumoral heterogeneity of HER2 gene amplification in breast cancer: its clinicopathological significance. Mod Pathol 2012; 25: 938-948.
219. Kuukasjärvi T, Karhu R, Tanner M et al. Genetic heterogeneity and clonal evolution underlying development of asynchronous metastasis in human breast cancer. Cancer Res 1997; 57: 1597-1604.
220. Schiavon G, Hrebien S, Garcia-Murillas I et al. Analysis of ESR1 mutation in circulating tumor DNA demonstrates evolution during therapy for metastatic breast cancer. Sci Transl Med 2015; 7: 313ra182.
221. Cottu PH, Asselah J, Lae M et al. Intratumoral heterogeneity of HER2/neu expression and its consequences for the management of advanced breast cancer. Ann Oncol 2008; 19: 595-597.
222. Wu JM, Halushka MK, Argani P. Intratumoral heterogeneity of HER-2 gene amplification and protein overexpression in breast cancer. Human pathology 2010; 41: 914-917.
223. Gerlinger M, Rowan AJ, Horswell S et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. The New England journal of medicine 2012; 366: 883-892.
224. Carvajal-Hausdorf DE, Schalper KA, Neumeister VM, Rimm DL. Quantitative measurement of cancer tissue biomarkers in the lab and in the clinic. Lab Invest 2015; 95: 385-396.
225. Zardavas D, Irrthum A, Swanton C, Piccart M. Clinical management of breast cancer heterogeneity. Nature Reviews Clinical Oncology 2015; 12: 381-394.
226. Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham Prognostic Index in primary breast cancer. Breast Cancer Res Treat 1992; 22: 207-219.
227. Habashy HO, Powe DG, Rakha EA et al. Forkhead-box A1 (FOXA1) expression in breast cancer and its prognostic significance. Eur J Cancer 2008; 44: 1541-1551.

228. Ahmed MA, Aleskandarany MA, Rakha EA et al. A $CD44^{-}/CD24^{+}$ phenotype is a poor prognostic marker in early invasive breast cancer. Breast Cancer Res Treat 2012; 133: 979-995.
229. Joseph C, Arshad M, Kurozomi S et al. Overexpression of the cancer stem cell marker CD133 confers a poor prognosis in invasive breast cancer. Breast Cancer Res Treat 2019; 174: 387-399.
230. Althobiti M, El Ansari R, Aleskandarany M et al. The prognostic significance of ALDH1A1 expression in early invasive breast cancer. Histopathology 2020.
231. Mahmoud SM, Paish EC, Powe DG et al. An evaluation of the clinical significance of FOXP3+ infiltrating cells in human breast cancer. Breast Cancer Res Treat 2011; 127: 99-108.
232. Mahmoud SM, Paish EC, Powe DG et al. Tumor-infiltrating CD8+ lymphocytes predict clinical outcome in breast cancer. J Clin Oncol 2011; 29: 1949-1955.
233. Mahmoud SM, Lee AH, Paish EC et al. Tumour-infiltrating macrophages and clinical outcome in breast cancer. J Clin Pathol 2012; 65: 159-163.
234. Mahmoud SM, Lee AH, Paish EC et al. The prognostic significance of B lymphocytes in invasive carcinoma of the breast. Breast Cancer Res Treat 2012; 132: 545-553.
235. Curtis C, Shah SP, Chin SF et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012; 486: 346-352.
236. McCarty KS, Jr., Miller LS, Cox EB et al. Estrogen receptor analyses. Correlation of biochemical and immunohistochemical methods using monoclonal antireceptor antibodies. Arch Pathol Lab Med 1985; 109: 716-721.
237. Wang F, Flanagan J, Su N et al. RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. J Mol Diagn 2012; 14: 22-29.
238. Grabinski TM, Kneynsberg A, Manfredsson FP, Kanaan NM. A method for combining RNAscope in situ hybridization with immunohistochemistry in thick free-floating brain sections and primary neuronal cultures. PLoS One 2015; 10: e0120120.
239. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin Cancer Res 2004; 10: 7252-7259.
240. Asano J, Hirakawa A, Hamada C. Assessing the prediction accuracy of cure in the Cox proportional hazards cure model: an application to breast cancer data. Pharm Stat 2014; 13: 357-363.
241. Jackson B, Brocker C, Thompson DC et al. Update on the aldehyde dehydrogenase gene (ALDH) superfamily. Human genomics 2011; 5: 283-303.
242. Sophos NA, Vasiliou V. Aldehyde dehydrogenase gene superfamily: the 2002 update. Chem Biol Interact 2003; 143-144: 5-22.
243. Marchitti SA, Brocker C, Stagos D, Vasiliou V. Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily. Expert opinion on drug metabolism & toxicology 2008; 4: 697-720.
244. Braun T, Bober E, Singh S et al. Evidence for a signal peptide at the amino-terminal end of human mitochondrial aldehyde dehydrogenase. FEBS Lett 1987; 215: 233-236.
245. Moreb JS, Ucar D, Han S et al. The enzymatic activity of human aldehyde dehydrogenases 1A2 and 2 (ALDH1A2 and ALDH2) is detected by Aldefluor, inhibited by diethylaminobenzaldehyde and has significant effects on cell proliferation and drug resistance. Chem Biol Interact 2012; 195: 52-60.
246. Vasiliou V, Nebert DW. Analysis and update of the human aldehyde dehydrogenase (ALDH) gene family. Human genomics 2005; 2: 138-143.
247. Marchitti SA, Brocker C, Stagos D, Vasiliou V. Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily. Expert Opin Drug Metab Toxicol 2008; 4: 697-720.
248. Hou Q, Bartels D. Comparative study of the aldehyde dehydrogenase (ALDH) gene superfamily in the glycophyte Arabidopsis thaliana and Eutrema halophytes. Annals of botany 2015; 115: 465-479.
249. Tomita H, Tanaka K, Tanaka T, Hara A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. Oncotarget 2016; 7: 11018-11032.

250. Vasiliou V, Thompson DC, Smith C et al. Aldehyde dehydrogenases: from eye crystallins to metabolic disease and cancer stem cells. Chemico-biological interactions 2013; 202: 2-10.
251. Tomita H, Tanaka K, Tanaka T, Hara A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. Oncotarget 2016; 7: 11018-11032.
252. Ross-Innes CS, Stark R, Holmes KA et al. Cooperative interaction between retinoic acid receptor-alpha and estrogen receptor in breast cancer. Genes Dev 2010; 24: 171-182.
253. Johansson HJ, Sanchez BC, Mundt F et al. Retinoic acid receptor alpha is associated with tamoxifen resistance in breast cancer. Nat Commun 2013; 4: 2175.
254. Vassalli G. Aldehyde Dehydrogenases: Not Just Markers, but Functional Regulators of Stem Cells. Stem cells international 2019; 2019: 3904645-3904645.
255. Vassalli G. Aldehyde Dehydrogenases: Not Just Markers, but Functional Regulators of Stem Cells. Stem Cells Int 2019; 2019: 3904645.
256. Cojoc M, Peitzsch C, Kurth I et al. Aldehyde Dehydrogenase Is Regulated by beta-Catenin/TCF and Promotes Radioresistance in Prostate Cancer Progenitor Cells. Cancer Res 2015; 75: 1482-1494.
257. Martinez-Cruzado L, Tornin J, Santos L et al. Aldh1 Expression and Activity Increase During Tumor Evolution in Sarcoma Cancer Stem Cell Populations. Sci Rep 2016; 6: 27878.
258. Nourbakhsh M, Farzaneh S, Taghikhani A et al. The Effect of a Newly Synthesized Ferrocene Derivative AGAINST MCF-7 Breast Cancer Cells and Spheroid Stem Cells through ROS Production and Inhibition of JAK2/STAT3 Signaling Pathway. Anticancer Agents Med Chem 2020.
259. Koenders STA, Wijaya LS, Erkelens MN et al. Development of a Retinal-Based Probe for the Profiling of Retinaldehyde Dehydrogenases in Cancer Cells. ACS Cent Sci 2019; 5: 1965-1974.
260. Ginestier C, Hur MH, Charafe-Jauffret E et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell 2007; 1: 555-567.
261. Yang W, Wang Y, Wang W et al. Expression of Aldehyde Dehydrogenase 1A1 (ALDH1A1) as a Prognostic Biomarker in Colorectal Cancer Using Immunohistochemistry. Medical science monitor : international medical journal of experimental and clinical research 2018; 24: 2864-2872.
262. Alamgeer M, Ganju V, Szczepny A et al. The prognostic significance of aldehyde dehydrogenase 1A1 (ALDH1A1) and CD133 expression in early stage non-small cell lung cancer. Thorax 2013; 68: 1095-1104.
263. Liu Y, Baglia M, Zheng Y et al. ALDH1A1 mRNA expression in association with prognosis of triple-negative breast cancer. Oncotarget 2015; 6: 41360-41369.
264. Croker AK, Rodriguez-Torres M, Xia Y et al. Differential Functional Roles of ALDH1A1 and ALDH1A3 in Mediating Metastatic Behavior and Therapy Resistance of Human Breast Cancer Cells. International journal of molecular sciences 2017; 18: 2039.
265. Ciccone V, Terzuoli E, Donnini S et al. Correction to: Stemness marker ALDH1A1 promotes tumor angiogenesis via retinoic acid/HIF-1alpha/VEGF signalling in MCF-7 breast cancer cells. J Exp Clin Cancer Res 2019; 38: 45.
266. Aghajani M, Mansoori B, Mohammadi A et al. New emerging roles of CD133 in cancer stem cell: Signaling pathway and miRNA regulation. J Cell Physiol 2019; 234: 21642-21661.
267. Aguilar-Medina M, Avendaño-Félix M, Lizárraga-Verdugo E et al. SOX9 Stem-Cell Factor: Clinical and Functional Relevance in Cancer. Journal of oncology 2019; 2019: 6754040-6754040.
268. Imrich S, Hachmeister M, Gires O. EpCAM and its potential role in tumor-initiating cells. Cell adhesion & migration 2012; 6: 30-38.
269. Jamidi SK, Hu J, Aphivatanasiri C et al. SOX10 as a sensitive marker for triple negative breast cancer. Histopathology 2020.
270. Li W, Ma H, Zhang J et al. Unraveling the roles of CD44/CD24 and ALDH1 as cancer stem cell markers in tumorigenesis and metastasis. Scientific Reports 2017; 7: 13856.

271. Sin WC, Lim CL. Breast cancer stem cells-from origins to targeted therapy. Stem cell investigation 2017; 4: 96-96.
272. Alamgeer M, Ganju V, Szczepny A et al. The prognostic significance of aldehyde dehydrogenase 1A1 (ALDH1A1) and CD133 expression in early stage non-small cell lung cancer. Thorax 2013; 68: 1095-1104.
273. Yang W, Wang Y, Wang W et al. Expression of Aldehyde Dehydrogenase 1A1 (ALDH1A1) as a Prognostic Biomarker in Colorectal Cancer Using Immunohistochemistry. Med Sci Monit 2018; 24: 2864-2872.
274. Qian X, Wagner S, Ma C et al. Prognostic significance of ALDH1A1-positive cancer stem cells in patients with locally advanced, metastasized head and neck squamous cell carcinoma. J Cancer Res Clin Oncol 2014; 140: 1151-1158.
275. van der Waals LM, Borel Rinkes IHM, Kranenburg O. ALDH1A1 expression is associated with poor differentiation, 'right-sidedness' and poor survival in human colorectal cancer. PLoS One 2018; 13: e0205536.
276. Li XS, Xu Q, Fu XY, Luo WS. ALDH1A1 overexpression is associated with the progression and prognosis in gastric cancer. BMC Cancer 2014; 14: 705.
277. Kong Y, Lyu N, Wu J et al. Breast cancer stem cell markers CD44 and ALDH1A1 in serum: distribution and prognostic value in patients with primary breast cancer. J Cancer 2018; 9: 3728-3735.
278. Chang B, Liu G, Xue F et al. ALDH1 expression correlates with favorable prognosis in ovarian cancers. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc 2009; 22: 817-823.
279. Wang L, Lin D, Fu Y, Lai M. Nuclear aldehyde dehydrogenase 1A1 (ALDH1A1) expression is a favorable prognostic indicator in colorectal carcinoma. Pathol Res Pract 2016; 212: 791-799.
280. Kida K, Ishikawa T, Yamada A et al. Effect of ALDH1 on prognosis and chemoresistance by breast cancer subtype. Breast Cancer Res Treat 2016; 156: 261-269.
281. Li H, Ma F, Wang H et al. Stem cell marker aldehyde dehydrogenase 1 (ALDH1)-expressing cells are enriched in triple-negative breast cancer. Int J Biol Markers 2013; 28: e357-364.
282. Opdenaker LM, Arnold KM, Pohlig RT et al. Immunohistochemical analysis of aldehyde dehydrogenase isoforms and their association with estrogen-receptor status and disease progression in breast cancer. Breast cancer (Dove Medical Press) 2014; 6: 205-209.
283. El Ansari R, Craze ML, Miligy I et al. The amino acid transporter SLC7A5 confers a poor prognosis in the highly proliferative breast cancer subtypes and is a key therapeutic target in luminal B tumours. Breast Cancer Research 2018; 20: 21.
284. Aleskandarany MA, Sonbul S, Surridge R et al. Rho-GTPase activating-protein 18: a biomarker associated with good prognosis in invasive breast cancer. British journal of cancer 2017; 117: 1176-1184.
285. Nalwoga H, Arnes JB, Wabinga H, Akslen LA. Expression of aldehyde dehydrogenase 1 (ALDH1) is associated with basal-like markers and features of aggressive tumours in African breast cancer. British journal of cancer 2010; 102: 369-375.
286. Li Y, Kong D, Ahmad A et al. Pancreatic cancer stem cells: emerging target for designing novel therapy. Cancer Lett 2013; 338: 94-100.
287. Toll AD, Dasgupta A, Potoczek M et al. Implications of enhancer of zeste homologue 2 expression in pancreatic ductal adenocarcinoma. Hum Pathol 2010; 41: 1205-1209.
288. Neumeister V, Agarwal S, Bordeaux J et al. In situ identification of putative cancer stem cells by multiplexing ALDH1, CD44, and cytokeratin identifies breast cancer patients with poor prognosis. Am J Pathol 2010; 176: 2131-2138.
289. Ma F, Li H, Li Y et al. Aldehyde dehydrogenase 1 (ALDH1) expression is an independent prognostic factor in triple negative breast cancer (TNBC). Medicine 2017; 96: e6561-e6561.
290. Waterworth A. Introducing the concept of breast cancer stem cells. Breast Cancer Research 2003; 6: 53.

291. Ahmed MA, Aleskandarany MA, Rakha EA et al. A CD44(-)/CD24(+) phenotype is a poor prognostic marker in early invasive breast cancer. Breast Cancer Res Treat 2012; 133: 979-995.
292. Horst D, Kriegl L, Engel J et al. CD133 expression is an independent prognostic marker for low survival in colorectal cancer. Br J Cancer 2008; 99: 1285-1289.
293. Gonzalez B, Denzel S, Mack B et al. EpCAM is involved in maintenance of the murine embryonic stem cell phenotype. Stem Cells 2009; 27: 1782-1791.
294. Grzegrzolka J, Biala M, Wojtyra P et al. Expression of EMT Markers SLUG and TWIST in Breast Cancer. Anticancer Res 2015; 35: 3961-3968.
295. Domenici G, Aurrekoetxea-Rodriguez I, Simoes BM et al. A Sox2-Sox9 signalling axis maintains human breast luminal progenitor and breast cancer stem cells. Oncogene 2019; 38: 3151-3169.
296. Chakravarty G, Moroz K, Makridakis NM et al. Prognostic significance of cytoplasmic SOX9 in invasive ductal carcinoma and metastatic breast cancer. Exp Biol Med (Maywood) 2011; 236: 145-155.
297. Yao J, Jin Q, Wang X-D et al. Aldehyde dehydrogenase 1 expression is correlated with poor prognosis in breast cancer. Medicine 2017; 96: e7171-e7171.
298. Inwald EC, Klinkhammer-Schalke M, Hofstädter F et al. Ki-67 is a prognostic parameter in breast cancer patients: results of a large population-based cohort of a cancer registry. Breast cancer research and treatment 2013; 139: 539-552.
299. Yao J, Jin Q, Wang XD et al. Aldehyde dehydrogenase 1 expression is correlated with poor prognosis in breast cancer. Medicine (Baltimore) 2017; 96: e7171.
300. Lee K-L, Kuo Y-C, Ho Y-S, Huang Y-H. Triple-Negative Breast Cancer: Current Understanding and Future Therapeutic Breakthrough Targeting Cancer Stemness. Cancers 2019; 11: 1334.
301. Mieog JS, de Kruijf EM, Bastiaannet E et al. Age determines the prognostic role of the cancer stem cell marker aldehyde dehydrogenase-1 in breast cancer. BMC Cancer 2012; 12: 42.
302. Liu Y, Baglia M, Zheng Y et al. ALDH1A1 mRNA expression in association with prognosis of triple-negative breast cancer. Oncotarget 2015; 6: 41360-41369.
303. Sjostrom M, Hartman L, Honeth G et al. Stem cell biomarker ALDH1A1 in breast cancer shows an association with prognosis and clinicopathological variables that is highly cut-off dependent. J Clin Pathol 2015; 68: 1012-1019.
304. Bednarz-Knoll N, Nastały P, Żaczek A et al. Stromal expression of ALDH1 in human breast carcinomas indicates reduced tumor progression. Oncotarget 2015; 6: 26789-26803.
305. M JR, S V. BMI1 and PTEN are key determinants of breast cancer therapy: A plausible therapeutic target in breast cancer. Gene 2018; 678: 302-311.
306. Paranjape AN, Balaji SA, Mandal T et al. Bmi1 regulates self-renewal and epithelial to mesenchymal transition in breast cancer cells through Nanog. BMC cancer 2014; 14: 785-785.
307. Alkema MJ, Wiegant J, Raap AK et al. Characterization and chromosomal localization of the human proto-oncogene BMI-1. Hum Mol Genet 1993; 2: 1597-1603.
308. Bhattacharya R, Mustafi SB, Street M et al. Bmi-1: At the crossroads of physiological and pathological biology. Genes Dis 2015; 2: 225-239.
309. Yang DD, Cui BB, Sun LY et al. The co-expression of USP22 and BMI-1 may promote cancer progression and predict therapy failure in gastric carcinoma. Cell Biochem Biophys 2011; 61: 703-710.
310. Yang G-F, He W-P, Cai M-Y et al. Intensive expression of Bmi-1 is a new independent predictor of poor outcome in patients with ovarian carcinoma. BMC cancer 2010; 10: 133-133.
311. Häyry V, Mäkinen LK, Atula T et al. Bmi-1 expression predicts prognosis in squamous cell carcinoma of the tongue. British journal of cancer 2010; 102: 892-897.
312. Guo B-H, Feng Y, Zhang R et al. Bmi-1 promotes invasion and metastasis, and its elevated expression is correlated with an advanced stage of breast cancer. Molecular cancer 2011; 10: 10-10.

313. Chang Z, Li Z, Wang X et al. Deciphering the mechanisms of tumorigenesis in human pancreatic ductal epithelial cells. Clinical cancer research : an official journal of the American Association for Cancer Research 2013; 19: 549-559.
314. Zhang X, Sun J, Wang H et al. IGF-1R and Bmi-1 expressions in lung adenocarcinoma and their clinicopathologic and prognostic significance. Tumour Biol 2014; 35: 739-745.
315. Li X, Yang Z, Song W et al. Overexpression of Bmi-1 contributes to the invasion and metastasis of hepatocellular carcinoma by increasing the expression of matrix metalloproteinase (MMP)2, MMP-9 and vascular endothelial growth factor via the PTEN/PI3K/Akt pathway. Int J Oncol 2013; 43: 793-802.
316. Engelsen IB, Mannelqvist M, Stefansson IM et al. Low BMI-1 expression is associated with an activated BMI-1-driven signature, vascular invasion, and hormone receptor loss in endometrial carcinoma. British journal of cancer 2008; 98: 1662-1669.
317. Liu PW, Lin Y, Chen XY. Expression of B-cell-specific Moloney murine leukemia virus integration site 1 mRNA and protein in gastric cancer. J Dig Dis 2014; 15: 166-173.
318. Meng X, Wang Y, Zheng X et al. shRNA-mediated knockdown of Bmi-1 inhibit lung adenocarcinoma cell migration and metastasis. Lung Cancer 2012; 77: 24-30.
319. Huang J, Qiu Y, Chen G et al. The relationship between Bmi-1 and the epithelial-mesenchymal transition in lung squamous cell carcinoma. Med Oncol 2012; 29: 1606-1613.
320. Colas E, Pedrola N, Devis L et al. The EMT signaling pathways in endometrial carcinoma. Clin Transl Oncol 2012; 14: 715-720.
321. Huber GF, Albinger-Hegyi A, Soltermann A et al. Expression patterns of Bmi-1 and p16 significantly correlate with overall, disease-specific, and recurrence-free survival in oropharyngeal squamous cell carcinoma. Cancer 2011; 117: 4659-4670.
322. Dimri GP, Martinez JL, Jacobs JJ et al. The Bmi-1 oncogene induces telomerase activity and immortalizes human mammary epithelial cells. Cancer Res 2002; 62: 4736-4745.
323. Joensuu K, Hagstrom J, Leidenius M et al. Bmi-1, c-myc, and Snail expression in primary breast cancers and their metastases--elevated Bmi-1 expression in late breast cancer relapses. Virchows Arch 2011; 459: 31-39.
324. Wang M-C, Li C-L, Cui J et al. BMI-1, a promising therapeutic target for human cancer. Oncology letters 2015; 10: 583-588.
325. Jacobs JJ, Scheijen B, Voncken JW et al. Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via INK4a/ARF. Genes & development 1999; 13: 2678-2690.
326. Guo WJ, Datta S, Band V, Dimri GP. Mel-18, a polycomb group protein, regulates cell proliferation and senescence via transcriptional repression of Bmi-1 and c-Myc oncoproteins. Mol Biol Cell 2007; 18: 536-546.
327. Cho J-H, Dimri M, Dimri GP. A positive feedback loop regulates the expression of polycomb group protein BMI1 via WNT signaling pathway. The Journal of biological chemistry 2013; 288: 3406-3418.
328. Gao C, Kong NR, Chai L. The role of stem cell factor SALL4 in leukemogenesis. Crit Rev Oncog 2011; 16: 117-127.
329. Yang J, Chai L, Liu F et al. Bmi-1 is a target gene for SALL4 in hematopoietic and leukemic cells. Proc Natl Acad Sci U S A 2007; 104: 10494-10499.
330. Zhu J, Wang H, Sun Q et al. Nrf2 is required to maintain the self-renewal of glioma stem cells. BMC Cancer 2013; 13: 380.
331. Song LB, Li J, Liao WT et al. The polycomb group protein Bmi-1 represses the tumor suppressor PTEN and induces epithelial-mesenchymal transition in human nasopharyngeal epithelial cells. J Clin Invest 2009; 119: 3626-3636.
332. Wu J, Hu D, Zhang R. Depletion of Bmi-1 enhances 5-fluorouracil-induced apoptosis and autophagy in hepatocellular carcinoma cells. Oncol Lett 2012; 4: 723-726.
333. Vlachostergios PJ, Papandreou CN. The Bmi-1/NF-kappaB/VEGF story: another hint for proteasome involvement in glioma angiogenesis? J Cell Commun Signal 2013; 7: 235-237.

334. Lee YC, Chang WW, Chen YY et al. Hsp90alpha Mediates BMI1 Expression in Breast Cancer Stem/Progenitor Cells through Facilitating Nuclear Translocation of c-Myc and EZH2. Int J Mol Sci 2017; 18.
335. Nalwoga H, Arnes JB, Wabinga H, Akslen LA. Expression of aldehyde dehydrogenase 1 (ALDH1) is associated with basal-like markers and features of aggressive tumours in African breast cancer. Br J Cancer 2010; 102: 369-375.
336. Wang H, Liu H, Li X et al. Estrogen receptor α-coupled Bmi1 regulation pathway in breast cancer and its clinical implications. BMC cancer 2014; 14: 122-122.
337. Li Y, Wang L, Liu J et al. O-GlcNAcylation modulates Bmi-1 protein stability and potential oncogenic function in prostate cancer. Oncogene 2017; 36: 6293-6305.
338. Li N, Jiang K, Fang LP et al. Knockdown of long noncoding RNA CCAT1 inhibits cell growth, invasion and peritoneal metastasis via downregulation of Bmi-1 in gastric cancer. Neoplasma 2018; 65: 736-744.
339. Kim JH, Yoon SY, Jeong SH et al. Overexpression of Bmi-1 oncoprotein correlates with axillary lymph node metastases in invasive ductal breast cancer. Breast 2004; 13: 383-388.
340. Januskeviciene I, Petrikaite V. Heterogeneity of breast cancer: The importance of interaction between different tumor cell populations. Life Sci 2019; 239: 117009.
341. Aleskandarany MA, Sonbul S, Surridge R et al. Rho-GTPase activating-protein 18: a biomarker associated with good prognosis in invasive breast cancer. Br J Cancer 2017; 117: 1176-1184.
342. Toss MS, Miligy IM, Gorringe KL et al. Prolyl-4-hydroxylase Alpha subunit 2 (P4HA2) expression is a predictor of poor outcome in breast ductal carcinoma in situ (DCIS). Br J Cancer 2018; 119: 1518-1526.
343. El Ansari R, Craze ML, Miligy I et al. The amino acid transporter SLC7A5 confers a poor prognosis in the highly proliferative breast cancer subtypes and is a key therapeutic target in luminal B tumours. Breast Cancer Res 2018; 20: 21.
344. Glinsky GV, Berezovska O, Glinskii AB. Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer. J Clin Invest 2005; 115: 1503-1521.
345. Song LB, Zeng MS, Liao WT et al. Bmi-1 is a novel molecular marker of nasopharyngeal carcinoma progression and immortalizes primary human nasopharyngeal epithelial cells. Cancer Res 2006; 66: 6225-6232.
346. Wang Y, Zhe H, Ding Z et al. Cancer stem cell marker Bmi-1 expression is associated with basal-like phenotype and poor survival in breast cancer. World J Surg 2012; 36: 1189-1194.
347. Choi YJ, Choi YL, Cho EY et al. Expression of Bmi-1 protein in tumor tissues is associated with favorable prognosis in breast cancer patients. Breast Cancer Research and Treatment 2009; 113: 83-93.
348. Choi YJ, Choi YL, Cho EY et al. Expression of Bmi-1 protein in tumor tissues is associated with favorable prognosis in breast cancer patients. Breast Cancer Res Treat 2009; 113: 83-93.
349. !!! INVALID CITATION !!! [40].
350. Lessard J, Sauvageau G. Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. Nature 2003; 423: 255-260.
351. Park IK, Qian D, Kiel M et al. Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. Nature 2003; 423: 302-305.
352. Abdouh M, Facchino S, Chatoo W et al. BMI1 sustains human glioblastoma multiforme stem cell renewal. J Neurosci 2009; 29: 8884-8896.
353. Akbar MW, Isbilen M, Belder N et al. A Stemness and EMT Based Gene Expression Signature Identifies Phenotypic Plasticity and is A Predictive but Not Prognostic Biomarker for Breast Cancer. Journal of Cancer 2020; 11: 949-961.
354. Margaryan NV, Hazard-Jenkins H, Salkeni MA et al. The Stem Cell Phenotype of Aggressive Breast Cancer Cells. Cancers (Basel) 2019; 11.
355. Dey P, Rathod M, De A. Targeting stem cells in the realm of drug-resistant breast cancer. Breast cancer (Dove Medical Press) 2019; 11: 115-135.

356. Mannello F. Understanding breast cancer stem cell heterogeneity: time to move on to a new research paradigm. BMC medicine 2013; 11: 169-169.
357. Schatton T, Frank NY, Frank MH. Identification and targeting of cancer stem cells. Bioessays 2009; 31: 1038-1049.
358. Bao B, Ahmad A, Azmi AS et al. Overview of cancer stem cells (CSCs) and mechanisms of their regulation: implications for cancer therapy. Current protocols in pharmacology 2013; Chapter 14: Unit-14.25.
359. Tang H, Song C, Ye F et al. miR-200c suppresses stemness and increases cellular sensitivity to trastuzumab in HER2+ breast cancer. Journal of cellular and molecular medicine 2019; 23: 8114-8127.
360. Palomeras S, Ruiz-Martinez S, Puig T. Targeting Breast Cancer Stem Cells to Overcome Treatment Resistance. Molecules 2018; 23.
361. Piva M, Domenici G, Iriondo O et al. Sox2 promotes tamoxifen resistance in breast cancer cells. EMBO Mol Med 2014; 6: 66-79.
362. Colacino JA, Azizi E, Brooks MD et al. Heterogeneity of Human Breast Stem and Progenitor Cells as Revealed by Transcriptional Profiling. Stem Cell Reports 2018; 10: 1596-1609.
363. Orlova Z, Pruefer F, Castro-Oropeza R et al. IKKepsilon regulates the breast cancer stem cell phenotype. Biochim Biophys Acta Mol Cell Res 2019; 1866: 598-611.
364. Almozyan S, Colak D, Mansour F et al. PD-L1 promotes OCT4 and Nanog expression in breast cancer stem cells by sustaining PI3K/AKT pathway activation. Int J Cancer 2017; 141: 1402-1412.
365. Gutierrez Diez PJ, Su Y, Russo J. Immunocytochemical stem cell markers can predict clinical stage of breast cancer. Oncol Rep 2017; 38: 1507-1516.
366. Li W, Ma H, Zhang J et al. Unraveling the roles of CD44/CD24 and ALDH1 as cancer stem cell markers in tumorigenesis and metastasis. Sci Rep 2017; 7: 13856.
367. Wang Z, Wang Q, Wang Q et al. Prognostic significance of CD24 and CD44 in breast cancer: a meta-analysis. Int J Biol Markers 2017; 32: e75-e82.
368. Li Z, Yin S, Zhang L et al. Clinicopathological characteristics and prognostic value of cancer stem cell marker CD133 in breast cancer: a meta-analysis. Onco Targets Ther 2017; 10: 859-870.
369. Shien T, Tanaka T, Tanabe M et al. Evaluation of ALDH1 expression in ipsilateral breast cancer recurrence. Oncol Lett 2017; 13: 1071-1077.
370. Ciccone V, Terzuoli E, Donnini S et al. Stemness marker ALDH1A1 promotes tumor angiogenesis via retinoic acid/HIF-1alpha/VEGF signalling in MCF-7 breast cancer cells. J Exp Clin Cancer Res 2018; 37: 311.
371. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med 1997; 3: 730-737.
372. Jang MH, Kang HJ, Jang KS et al. Clinicopathological analysis of CD44 and CD24 expression in invasive breast cancer. Oncol Lett 2016; 12: 2728-2733.
373. Rabinovich I, Sebastião APM, Lima RS et al. Cancer stem cell markers ALDH1 and CD44+/CD24- phenotype and their prognosis impact in invasive ductal carcinoma. European journal of histochemistry : EJH 2018; 62: 2943.
374. Spizzo G, Fong D, Wurm M et al. EpCAM expression in primary tumour tissues and metastases: an immunohistochemical analysis. J Clin Pathol 2011; 64: 415-420.
375. Kakarala M, Wicha MS. Implications of the cancer stem-cell hypothesis for breast cancer prevention and therapy. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2008; 26: 2813-2820.
376. Saeg F, Anbalagan M. Breast cancer stem cells and the challenges of eradication: a review of novel therapies. Stem cell investigation 2018; 5: 39-39.
377. Ghebeh H, Sleiman GM, Manogaran PS et al. Profiling of normal and malignant breast tissue show CD44high/CD24low phenotype as a predominant stem/progenitor marker when used in combination with Ep-CAM/CD49f markers. BMC Cancer 2013; 13: 289.

378. Pommier SJ, Quan GG, Christante D et al. Characterizing the HER2/neu status and metastatic potential of breast cancer stem/progenitor cells. Ann Surg Oncol 2010; 17: 613-623.
379. Idowu MO, Kmieciak M, Dumur C et al. 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 2012; 43: 364-373.
380. Agboola AJ, Paish EC, Rakha EA et al. EpCAM expression is an indicator of recurrence in basal-like breast cancer. Breast Cancer Res Treat 2012; 133: 575-582.
381. Jain RK, Mehta R, Dimitrov R et al. Atypical ductal hyperplasia: interobserver and intraobserver variability. Modern Pathology 2011; 24: 917-923.
382. Trzpis M, McLaughlin PMJ, de Leij LMFH, Harmsen MC. Epithelial cell adhesion molecule: more than a carcinoma marker and adhesion molecule. The American journal of pathology 2007; 171: 386-395.
383. Pulido C, Vendrell I, Ferreira AR et al. Bone metastasis risk factors in breast cancer. Ecancermedicalscience 2017; 11: 715-715.
384. Kennecke H, Yerushalmi R, Woods R et al. Metastatic behavior of breast cancer subtypes. J Clin Oncol 2010; 28: 3271-3277.
385. Yang F, Xu J, Tang L, Guan X. Breast cancer stem cell: the roles and therapeutic implications. Cell Mol Life Sci 2017; 74: 951-966.
386. Bai X, Ni J, Beretov J et al. Cancer stem cell in breast cancer therapeutic resistance. Cancer Treat Rev 2018; 69: 152-163.
387. Ghasemi F, Sarabi PZ, Athari SS, Esmaeilzadeh A. Therapeutics strategies against cancer stem cell in breast cancer. Int J Biochem Cell Biol 2019; 109: 76-81.
388. Leao R, Domingos C, Figueiredo A et al. Cancer Stem Cells in Prostate Cancer: Implications for Targeted Therapy. Urol Int 2017; 99: 125-136.
389. Jang MH, Kang HJ, Jang KS et al. Clinicopathological analysis of CD44 and CD24 expression in invasive breast cancer. Oncology letters 2016; 12: 2728-2733.
390. Yin B-b, Wu S-j, Zong H-j et al. Preliminary screening and identification of stem cell-like sphere clones in a gallbladder cancer cell line GBC-SD. Journal of Zhejiang University. Science. B 2011; 12: 256-263.
391. Das PK, Rakib MA, Khanam JA et al. Novel Therapeutics Against Breast Cancer Stem Cells by Targeting Surface Markers and Signaling Pathways. Curr Stem Cell Res Ther 2019; 14: 669-682.
392. Kozovska Z, Patsalias A, Bajzik V et al. ALDH1A inhibition sensitizes colon cancer cells to chemotherapy. BMC cancer 2018; 18: 656-656.
393. Tanaka K, Tomita H, Hisamatsu K et al. ALDH1A1-overexpressing cells are differentiated cells but not cancer stem or progenitor cells in human hepatocellular carcinoma. Oncotarget 2015; 6: 24722-24732.
394. Winter MJ, Nagelkerken B, Mertens AE et al. Expression of Ep-CAM shifts the state of cadherin-mediated adhesions from strong to weak. Exp Cell Res 2003; 285: 50-58.
395. Gaiser MR, Lammermann T, Feng X et al. Cancer-associated epithelial cell adhesion molecule (EpCAM; CD326) enables epidermal Langerhans cell motility and migration in vivo. Proc Natl Acad Sci U S A 2012; 109: E889-897.
396. Schnell U, Cirulli V, Giepmans BN. EpCAM: structure and function in health and disease. Biochim Biophys Acta 2013; 1828: 1989-2001.
397. Patriarca C, Macchi RM, Marschner AK, Mellstedt H. Epithelial cell adhesion molecule expression (CD326) in cancer: a short review. Cancer Treat Rev 2012; 38: 68-75.
398. Munz M, Baeuerle PA, Gires O. The emerging role of EpCAM in cancer and stem cell signaling. Cancer Res 2009; 69: 5627-5629.
399. Hiraga T, Ito S, Nakamura H. EpCAM expression in breast cancer cells is associated with enhanced bone metastasis formation. Int J Cancer 2016; 138: 1698-1708.
400. Osta WA, Chen Y, Mikhitarian K et al. EpCAM is overexpressed in breast cancer and is a potential target for breast cancer gene therapy. Cancer Res 2004; 64: 5818-5824.
401. Sudbeck P, Scherer G. Two independent nuclear localization signals are present in the DNA-binding high-mobility group domains of SRY and SOX9. J Biol Chem 1997; 272: 27848-27852.

402. Jana S, Madhu Krishna B, Singhal J et al. SOX9: The master regulator of cell fate in breast cancer. Biochem Pharmacol 2020; 174: 113789.
403. Kamachi Y, Kondoh H. Sox proteins: regulators of cell fate specification and differentiation. Development 2013; 140: 4129-4144.
404. She ZY, Yang WX. Nucleocytoplasmic shuttling of SOX14A and SOX14B transcription factors. Oncotarget 2017; 8: 46955-46968.
405. Jiang SS, Fang WT, Hou YH et al. Upregulation of SOX9 in lung adenocarcinoma and its involvement in the regulation of cell growth and tumorigenicity. Clin Cancer Res 2010; 16: 4363-4373.
406. Huang JQ, Wei FK, Xu XL et al. SOX9 drives the epithelial-mesenchymal transition in non-small-cell lung cancer through the Wnt/beta-catenin pathway. J Transl Med 2019; 17: 143.
407. Aguilar-Medina M, Avendaño-Félix M, Lizárraga-Verdugo E et al. SOX9 Stem-Cell Factor: Clinical and Functional Relevance in Cancer. J Oncol 2019; 2019: 6754040.
408. Fazilaty H, Gardaneh M, Akbari P et al. SLUG and SOX9 Cooperatively Regulate Tumor Initiating Niche Factors in Breast Cancer. Cancer Microenviron 2016; 9: 71-74.
409. Chakravarty G, Rider B, Mondal D. Cytoplasmic compartmentalization of SOX9 abrogates the growth arrest response of breast cancer cells that can be rescued by trichostatin A treatment. Cancer Biol Ther 2011; 11: 71-83.
410. Britton KM, Kirby JA, Lennard TWJ, Meeson AP. Cancer stem cells and side population cells in breast cancer and metastasis. Cancers 2011; 3: 2106-2130.
411. Tsyganov MM, Ibragimova MK, Pevzner AM et al. Amplification of Stem Genes: New Potential Metastatic Makers in Patients with an Early Form of Breast Cancer. Journal of Korean medical science 2019; 34: e312-e312.
412. Shah D, Osipo C. Cancer stem cells and HER2 positive breast cancer: The story so far. Genes & diseases 2016; 3: 114-123.
413. Osipo C, Patel P, Rizzo P et al. ErbB-2 inhibition activates Notch-1 and sensitizes breast cancer cells to a gamma-secretase inhibitor. Oncogene 2008; 27: 5019-5032.
414. Hallett RM, Kondratyev MK, Giacomelli AO et al. Small molecule antagonists of the Wnt/β-catenin signaling pathway target breast tumor-initiating cells in a Her2/Neu mouse model of breast cancer. PLoS One 2012; 7: e33976.
415. Cui W, Wang LH, Wen YY et al. Expression and regulation mechanisms of Sonic Hedgehog in breast cancer. Cancer Sci 2010; 101: 927-933.
416. Yoshida GJ, Saya H. Therapeutic strategies targeting cancer stem cells. Cancer science 2016; 107: 5-11.
417. Shima H, Yamada A, Ishikawa T, Endo I. Are breast cancer stem cells the key to resolving clinical issues in breast cancer therapy? Gland Surg 2017; 6: 82-88.
418. Hodi Z, Chakrabarti J, Lee AHS et al. The reliability of assessment of oestrogen receptor expression on needle core biopsy specimens of invasive carcinomas of the breast. Journal of clinical pathology 2007; 60: 299-302.
419. Lumachi F, Brunello A, Maruzzo M et al. Treatment of estrogen receptor-positive breast cancer. Curr Med Chem 2013; 20: 596-604.
420. Badve S, Vladislav IT, Spaulding B et al. EP1: a novel rabbit monoclonal antibody for detection of oestrogen receptor alpha. J Clin Pathol 2013; 66: 1051-1057.
421. Layfield LJ, Gupta D, Mooney EE. Assessment of Tissue Estrogen and Progesterone Receptor Levels: A Survey of Current Practice, Techniques, and Quantitation Methods. Breast J 2000; 6: 189-196.
422. Dowsett M, Cuzick J, Wale C et al. Retrospective analysis of time to recurrence in the ATAC trial according to hormone receptor status: an hypothesis-generating study. J Clin Oncol 2005; 23: 7512-7517.
423. Putti TC, Pinder SE, Elston CW et al. Breast pathology practice: most common problems in a consultation service. Histopathology 2005; 47: 445-457.
424. Rakha EA, Ellis IO. An overview of assessment of prognostic and predictive factors in breast cancer needle core biopsy specimens. Journal of clinical pathology 2007; 60: 1300-1306.

425. Hammond MEH, Hayes DF, Dowsett M et al. American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2010; 28: 2784-2795.
426. Prabhu JS, Korlimarla A, Desai K et al. A Majority of Low (1-10%) ER Positive Breast Cancers Behave Like Hormone Receptor Negative Tumors. J Cancer 2014; 5: 156-165.
427. Yi M, Huo L, Koenig KB et al. Which threshold for ER positivity? a retrospective study based on 9639 patients. Annals of oncology : official journal of the European Society for Medical Oncology 2014; 25: 1004-1011.
428. Chen T, Zhang N, Moran MS et al. Borderline ER-Positive Primary Breast Cancer Gains No Significant Survival Benefit From Endocrine Therapy: A Systematic Review and Meta-Analysis. Clin Breast Cancer 2018; 18: 1-8.
429. Iwamoto T, Booser D, Valero V et al. Estrogen receptor (ER) mRNA and ER-related gene expression in breast cancers that are 1% to 10% ER-positive by immunohistochemistry. J Clin Oncol 2012; 30: 729-734.
430. Allison KH, Hammond MEH, Dowsett M et al. Estrogen and Progesterone Receptor Testing in Breast Cancer: ASCO/CAP Guideline Update. J Clin Oncol 2020; 38: 1346-1366.
431. Hammond ME, Hayes DF, Dowsett M et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer (unabridged version). Arch Pathol Lab Med 2010; 134: e48-72.
432. Yaşar P, Ayaz G, User SD et al. Molecular mechanism of estrogen-estrogen receptor signaling. Reproductive medicine and biology 2016; 16: 4-20.
433. Freedman OC, Fletcher GG, Gandhi S et al. Adjuvant endocrine therapy for early breast cancer: a systematic review of the evidence for the 2014 Cancer Care Ontario systemic therapy guideline. Current oncology (Toronto, Ont.) 2015; 22: S95-S113.
434. Lim E, Metzger-Filho O, Winer EP. The natural history of hormone receptor-positive breast cancer. Oncology (Williston Park) 2012; 26: 688-694, 696.
435. Raghav KP, Hernandez-Aya LF, Lei X et al. Impact of low estrogen/progesterone receptor expression on survival outcomes in breast cancers previously classified as triple negative breast cancers. Cancer 2012; 118: 1498-1506.
436. Yi M, Huo L, Koenig KB et al. Which threshold for ER positivity? a retrospective study based on 9639 patients. Ann Oncol 2014; 25: 1004-1011.
437. Fujii T, Kogawa T, Dong W et al. Revisiting the definition of estrogen receptor positivity in HER2-negative primary breast cancer. Annals of oncology : official journal of the European Society for Medical Oncology 2017; 28: 2420-2428.
438. Petrau C, Clatot F, Cornic M et al. Reliability of Prognostic and Predictive Factors Evaluated by Needle Core Biopsies of Large Breast Invasive Tumors. Am J Clin Pathol 2015; 144: 555-562.
439. Munch-Petersen HD, Rasmussen BB, Balslev E. Reliability of histological malignancy grade, ER and HER2 status on core needle biopsy vs surgical specimen in breast cancer. Apmis 2014; 122: 750-754.
440. Dekker TJ, Smit VT, Hooijer GK et al. Reliability of core needle biopsy for determining ER and HER2 status in breast cancer. Ann Oncol 2013; 24: 931-937.
441. Collins LC, Botero ML, Schnitt SJ. Bimodal frequency distribution of estrogen receptor immunohistochemical staining results in breast cancer: an analysis of 825 cases. Am J Clin Pathol 2005; 123: 16-20.
442. Muftah AA, Aleskandarany M, Sonbul SN et al. Further evidence to support bimodality of oestrogen receptor expression in breast cancer. Histopathology 2017; 70: 456-465.
443. Cohen DA, Dabbs DJ, Cooper KL et al. Interobserver agreement among pathologists for semiquantitative hormone receptor scoring in breast carcinoma. Am J Clin Pathol 2012; 138: 796-802.

444. Rakha EA, Lee AH, Roberts J et al. Low-estrogen receptor-positive breast cancer: the impact of tissue sampling, choice of antibody, and molecular subtyping. J Clin Oncol 2012; 30: 2929-2930; author reply 2931.
445. Diorio C, Laberge S, Caron C et al. Validation of EP1 antibody clone for estrogen receptor immunohistochemistry in breast cancer. Appl Immunohistochem Mol Morphol 2014; 22: 613-618.
446. Yu X, Guo S, Song W et al. Estrogen receptor alpha (ERalpha) status evaluation using RNAscope in situ hybridization: a reliable and complementary method for IHC in breast cancer tissues. Hum Pathol 2017; 61: 121-129.
447. Nadji M, Gomez-Fernandez C, Ganjei-Azar P, Morales AR. Immunohistochemistry of estrogen and progesterone receptors reconsidered: experience with 5,993 breast cancers. Am J Clin Pathol 2005; 123: 21-27.
448. Arpino G, Weiss H, Lee AV et al. Estrogen receptor-positive, progesterone receptor-negative breast cancer: association with growth factor receptor expression and tamoxifen resistance. J Natl Cancer Inst 2005; 97: 1254-1261.
449. Horwitz KB, McGuire WL. Specific progesterone receptors in human breast cancer. Steroids 1975; 25: 497-505.
450. Yip CH, Rhodes A. Estrogen and progesterone receptors in breast cancer. Future Oncol 2014; 10: 2293-2301.
451. Szostakowska M, Trębińska-Stryjewska A, Grzybowska EA, Fabisiewicz A. Resistance to endocrine therapy in breast cancer: molecular mechanisms and future goals. Breast cancer research and treatment 2019; 173: 489-497.
452. Loi S, Haibe-Kains B, Desmedt C et al. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. BMC Genomics 2008; 9: 239.
453. Potts SJ, Krueger JS, Landis ND et al. Evaluating tumor heterogeneity in immunohistochemistry-stained breast cancer tissue. Lab Invest 2012; 92: 1342-1357.
454. Allison KH, Hammond MEH, Dowsett M et al. Estrogen and Progesterone Receptor Testing in Breast Cancer: ASCO/CAP Guideline Update. J Clin Oncol 2020; Jco1902309.
455. Turashvili G, Brogi E. Tumor Heterogeneity in Breast Cancer. Frontiers in medicine 2017; 4: 227-227.
456. Joseph C, Papadaki A, Althobiti M et al. Breast cancer intratumour heterogeneity: current status and clinical implications. Histopathology 2018; 73: 717-731.
457. Masuda S, Nitta H, Kelly BD et al. Intratumoral Estrogen Receptor Heterogeneity of Expression in Human Epidermal Growth Factor Receptor 2-Positive Breast Cancer as Evaluated by a Brightfield Multiplex Assay. J Histochem Cytochem 2019; 67: 563-574.
458. Inic Z, Zegarac M, Inic M et al. Difference between Luminal A and Luminal B Subtypes According to Ki-67, Tumor Size, and Progesterone Receptor Negativity Providing Prognostic Information. Clinical Medicine Insights. Oncology 2014; 8: 107-111.
459. Loi S, Haibe-Kains B, Desmedt C et al. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. BMC genomics 2008; 9: 239-239.
460. Fragomeni SM, Sciallis A, Jeruss JS. Molecular Subtypes and Local-Regional Control of Breast Cancer. Surgical oncology clinics of North America 2018; 27: 95-120.
461. Mohammed H, Russell IA, Stark R et al. Progesterone receptor modulates ERα action in breast cancer. Nature 2015; 523: 313-317.
462. Rakha EA, El-Sayed ME, Green AR et al. Biologic and clinical characteristics of breast cancer with single hormone receptor positive phenotype. J Clin Oncol 2007; 25: 4772-4778.
463. Kurozumi S, Matsumoto H, Hayashi Y et al. Power of PgR expression as a prognostic factor for ER-positive/HER2-negative breast cancer patients at intermediate risk classified by the Ki67 labeling index. BMC Cancer 2017; 17: 354.
464. Umekita Y, Souda M, Ohi Y et al. Expression of estrogen receptor alpha and progesterone receptor in normal human breast epithelium. In Vivo 2007; 21: 535-539.

465. Lindström LS, Yau C, Czene K et al. Intratumor Heterogeneity of the Estrogen Receptor and the Long-term Risk of Fatal Breast Cancer. Journal of the National Cancer Institute 2018; 110: 726-733.
466. Simpson EH. Measurement of Diversity. Nature 1949; 163: 688-688.
467. Park SY, Gönen M, Kim HJ et al. Cellular and genetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype. J Clin Invest 2010; 120: 636-644.
468. Martins FC, De S, Almendro V et al. Evolutionary pathways in BRCA1-associated breast tumors. Cancer Discov 2012; 2: 503-511.
469. Kucharavy A, Rubinstein B, Zhu J, Li R. Robustness and evolvability of heterogeneous cell populations. Molecular biology of the cell 2018; 29: 1400-1409.
470. Gay L, Baker A-M, Graham TA. Tumour Cell Heterogeneity. F1000Research 2016; 5: F1000 Faculty Rev-1238.
471. Althobiti M, Aleskandarany MA, Joseph C et al. Heterogeneity of tumour-infiltrating lymphocytes in breast cancer and its prognostic significance. Histopathology 2018; 73: 887-896.
472. Marusyk A, Polyak K. Tumor heterogeneity: causes and consequences. Biochim Biophys Acta 2010; 1805: 105-117.
473. Gerlinger M, Rowan AJ, Horswell S et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 2012; 366: 883-892.
474. Liu J, Dang H, Wang XW. The significance of intertumor and intratumor heterogeneity in liver cancer. Experimental & molecular medicine 2018; 50: e416-e416.
475. Michor F, Polyak K. The origins and implications of intratumor heterogeneity. Cancer prevention research (Philadelphia, Pa.) 2010; 3: 1361-1364.
476. Denison TA, Bae YH. Tumor heterogeneity and its implication for drug delivery. Journal of controlled release : official journal of the Controlled Release Society 2012; 164: 187-191.
477. Russnes HG, Navin N, Hicks J, Borresen-Dale AL. Insight into the heterogeneity of breast cancer through next-generation sequencing. J Clin Invest 2011; 121: 3810-3818.
478. Rossi L, Verrico M, Tomao S et al. Expression of ER, PgR, HER-2, and Ki-67 in core biopsies and in definitive histological specimens in patients with locally advanced breast cancer treated with neoadjuvant chemotherapy. Cancer Chemotherapy and Pharmacology 2020; 85: 105-111.
479. Rhodes A, Jasani B. The oestrogen receptor-negative/progesterone receptor-positive breast tumour: a biological entity or a technical artefact? J Clin Pathol 2009; 62: 95-96.
480. Harrell JC, Dye WW, Allred DC et al. Estrogen receptor positive breast cancer metastasis: altered hormonal sensitivity and tumor aggressiveness in lymphatic vessels and lymph nodes. Cancer Res 2006; 66: 9308-9315.
481. Zhao J, Hu C, Wang C et al. Breast cancer primary tumor ER expression pattern predicts its expression concordance in matched synchronous lymph node metastases. BMC Cancer 2018; 18: 1290.
482. Lindström LS, Yau C, Czene K et al. Intratumor Heterogeneity of the Estrogen Receptor and the Long-term Risk of Fatal Breast Cancer. J Natl Cancer Inst 2018; 110: 726-733.
483. Voduc D, Cheang M, Nielsen T. GATA-3 expression in breast cancer has a strong association with estrogen receptor but lacks independent prognostic value. Cancer Epidemiol Biomarkers Prev 2008; 17: 365-373.
484. Garcia-Becerra R, Santos N, Diaz L, Camacho J. Mechanisms of resistance to endocrine therapy in breast cancer: focus on signaling pathways, miRNAs and genetically based resistance. Int J Mol Sci 2012; 14: 108-145.
485. Kjallquist U, Erlandsson R, Tobin NP et al. Exome sequencing of primary breast cancers with paired metastatic lesions reveals metastasis-enriched mutations in the A-kinase anchoring protein family (AKAPs). BMC Cancer 2018; 18: 174.

486. Kumar B, Prasad M, Bhat-Nakshatri P et al. Normal Breast-Derived Epithelial Cells with Luminal and Intrinsic Subtype-Enriched Gene Expression Document Interindividual Differences in Their Differentiation Cascade. Cancer Res 2018; 78: 5107-5123.
487. Pereira B, Chin SF, Rueda OM et al. Erratum: The somatic mutation profiles of 2,433 breast cancers refine their genomic and transcriptomic landscapes. Nat Commun 2016; 7: 11908.
488. Curtis C, Shah SP, Chin S-F et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012; 486: 346-352.
489. Chaudhary S, Krishna BM, Mishra SK. A novel FOXA1/ESR1 interacting pathway: A study of Oncomine breast cancer microarrays. Oncol Lett 2017; 14: 1247-1264.
490. Haynes BP, Ginsburg O, Gao Q et al. Menstrual cycle associated changes in hormone-related gene expression in oestrogen receptor positive breast cancer. NPJ Breast Cancer 2019; 5: 42.
491. Andres SA, Brock GN, Wittliff JL. Interrogating differences in expression of targeted gene sets to predict breast cancer outcome. BMC Cancer 2013; 13: 326.
492. Tozlu S, Girault I, Vacher S et al. Identification of novel genes that co-cluster with estrogen receptor alpha in breast tumor biopsy specimens, using a large-scale real-time reverse transcription-PCR approach. Endocr Relat Cancer 2006; 13: 1109-1120.
493. Abd El-Rehim DM, Ball G, Pinder SE et al. High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. Int J Cancer 2005; 116: 340-350.
494. Grattan BJ, Freake HC. Zinc and cancer: implications for LIV-1 in breast cancer. Nutrients 2012; 4: 648-675.
495. Taylor KM, Morgan HE, Johnson A et al. Structure-function analysis of LIV-1, the breast cancer-associated protein that belongs to a new subfamily of zinc transporters. The Biochemical journal 2003; 375: 51-59.
496. Hogstrand C, Kille P, Ackland ML et al. A mechanism for epithelial-mesenchymal transition and anoikis resistance in breast cancer triggered by zinc channel ZIP6 and STAT3 (signal transducer and activator of transcription 3). Biochem J 2013; 455: 229-237.
497. Gaither LA, Eide DJ. Eukaryotic zinc transporters and their regulation. Biometals 2001; 14: 251-270.
498. McClelland RA, Manning DL, Gee JM et al. Oestrogen-regulated genes in breast cancer: association of pLIV1 with response to endocrine therapy. Br J Cancer 1998; 77: 1653-1656.
499. Taylor KM, Morgan HE, Smart K et al. The emerging role of the LIV-1 subfamily of zinc transporters in breast cancer. Molecular medicine (Cambridge, Mass.) 2007; 13: 396-406.
500. Shen H, Qin H, Guo J. Concordant correlation of LIV-1 and E-cadherin expression in human breast cancer cell MCF-7. Mol Biol Rep 2009; 36: 653-659.
501. Truong-Tran AQ, Carter J, Ruffin RE, Zalewski PD. The role of zinc in caspase activation and apoptotic cell death. Biometals 2001; 14: 315-330.
502. !!! INVALID CITATION !!! [9].
503. Hogstrand C, Kille P, Ackland ML et al. A mechanism for epithelial-mesenchymal transition and anoikis resistance in breast cancer triggered by zinc channel ZIP6 and STAT3 (signal transducer and activator of transcription 3). The Biochemical journal 2013; 455: 229-237.
504. Grattan BJ, Freake HC. Zinc and cancer: implications for LIV-1 in breast cancer. Nutrients 2012; 4: 648-675.
505. Taylor KM, Morgan HE, Smart K et al. The emerging role of the LIV-1 subfamily of zinc transporters in breast cancer. Mol Med 2007; 13: 396-406.
506. Lue HW, Yang X, Wang R et al. LIV-1 promotes prostate cancer epithelial-to-mesenchymal transition and metastasis through HB-EGF shedding and EGFR-mediated ERK signaling. PLoS One 2011; 6: e27720.

507. Unno J, Satoh K, Hirota M et al. LIV-1 enhances the aggressive phenotype through the induction of epithelial to mesenchymal transition in human pancreatic carcinoma cells. Int J Oncol 2009; 35: 813-821.
508. Zhao L, Chen W, Taylor KM et al. LIV-1 suppression inhibits HeLa cell invasion by targeting ERK1/2-Snail/Slug pathway. Biochem Biophys Res Commun 2007; 363: 82-88.
509. Shen R, Xie F, Shen H et al. Negative correlation of LIV-1 and E-cadherin expression in hepatocellular carcinoma cells. PLoS One 2013; 8: e56542.
510. Kasper G, Weiser AA, Rump A et al. Expression levels of the putative zinc transporter LIV-1 are associated with a better outcome of breast cancer patients. Int J Cancer 2005; 117: 961-973.
511. Taylor KM. A distinct role in breast cancer for two LIV-1 family zinc transporters. Biochem Soc Trans 2008; 36: 1247-1251.
512. Aleskandarany MA, Vandenberghe ME, Marchio C et al. Tumour Heterogeneity of Breast Cancer: From Morphology to Personalised Medicine. Pathobiology 2018; 85: 23-34.
513. Perou CM, Sorlie T, Eisen MB et al. Molecular portraits of human breast tumours. Nature 2000; 406: 747-752.
514. Sorlie T, Perou CM, Tibshirani R et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001; 98: 10869-10874.
515. McClelland RA, Manning DL, Gee JM et al. Oestrogen-regulated genes in breast cancer: association of pLIV1 with response to endocrine therapy. British journal of cancer 1998; 77: 1653-1656.
516. Taylor KM, Hiscox S, Nicholson RI. Zinc transporter LIV-1: a link between cellular development and cancer progression. Trends Endocrinol Metab 2004; 15: 461-463.
517. Lopez V, Kelleher SL. Zip6-attenuation promotes epithelial-to-mesenchymal transition in ductal breast tumor (T47D) cells. Exp Cell Res 2010; 316: 366-375.
518. Carroll JS. Mechanisms of oestrogen receptor (ER) gene regulation in breast cancer. European journal of endocrinology 2016; 175: R41-R49.
519. Rangel N, Fortunati N, Osella-Abate S et al. FOXA1 and AR in invasive breast cancer: new findings on their co-expression and impact on prognosis in ER-positive patients. BMC cancer 2018; 18: 703-703.
520. Theodorou V, Stark R, Menon S, Carroll JS. GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. Genome Res 2013; 23: 12-22.
521. Kouros-Mehr H, Slorach EM, Sternlicht MD, Werb Z. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. Cell 2006; 127: 1041-1055.
522. Corte MD, Tamargo F, Alvarez A et al. Cytosolic levels of TFF1/pS2 in breast cancer: Their relationship with clinical-pathological parameters and their prognostic significance. Breast Cancer Res Treat 2006; 96: 63-72.
523. Zheng ZY, Bay BH, Aw SE, Lin VC. A novel antiestrogenic mechanism in progesterone receptor-transfected breast cancer cells. J Biol Chem 2005; 280: 17480-17487.
524. Place AE, Jin Huh S, Polyak K. The microenvironment in breast cancer progression: biology and implications for treatment. Breast Cancer Res 2011; 13: 227.
525. Joyce JA, Pollard JW. Microenvironmental regulation of metastasis. Nat Rev Cancer 2009; 9: 239-252.
526. Tian T, Ruan M, Yang W, Shui R. Evaluation of the prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancers. Oncotarget 2016; 7: 44395-44405.
527. Hida AI, Sagara Y, Yotsumoto D et al. Prognostic and predictive impacts of tumor-infiltrating lymphocytes differ between Triple-negative and HER2-positive breast cancers treated with standard systemic therapies. Breast Cancer Res Treat 2016; 158: 1-9.
528. Loi S, Michiels S, Salgado R et al. Tumor infiltrating lymphocytes are prognostic in triple negative breast cancer and predictive for trastuzumab benefit in early breast cancer: results from the FinHER trial. Ann Oncol 2014; 25: 1544-1550.

529. Adams S, Gray RJ, Demaria S et al. Prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancers from two phase III randomized adjuvant breast cancer trials: ECOG 2197 and ECOG 1199. J Clin Oncol 2014; 32: 2959-2966.
530. Denkert C, Loibl S, Noske A et al. Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. J Clin Oncol 2010; 28: 105-113.
531. Ono M, Tsuda H, Shimizu C et al. Tumor-infiltrating lymphocytes are correlated with response to neoadjuvant chemotherapy in triple-negative breast cancer. Breast Cancer Res Treat 2012; 132: 793-805.
532. Issa-Nummer Y, Darb-Esfahani S, Loibl S et al. Prospective validation of immunological infiltrate for prediction of response to neoadjuvant chemotherapy in HER2-negative breast cancer--a substudy of the neoadjuvant GeparQuinto trial. PLoS One 2013; 8: e79775.
533. Salgado R, Denkert C, Demaria S et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. Ann Oncol 2015; 26: 259-271.
534. Heindl A, Sestak I, Naidoo K et al. Relevance of Spatial Heterogeneity of Immune Infiltration for Predicting Risk of Recurrence After Endocrine Therapy of ER+ Breast Cancer. Journal of the National Cancer Institute 2018; 110: 10.1093/jnci/djx1137.
535. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. Nat Med 2013; 19: 1423-1437.
536. Liu S, Lachapelle J, Leung S et al. CD8+ lymphocyte infiltration is an independent favorable prognostic indicator in basal-like breast cancer. Breast Cancer Res 2012; 14: R48.
537. Bailey SR, Nelson MH, Himes RA et al. Th17 cells in cancer: the ultimate identity crisis. Front Immunol 2014; 5: 276.
538. DeNardo DG, Coussens LM. Inflammation and breast cancer. Balancing immune response: crosstalk between adaptive and innate immune cells during breast cancer progression. Breast cancer research : BCR 2007; 9: 212-212.
539. Schoenfeld AJ, Rizvi H, Bandlamudi C et al. Clinical and molecular correlates of PD-L1 expression in patients with lung adenocarcinomas. Ann Oncol 2020.
540. Le Saux O, Dubois B, Stern MH et al. [Current advances in immunotherapy in ovarian cancer]. Bull Cancer 2020.
541. Giuffrida P, Arpa G, Grillo F et al. PD-L1 in small bowel adenocarcinoma is associated with etiology and tumor-infiltrating lymphocytes, in addition to microsatellite instability. Mod Pathol 2020.
542. Planes-Laine G, Rochigneux P, Bertucci F et al. PD-1/PD-L1 Targeting in Breast Cancer: The First Clinical Evidences Are Emerging. A Literature Review. Cancers 2019; 11: 1033.
543. Hida AI, Sagara Y, Yotsumoto D et al. Prognostic and predictive impacts of tumor-infiltrating lymphocytes differ between Triple-negative and HER2-positive breast cancers treated with standard systemic therapies. Breast Cancer Research and Treatment 2016; 158: 1-9.
544. Nanda R, Chow LQM, Dees EC et al. Pembrolizumab in Patients With Advanced Triple-Negative Breast Cancer: Phase Ib KEYNOTE-012 Study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2016; 34: 2460-2467.
545. Stanton SE, Disis ML. Clinical significance of tumor-infiltrating lymphocytes in breast cancer. J Immunother Cancer 2016; 4: 59.
546. Wang K, Xu J, Zhang T, Xue D. Tumor-infiltrating lymphocytes in breast cancer predict the response to chemotherapy and survival outcome: A meta-analysis. Oncotarget 2016; 7: 44288-44298.
547. Brown JR, Wimberly H, Lannin DR et al. Multiplexed Quantitative Analysis of CD3, CD8, and CD20 Predicts Response to Neoadjuvant Chemotherapy in Breast Cancer. Clinical cancer research : an official journal of the American Association for Cancer Research 2014; 20: 5995-6005.

548. Mani NL, Schalper KA, Hatzis C et al. Quantitative assessment of the spatial heterogeneity of tumor-infiltrating lymphocytes in breast cancer. Breast Cancer Res 2016; 18: 78.
549. Nawaz S, Heindl A, Koelble K, Yuan Y. Beyond immune density: critical role of spatial heterogeneity in estrogen receptor-negative breast cancer. Mod Pathol 2015; 28: 1621.
550. Loi S, Drubay D, Adams S et al. Tumor-Infiltrating Lymphocytes and Prognosis: A Pooled Individual Patient Analysis of Early-Stage Triple-Negative Breast Cancers. J Clin Oncol 2019; 37: 559-569.
551. Kurozumi S, Matsumoto H, Kurosumi M et al. Prognostic significance of tumour-infiltrating lymphocytes for oestrogen receptor-negative breast cancer without lymph node metastasis. Oncology letters 2019; 17: 2647-2656.
552. Ogiya R, Niikura N, Kumaki N et al. Comparison of tumor-infiltrating lymphocytes between primary and metastatic tumors in breast cancer patients. Cancer Science 2016; 107: 1730-1735.
553. Carbognin L, Pilotto S, Nortilli R et al. Predictive and Prognostic Role of Tumor-Infiltrating Lymphocytes for Early Breast Cancer According to Disease Subtypes: Sensitivity Analysis of Randomized Trials in Adjuvant and Neoadjuvant Setting. The oncologist 2016; 21: 283-291.
554. Oelkrug C, Ramage JM. Enhancement of T cell recruitment and infiltration into tumours. Clin Exp Immunol 2014; 178: 1-8.
555. Zhang Y, Gallastegui N, Rosenblatt JD. Regulatory B cells in anti-tumor immunity. Int Immunol 2015; 27: 521-530.
556. Buisseret L, Desmedt C, Garaud S et al. Reliability of tumor-infiltrating lymphocyte and tertiary lymphoid structure assessment in human breast cancer. Mod Pathol 2017; 30: 1204-1212.
557. Sikandar B, Qureshi MA, Naseem S et al. Increased Tumour Infiltration of CD4+ and CD8+ T-Lymphocytes in Patients with Triple Negative Breast Cancer Suggests Susceptibility to Immune Therapy. Asian Pac J Cancer Prev 2017; 18: 1827-1832.
558. Burstein MD, Tsimelzon A, Poage GM et al. Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer. Clin Cancer Res 2015; 21: 1688-1698.
559. Yeong J, Thike AA, Lim JC et al. Higher densities of Foxp3(+) regulatory T cells are associated with better prognosis in triple-negative breast cancer. Breast Cancer Res Treat 2017; 163: 21-35.
560. Medrek C, Ponten F, Jirstrom K, Leandersson K. The presence of tumor associated macrophages in tumor stroma as a prognostic marker for breast cancer patients. BMC Cancer 2012; 12: 306.
561. Yang J, Li X, Liu X, Liu Y. The role of tumor-associated macrophages in breast carcinoma invasion and metastasis. International Journal of Clinical and Experimental Pathology 2015; 8: 6656-6664.
562. Zhu B, Tse LA, Wang D et al. Immune gene expression profiling reveals heterogeneity in luminal breast tumors. Breast Cancer Research 2019; 21: 147.
563. Dieci MV, Prat A, Tagliafico E et al. Integrated evaluation of PAM50 subtypes and immune modulation of pCR in HER2-positive breast cancer patients treated with chemotherapy and HER2-targeted agents in the CherLOB trial. Ann Oncol 2016; 27: 1867-1873.
564. Noske A, Mobus V, Weber K et al. Relevance of tumour-infiltrating lymphocytes, PD-1 and PD-L1 in patients with high-risk, nodal-metastasised breast cancer of the German Adjuvant Intergroup Node-positive study. Eur J Cancer 2019; 114: 76-88.
565. Loi S, Sirtaine N, Piette F et al. Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. J Clin Oncol 2013; 31: 860-867.
566. Kurozumi S, Inoue K, Matsumoto H et al. Clinicopathological values of PD-L1 expression in HER2-positive breast cancer. Scientific reports 2019; 9: 16662-16662.

567. Enwere EK, Kornaga EN, Dean M et al. Expression of PD-L1 and presence of CD8-positive T cells in pre-treatment specimens of locally advanced cervical cancer. Modern Pathology 2017; 30: 577-586.
568. Yang F, Cao L, Sun Z et al. Evaluation of Breast Cancer Stem Cells and Intratumor Stemness Heterogeneity in Triple-negative Breast Cancer as Prognostic Factors. International journal of biological sciences 2016; 12: 1568-1577.
569. Al-Thoubaity FK. Molecular classification of breast cancer: A retrospective cohort study. Annals of medicine and surgery (2012) 2019; 49: 44-48.
570. Nguyen TH, Nguyen VH, Nguyen TL et al. Evaluations of Biomarker Status Changes between Primary and Recurrent Tumor Tissue Samples in Breast Cancer Patients. Biomed Res Int 2019; 2019: 7391237.
571. Gerber B, Freund M, Reimer T. Recurrent breast cancer: treatment strategies for maintaining and prolonging good quality of life. Dtsch Arztebl Int 2010; 107: 85-91.
572. Marusyk A, Polyak K. Tumor heterogeneity: causes and consequences. Biochimica et biophysica acta 2010; 1805: 105-117.
573. Caiado F, Silva-Santos B, Norell H. Intra-tumour heterogeneity - going beyond genetics. Febs j 2016; 283: 2245-2258.
574. Assenov Y, Brocks D, Gerhäuser C. Intratumor heterogeneity in epigenetic patterns. Semin Cancer Biol 2018; 51: 12-21.
575. Roulot A, Héquet D, Guinebretière JM et al. Tumoral heterogeneity of breast cancer. Ann Biol Clin (Paris) 2016; 74: 653-660.
576. Rahman M, Deleyrolle L, Vedam-Mai V et al. The cancer stem cell hypothesis: failures and pitfalls. Neurosurgery 2011; 68: 531-545; discussion 545.
577. Martelotto LG, Ng CKY, Piscuoglio S et al. Breast cancer intra-tumor heterogeneity. Breast cancer research : BCR 2014; 16: 210-210.
578. Bertos NR, Park M. Breast cancer - one term, many entities? The Journal of clinical investigation 2011; 121: 3789-3796.
579. Hwang-Verslues WW, Lee W-H, Lee EYHP. Biomarkers to Target Heterogeneous Breast Cancer Stem Cells. Journal of molecular biomarkers & diagnosis 2012; Suppl 8: 6-6.
580. Feng Y, Spezia M, Huang S et al. Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. Genes Dis 2018; 5: 77-106.
581. Sousa B, Ribeiro AS, Paredes J. Heterogeneity and Plasticity of Breast Cancer Stem Cells. Adv Exp Med Biol 2019; 1139: 83-103.
582. Palomeras S, Ruiz-Martínez S, Puig T. Targeting Breast Cancer Stem Cells to Overcome Treatment Resistance. Molecules 2018; 23.
583. O'Conor CJ, Chen T, González I et al. Cancer stem cells in triple-negative breast cancer: a potential target and prognostic marker. Biomark Med 2018; 12: 813-820.
584. Shipitsin M, Campbell LL, Argani P et al. Molecular definition of breast tumor heterogeneity. Cancer Cell 2007; 11: 259-273.
585. Wright MH, Calcagno AM, Salcido CD et al. Brca1 breast tumors contain distinct CD44+/CD24- and CD133+ cells with cancer stem cell characteristics. Breast Cancer Res 2008; 10: R10.
586. Ye F, Qiu Y, Li L et al. The Presence of EpCAM(-)/CD49f(+) Cells in Breast Cancer Is Associated with a Poor Clinical Outcome. Journal of breast cancer 2015; 18: 242-248.
587. Adam SA, Schnell O, Pöschl J et al. ALDH1A1 is a marker of astrocytic differentiation during brain development and correlates with better survival in glioblastoma patients. Brain Pathol 2012; 22: 788-797.
588. Patel M, Lu L, Zander DS et al. ALDH1A1 and ALDH3A1 expression in lung cancers: correlation with histologic type and potential precursors. Lung Cancer 2008; 59: 340-349.
589. Yu C-C, Lo W-L, Chen Y-W et al. Bmi-1 Regulates Snail Expression and Promotes Metastasis Ability in Head and Neck Squamous Cancer-Derived ALDH1 Positive Cells. Journal of oncology 2011; 2011: 609259.
590. de Beça FF, Caetano P, Gerhard R et al. Cancer stem cells markers CD44, CD24 and ALDH1 in breast cancer special histological types. J Clin Pathol 2013; 66: 187-191.

591. Liu Y, Nenutil R, Appleyard MV et al. Lack of correlation of stem cell markers in breast cancer stem cells. British journal of cancer 2014; 110: 2063-2071.
592. Yersal O, Barutca S. Biological subtypes of breast cancer: Prognostic and therapeutic implications. World journal of clinical oncology 2014; 5: 412-424.
593. Kasper G, Weiser AA, Rump A et al. Expression levels of the putative zinc transporter LIV-1 are associated with a better outcome of breast cancer patients. Int J Cancer 2005; 117: 961-973.
594. Manning DL, Robertson JF, Ellis IO et al. Oestrogen-regulated genes in breast cancer: association of pLIV1 with lymph node involvement. Eur J Cancer 1994; 30a: 675-678.
595. Salgado R, Denkert C, Demaria S et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. Annals of oncology : official journal of the European Society for Medical Oncology 2015; 26: 259-271.
596. Zhang D, Salto-Tellez M, Putti TC et al. Reliability of Tissue Microarrays in Detecting Protein Expression and Gene Amplification in Breast Cancer. Modern Pathology 2003; 16: 79-85.
597. Buckley NE, Forde C, McArt DG et al. Quantification of HER2 heterogeneity in breast cancer-implications for identification of sub-dominant clones for personalised treatment. Scientific reports 2016; 6: 23383-23383.
598. Abe N, Matsumoto H, Takamatsu R et al. Quantitative digital image analysis of tumor-infiltrating lymphocytes in HER2-positive breast cancer. Virchows Arch 2020; 476: 701-709.